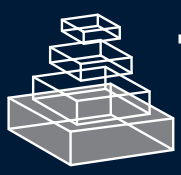
frontiers

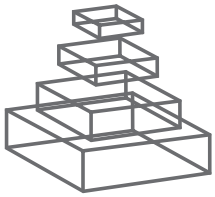
RESEARCH TOPICS

NEUROECONOMICS

Hosted by
Paul E. Phillips, Jeansok J. Kim and
Daeyeol Lee



frontiers in
BEHAVIORAL NEUROSCIENCE



frontiers

Cover image provided by Ibbl sarl, Lausanne CH

ISSN 1664-8714
ISBN 978-2-88919-027-0
DOI 10.3389/978-2-88919-027-0

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.
All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!
Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

NEUROECONOMICS

Hosted By
Paul E. Phillips, University of Washington, USA
Jeansok J. Kim, University of Washington, USA
Daeyeol Lee, Yale University School of Medicine, USA

The brain is the organ of decision making, and evolved to maximize the outcomes desirable for the survival and reproduction of the species. Compared to the sensory and motor functions of the brain, however, the process of decision making is less accessible to direct experimental manipulations and hence requires more careful theoretical analyses. Indeed, the principle of optimality and frequent departures of human behaviors from those predicted for optimal rational decision makers have long been studied experimentally as well as theoretically. However, it is only recently that neurobiological studies of decision making started exploiting the framework previously developed in economics and psychology systematically. This cross-disciplinary research program, known as neuroeconomics, has already been enormously successful. Increasingly, neuroscientists benefit from the studies of utility theory, game theory, prospect theory, and reinforcement learning theory, and contribute to the refinement of such theories by providing more relevant empirical data. This Research Topic will showcase the recent advances in neuroeconomics that combine economic and behavioral analyses in neurobiological studies of value-based decision making.

Table of Contents

frontiers in
BEHAVIORAL NEUROSCIENCE

EDITORIAL
published: 30 March 2012
doi: 10.3389/fnbeh.2012.00015

Neuroeconomics

Paul E. M. Phillips[1], Jeansok J. Kim[2] and Daeyeol Lee[3]*

[1] *Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA*
[2] *Psychology, University of Washington, Seattle, WA, USA*
[3] *Neurobiology, Yale University School of Medicine, New Haven, CT, USA*
**Correspondence: daeyeol.lee@yale.edu*

The neural process of decision making is not easily accessible to direct experimental manipulations and hence requires careful theoretical analyses. Nevertheless, neurobiological studies of decision making started exploiting the frameworks previously developed in economics and psychology. This cross-disciplinary research program, known as neuroeconomics, has already been enormously successful.

The articles included in this E-book are eclectic samples of the latest research on various topics in neuroeconomics. Many of these papers address the issues related to, utility functions. For example, Heldman et al. (2009) demonstrated that a combination of electroencephalographic recording and economic choice paradigms can be a useful research tool to examine the neural basis of utility function. Hwang et al. (2009) provided quantitative characterization of temporal discount function in non-human primates, whereas Luhman (2009) provides a review of behavioral and neurobiological studies on temporal discounting and intertemporal choice. Activity in the posterior parietal cortex, such as the lateral intraparietal cortex (LIP), encodes the signals related to utilities during decision making, and Pearson et al. (2010) provides a parsimonious model for the numerical coding in the LIP. Carter et al. (2009) have used functional neuroimaging to examine the neural activity related to monetary gains and losses directed to self and charity. Salamone et al. (2009) discuss the role of the dopamine neurons in effort-related choice behavior, whereas Delgado et al. (2009) demonstrated that the interaction between amygdala and striatum plays an important role during avoidance learning in human subjects.

Other papers in this collection focus on the neural basis of social decision making. Thevarajah et al. (2010) demonstrated that during a computer-simulated matching-pennies game, the behaviors of monkeys and the neural activity in the superior colliculus were parsimoniously accounted for by a hybrid learning model. Chang and Sanfey (2009) showed that the memory of partners encountered previously during an ultimatum game and the associated neural activity in many different brain areas were enhanced when the previous offers from these partners were contradictory to the initial expectations. Seymour et al. (2009) proposed that a combination of reinforcement learning and observation learning might give rise to altruistic behavior. Aragona and Wang (2009) showed that the prairie voles' monogamous pair bonding might be a model system for understanding the role of reward and hedonic mechanisms underlying social decision making. Finally, Kato et al. (2009) investigated the neural mechanisms mediating the effects of positive and negative advertisement on political preferences. This excellent collection of articles demonstrate that the insights from neuroeconomic studies have the potential to shed light on many topics in humanities and social sciences.

REFERENCES

Aragona, B. J., and Wang, Z. (2009). Dopamine regulation of social choice in a monogamous rodent species. *Front. Behav. Neurosci.* 3:15. doi: 10.3389/neuro.08.015.2009

Carter, R. M., MacInnes, J. J., Huettel, S. A., and Adcock, R. A. (2009). Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses. *Front. Behav. Neurosci.* 3:21. doi: 10.3389/neuro.08.021.2009

Chang, L. J., and Sanfey, A. G. (2009). Unforgettable ultimatums? Expectation violations promote enhanced social memory following economic bargaining. *Front. Behav. Neurosci.* 3:36. doi: 10.3389/neuro.08.036.2009

Delgado, M. R., Jou, R. L., LeDoux, J. E., and Phelps, E. A. (2009). Avoiding negative outcomes: tracking the mechanisms of avoidance learning in humans during fear conditioning. *Front. Behav. Neurosci.* 3:33. doi: 10.3389/neuro.08.033.2009

Heldmann, M.,Vogt, B., Heinze, H. -J., and Münte, T. F. (2009). Different methods to define utility functions yield similar results but engage different neural processes. *Front. Behav. Neurosci.* 3:43. doi: 10.3389/neuro.08.043.2009

Hwang, J., Kim, S., and Lee, D. (2009). Temporal discounting and inter-temporal choice in rhesus monkeys. *Front. Behav. Neurosci.* 3:9. doi:10.3389/neuro.08.009.2009.

Kato, J., Ide, H., Kabashima, I., Kadota, H., Takano, K., and Kansaku, K. (2009). Neural correlates of attitude change following positive and negative advertisements. *Front. Behav. Neurosci.* 3:6. doi:10.3389/neuro.08.006.2009

Luhmann, C. C. (2009). Temporal decision-making: insights from cognitive neuroscience. *Front. Behav. Neurosci.* 3:39. doi: 10.3389/neuro.08.039.2009

Pearson, J., Roitman, J. D., Brannon, E. M., Platt, M. L., and Raghavachari, S. (2010). A physiologically-inspired model of numerical classification based on graded stimulus coding. *Front. Behav. Neurosci.* 4:1. doi: 10.3389/neuro.08.001.2010

Salamone, J. D., Correa, M., Farrar, A. M., Nunes, E. J., and Pardo, M. (2009). Dopamine, behavioral economics, and effort. *Front. Behav. Neurosci.* 3:13. doi: 10.3389/neuro.08.013.2009

Seymour, B., Yoshida, W., and Dolan, R. (2009). Altruistic learning. *Front. Behav. Neurosci.* 3:23. doi: 10.3389/neuro.08.023.2009

Thevarajah, D., Webb, R., Ferrall, C., and Dorris, M. C. (2010). Modeling the value of strategic actions in the superior colliculus. *Front. Behav. Neurosci.* 3:57. doi: 10.3389/neuro.08.057.2009

Received: 07 March 2012; accepted: 11 March 2012; published online: 30 March 2012.
Citation: Phillips PEM, Kim JJ and Lee D (2012) Neuroeconomics. Front. Behav. Neurosci. ***6****:15. doi: 10.3389/fnbeh.2012.00015*

frontiers in
BEHAVIORAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 30 October 2009
doi: 10.3389/neuro.08.043.2009

Different methods to define utility functions yield similar results but engage different neural processes

Marcus Heldmann[1]*, Bodo Vogt[2], Hans-Jochen Heinze[1,3] and Thomas F. Münte[3,4]

[1] *Department of Neurology, University Magdeburg, Germany*
[2] *Faculty of Economics, Empirical Economics, University of Magdeburg, Magdeburg, Germany*
[3] *Center for Behavioral Brain Science, Magdeburg, Germany*
[4] *Department of Neuropsychology, University of Magdeburg, Magdeburg, Germany*

Edited by:
Daeyeol Lee, Yale University School of Medicine, USA

Reviewed by:
Mauricio R. Delgado, Rutgers University, USA
Joseph Kable, University of Pennsylvania, USA

****Correspondence:***
Marcus Heldmann, Neurology Department, Otto-von-Guericke University Magdeburg, Leipziger Strasse 44, 39120 Magdeburg, Germany.
e-mail: marcus.heldmann@med.ovgu.de

Although the concept of utility is fundamental to many economic theories, up to now a generally accepted method determining a subject's utility function is not available. We investigated two methods that are used in economic sciences for describing utility functions by using response-locked event-related potentials in order to assess their neural underpinnings. For determining the certainty equivalent, we used a lottery game with probabilities to win $p = 0.5$, for identifying the subjects' utility functions directly a standard bisection task was applied. Although the lottery tasks' payoffs were only hypothetical, a pronounced negativity was observed resembling the error related negativity (ERN) previously described in action monitoring research, but this occurred only for choices far away from the indifference point between money and lottery. By contrast, the bisection task failed to evoke an remarkable ERN irrespective of the responses' correctness. Based on these findings we are reasoning that only decisions made in the lottery task achieved a level of subjective relevance that activates cognitive-emotional monitoring. In terms of economic sciences, our findings support the view that the bisection method is unaffected by any kind of probability valuation or other parameters related to risk and in combination with the lottery task can, therefore, be used to differentiate between payoff and probability valuation.

Keywords: utility function, neuroeconomics, error-related negativity, executive functions, cognitive electrophysiology, lottery, bisection

INTRODUCTION

The concept of utility functions is fundamental to economics. Utility features prominent in most economic theories, as it is not a good's quantity or money *per se* that determines the actions of human beings, so called agents, but the utility they obtain from the good. Equilibrium concepts like the Nash equilibrium for strategic interactions of agents, the Walrasian equilibrium of economies, financial theory or the theory of political decision making are based on utility considerations. Expected utility theory and its modifications like Prospect Theory (Kahneman and Tversky, 1979) are the most established theories for decision making under risk. Utility theory is well-founded by an axiomatic approach with few intuitive axioms. In contrast to its theoretical importance, a generally accepted procedure how to measure utility does not yet exist. The need to have a method for determining utility functions is obvious, since violations of expected utility theory are frequent, in particular in the area of risky decision making. Without a generally accepted approach for identifying utility, it is impossible to figure out which theoretical predictions made by utility function related models do not fit observed decision making processes and hence many predictions of economic models are neither testable nor implementable. Several key questions have therefore be answered: Is the concept of utility functions a normative construct, does it capture the key features of decision making processes or is it just a tool for describing behavioral data?

In the present paper we focus on the utility of money. Two main methods to determine the utility of money are discussed in the literature: the evaluation of lotteries and the bisection method.

In the first approach, certainty equivalents (CE) of lotteries are determined for various amounts of money and various probabilities. Payoffs which are either 0 or x_1 ($x_1 = 100$ in **Figure 1A**) are offered by a lottery with a chance to win $p = 0.5$ or a sure win ("C" in **Figure 1A**) with the participant's task to choose one option (lottery or sure win). The CE is represented by the sure payoff for which a subject is indifferent between the two alternatives. Within this framework a utility function is determined based on its assumed functional form, a probability weighting and an econometric analysis. A key problem for econometric analyses is that each model, e.g. Regret Theory (Loomes and Sugden, 1982), Prospect Theory (Kahneman and Tversky, 1979), or Disappointment Theory (Bell, 1985), assumes a distinct functional form for the evaluation of money and probabilities. Consequently, the determination of utility functions for risky decisions differs across theoretical approaches. Another shortcoming of the binary lottery approach is that the evaluation of probabilities and money are not fully separable. Thus, risk aversion can be attributed either to the shape of the utility function or to the probability weighting. The statement of Prospect Theory that individuals are risk averse for positive payoffs or high probabilities therefore depends crucially on the assumptions of econometric analysis. On the other hand, one of this method's advantages is that agents really decide between options which lead to payoffs.

In the second approach, the bisection method, agents are asked to specify differences in the utility associated with monetary payoffs. In this method, a utility function is elicited by determining

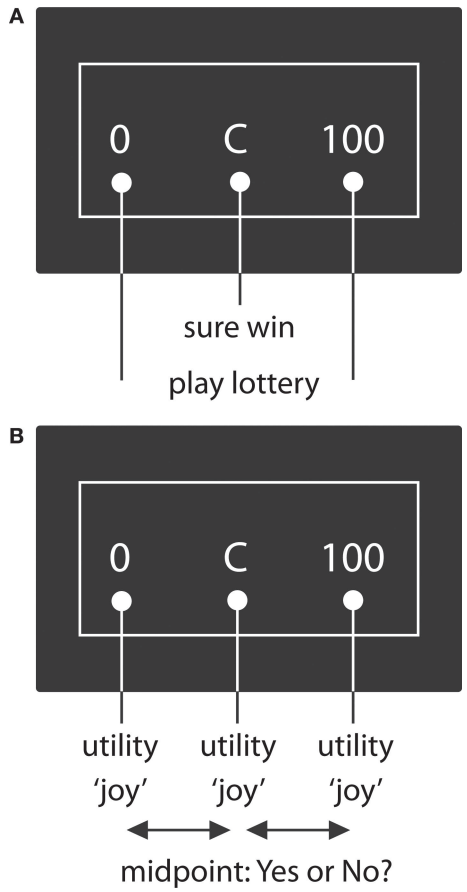


FIGURE 1 | Prototypical decision task for the binary lottery (A) and the bisection task (B).

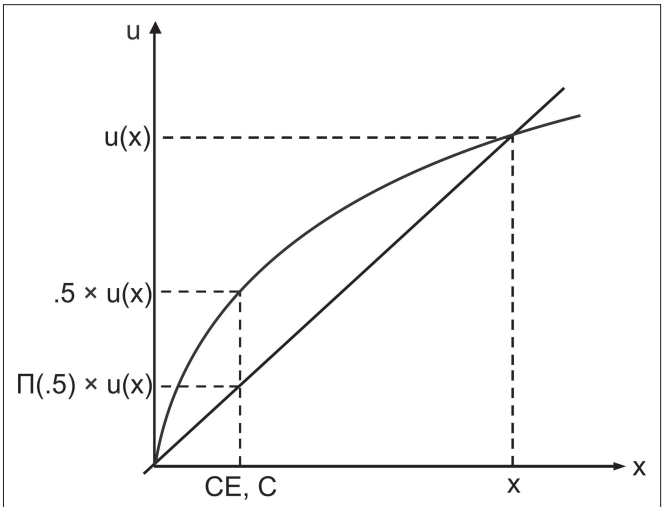


FIGURE 2 | Theoretical procedure how to determine a utility function if probabilities are perceived linearly.

mean points of utility between two utility values, which are x_1 and 0 in **Figure 1**. One possibility is, to present agents with two amounts of money, x_1 and 0, and ask them which amount of money M divides the utility difference $u(x_1) - u(0)$ into halves, i.e. $u(M) = (u(x_1) + u(0))/2 = u(x_1)/2$. It is also feasible to show a third amount of money x_3 ("C" in **Figure 1**) and to ask if this value divides the difference of x_1 and 0 into halves $[(u(x_1) + u(0))/2 = u(M)]$. To achieve a monetary valuation without using lotteries, subjects are asked to evaluate their perceived "happiness that money brings" (Galanter, 1962). In the present study the term "joy" is used (associated with receiving a specified amount of money) in order to induce a monetary valuation context. By varying the parameters, a utility function can be obtained. This method's advantage is that the resulting utility function does not depend on probabilities and the specification of a functional form, meaning that no theoretical presumptions are required. Its disadvantage is that decisions are neither connected to monetary nor to hypothetical payoffs.

Both methods have been widely used and discussed in the literature. In most studies in experimental economics the evaluation of lotteries is used to determine utility functions or to test theories. The key argument for preferring the lottery method over the bisection method is that economic decisions should involve monetary payoffs, otherwise the decisions "are not for real". **Figure 2** illustrates the theoretical foundation for both procedures. By asking for the CE or the midpoint the value on the x-axis is determined, i.e. CE or M, respectively. The value u(CE) or $u(M)$ on the y-axis is not given. For the bisection task it is by definition half of the utility of x_1. In the lottery condition this value depends on the theory describing the evaluation of lotteries. In Prospect Theory for example it is $\Pi(0.5)^* u(x_1)$. If the probability weight $\Pi(0.5)$ is small enough one might also obtain a linear utility function even if the CE is below the expected value. Therefore one has to know $\Pi(0.5)$ if one wants to determine the utility function by means of the lottery method. Assuming that both methods are based on the same utility function one could combine both methods to first get the utility function by means of the bisection method and then determine the probability weight of 0.5. This procedure can also be applied to probabilities different from 0.5, enabling one to determine a probability weighting function, and a utility function experimentally by combining both methods.

While economic scientists have pointed out differences between the two methods, e.g. stating that decisions in the bisection task are not for real, such difference might better be captured from a cognitive neuroscience point of view. Moreover, a cognitive neuroscience approach to this problem may also reveal differences in the neural processes involved in the two decision methods. In human beings decision making processes are supervised "online" by cognitive control mechanisms enabling adaptive behavior in a most flexible manner (Ridderinkhof et al., 2004; Botvinick, 2007). Typical simple situations in which these mechanisms have been studied involve response selection from several action alternatives or the evaluation of currently made decisions. By using event-related potentials (ERP) the neural underpinnings of these control mechanisms can be revealed. One ERP component related to response evaluation processes is the error-related negativity (error related negativity (ERN) Falkenstein et al., 1991; Gehring et al., 1993). This component was initially described to appear 50 to 100 ms following an incorrect response in choice-reaction tasks at fronto-central electrode sites and was postulated to reflect the perceived discrepancy

between the intended and the actually performed action. Source analysis as well as simultaneous analysis (Dehaene et al., 1994) of ERPs and functional magnetic resonance imaging (fMRI, Mathalon et al., 2003; Debener et al., 2006) have shown that the ERN is generated in the anterior cingulate cortex, an area that is closely linked to several cognitive control mechanisms involved in decision making (Gehring and Knight, 2000; Paus, 2001) and in the processing of risk related feedback information (Yeung and Sanfey, 2004; Cohen et al., 2007). Recent investigations have shown that the ERN is also sensitive to characteristics in error processing that are not directly linked to the violation of objective criteria. For example, an error may be more relevant by associating it with the loss of money (Hajcak et al., 2005) or by manipulating the participant's mood state (Tucker et al., 1999; Luu et al., 2000; Wiswede et al., 2009). These manipulations also influence ERN amplitudes. As Hewig et al. (2007) have shown, the high risk choices resulted also in an ERN, because such a selection implies a high chance not to get the response's intended outcome and will be processed as an error-like deviation from advantageous choice strategies. Based on their initial approach explaining error processing in terms of reinforcement learning (Holroyd and Coles, 2002; see also Munte et al., 2007 for electrophysiological evidences) Holroyd and Coles (2008) described the occurrence of an ERN in the absence of external ascertainable response criteria. Accordingly, responses are matched against internal criteria that were formed by individual learning histories and ERN amplitude is driven by the internal classification of a given response as "sub-optimal" (Holroyd and Coles, 2008). The ERN thus may reflect the subjective value of a potential response.

In the present investigation we use the ERN as a tool to characterize the neural implementation of decisions made in the lottery and bisection paradigms. Our prediction is that lottery decisions will be associated with increased monitoring, since the payoff instruction increases the subjective relevance or value in this task which should amplify the ERN amplitude. ERPs in the bisection task should not feature an ERN, because here responses have neither to be matched against set criteria nor are they associated with subjective relevance. Thus, we assume that fundamentally different neural processes will be engaged by the two methods. Importantly, we further ask whether the engagement of such different neural processes would also lead to differences in the estimates for the utility function of money.

MATERIALS AND METHODS

PARTICIPANTS

Sixteen neurological healthy, right-handed participants gave informed consent to take part in the study (10 women, age range 21–29). Two of these were excluded because of technical problems and one participant made no disadvantageous responses in the range from 390 to 100 and was excluded from data analysis as well. The final data-set thus comprised 13 participants. They were paid €7 per hour. The study protocol had been approved by the ethics committee of Magdeburg University.

GENERAL PROCEDURE

Participants were seated in a comfortable chair in front of a 19"-CRT monitor. A modified computer mouse was positioned under each index finger as a response device. The experiment consisted of two sessions which took place within 3–7 days. In both sessions identical stimulus material was presented but with differing task instructions. Every session began with 20 practice trials to familiarize subjects with the task. Thereafter the session started comprising 10 blocks of 82 trials each.

TASK

In each trial, lasting between 2700 to 3400 ms, a string of three numbers surrounded by a white box was presented (see **Figure 3**). The two outer numbers were shown first. After 1000 ms, the inner number was added and the completed array stayed on the screen for another 1000 ms. The array's left number was always zero. If numbers on the right were between 800 and 1000, the mid position numbers varied between 100 and 700, in case the right-sided numbers were between 1020 and 1200, mid position numbers were in the range of 300 to 900. Numbers in the middle were varied in steps of 50, right-sided numbers in steps of 20. Before presentation numbers within each string were multiplied by 1, 10 or 100, resulting in three classes of strings (e.g. "0 350 1120", "0 3500 11200" and "0 35000 112000").

In the binary lottery task participants had to choose to either get the amount of money corresponding to the center number or to play a lottery in which the outer numbers were the lottery's stakes played out at a fifty-fifty chance. Participants were explicitly told that the lottery game was hypothetical only; to that effect subjects expected no real payoff. They indicated their choices by pressing a button with the left or right index finger.

In the bisection task, the outer numbers corresponded to the utility interval's boundaries and the inner number to this interval's center. To keep the emotional framing of the bisection task comparable to the lottery task participants were instructed to imagine for each presented number the joy they would feel when getting this amount of money in Euro. By pressing the left or right index finger (YES/NO) subjects indicated whether the difference in perceived joy between the left – this number is always zero- and the center number and the center and the right number was felt to be equal (YES) or not (NO, see **Figure 1**). In both conditions subjects did not receive any performance feedback.

EEG-RECORDING AND ANALYSIS

The electroencephalogram was recorded from 28 tin electrodes, referenced against an electrode place on the left mastoid process, mounted in an elastic cap and placed according to the international 10–20 system. EEG was re-referenced offline to the mean activity at the left and right mastoid processes. All channels were amplified (bandpass 0.05–30 Hz) and digitized with 4 ms resolution. To control for eye movement artifacts, horizontal and vertical electrooculograms were recorded using bipolar montages. To eliminate eye movement contamination from EEG signals we used second order blind source separation as described by Joyce et al. (2004; for a comparison to other methods see Kierkels et al., 2006). Additionally, we controlled for other artifacts, e.g. muscle or heart rate, by visual inspection and removed afflicted epochs if necessary.

The generation of bins for ERP analysis was based on a difference value calculated for each trail. This difference value was computed by subtracting the arithmetic middle of the two outer numbers from the number presented in the center of the array.

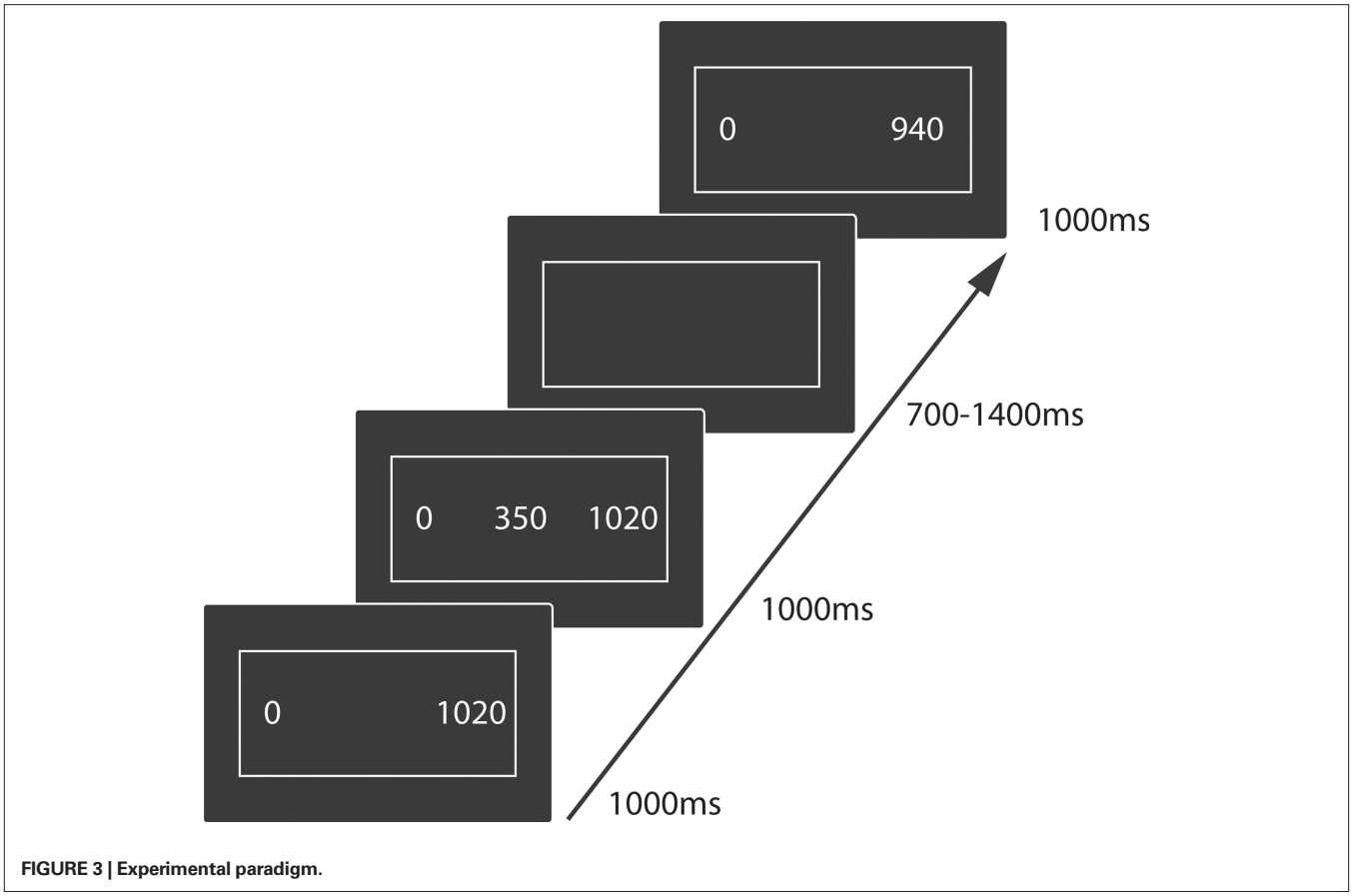


FIGURE 3 | Experimental paradigm.

To give an example the numbers shown in **Figure 3** are used: The outer numbers of the array are 0 and 1020, the number in the center is 350. The arithmetic middle of the two outer numbers is (0 + 1020)/2 = 510 and the resulting difference is 350 − 510 = −160. The negative value of this difference indicates that the number presented in the array's center is smaller than the arithmetic middle of the corresponding outer numbers. Accordingly, numbers presented in the array's center larger than the arithmetic middle of the outer numbers produce a positive difference value. Based on the difference values we sorted trials into five bins including the differences 390 to 100, 90 to 50, 40 to −40, −50 to −90 and −100 to −400. For each of these bins, we computed response-locked averages with an epoch-length of 900 ms (baseline −300 to 0) separately for "yes" and "no" responses. For each subject, averages were filtered using a 1–8 Hz band pass filter before calculating the mean amplitude 30–70 ms after response for statistical analysis. This time-window has been shown to capture the ERN component which typically has a maximum around 50 ms. To test for effects, we calculated an ANOVA with the factors condition (lottery/center judgement), response (yes/no, where "yes" in the lottery condition is related to choosing money and "no" choosing the lottery) and bin (five bins) for the electrode site Cz. Significance values will be reported Greenhouse-Geisser corrected, but the degrees of freedom uncorrected. In order to identify conditions causing significant interactions or main effects the corresponding *post-hoc t*-tests were performed. To adjust the significance level of one-tailed *post-hoc t*-tests for multiple comparisons α was set to 0.05 and an improved Bonferroni procedure based on the ordered *p*-values was applied (Simes, 1986). According to Simes (1986), let $p_{(1)} \leq p_{(2)} \leq \ldots \leq p_{(j)}$ be the ordered *p*-values for testing $H_0 = \{H_1, H_2, .., H_i\}$. H_0 will be rejected, whenever $p_i < i^*\alpha/j$ for $i = 1, \ldots, j$ (see also Samuel-Cahn, 1996; Sen, 1999 for a critical discussion and Wendt et al., 2007 for its application on EEG data). According to this procedure we grouped our *post-hoc*-testing into two sets: comparisons for decisions at the indifference point [bin (−40; 40)] and comparisons for decisions made at the endpoint [bins (−400; −100) and (100; 390)]. Correspondingly, we calculated critical *p*-values for 6 and 7 *post-hoc* comparisons. A similar correction was applied to the 5 *post-hoc* tests of the behavioral data. The resulting critical values are shown in **Tables 1 and 2** respectively.

BEHAVIORAL ANALYSIS

We calculated the means $\mu_{bisection}$, $\mu_{lottery}$ and standard deviations $\sigma_{bisection}$, $\sigma_{lottery}$ for the bisection method and the lottery method respectively. For the Yes-responses of both methods we determined histograms that are representing the best fit to a normal or to a cumulative normal distribution function. For these histograms we calculated means and standard deviations. We tested if the bisection method's mean $\mu_{bisection}$ can be the mean of the distribution (histogram) of the data under the lottery condition and if the lottery task's mean $\mu_{lottery}$ can be the mean of the distribution under the bisection condition.

In a second analysis the same bins as used for the EEG analysis were assessed. Response frequencies were calculated for every participant by computing the percentage of 'yes' and 'no' responses given to each class of difference values. For statistical analysis the mean percentage values of 'yes' response were used only, since percentage values of both response categories are inversely related. For global effect testing, we performed a repeated measures ANOVA with the factors condition (lottery/center judgement) and bin (five bins, see previous section for details). Significance values will be reported Greenhouse-Geisser corrected, but the degrees of freedom uncorrected. For correction of *post-hoc* test's alpha value see section "EEG-Recording and Analysis".

RESULTS

BEHAVIOR

Choices made for each class of difference value (in mean percentage) are depicted for the lottery and bisection tasks in **Figures 4A,B** respectively. We fitted a normally distributed density function for the data in **Figure 4A** and a normally distributed cumulative distribution function for the data in **Figure 4B**. This procedure results in $\mu_{bisection} = -14.33$ and $\sigma_{bisection} = 175.74$ for the bisection method and $\mu_{lottery} = -32.57$ and $\sigma_{lottery} = 191.97$ for the binary lottery method. In a *t*-test the Null hypothesis that $\mu_{bisection} = -32.57$ is not rejected on the 40% level (absolute *t*-values are smaller than 0.64). The same holds for the Null hypothesis $\mu_{lottery} = -14.33$. We observed a slightly concave utility function under both conditions ($\mu < 0$) which corresponds to risk aversion in the lottery condition. The average μ is smaller in the lottery condition, but the difference is not significant.

Collapsing choices into 5 bins as done for the EEG analysis clearly illustrates the differences between the tasks (see **Figure 4C**). Statistical analysis for the YES-responses reveals a significant interaction condition by bin ($F(4,48) = 33.38$, $p < 0.001$) as well as significant main effects (condition $F(1,12) = 14.39$, $p = 0.003$; bin $F(4,48) = 32.04$, $p < 0.001$). Comparing bins between conditions *post-hoc* contrasts are significant for the bins [100; 390] and [50; 90], but not for the remaining bins [−40; 40], [50; 90] and [100; 400]. For example, these tests show that the difference in the means of the CE and the mean point are not significant.

Table 1 | Ordered *p*-values of the comparison of the relative frequency of "YES"-responses between conditions.

Order	Contrast	*t*-value	p_{crit}	p_{emp}
1	[100; 390]	9.47	0.01	**<0.001**
2	[50; 90]	4.53	0.02	**<0.001**
3	[–90; –50]	–0.833	0.03	**0.21**
4	[–40; 40]	–0.345	0.04	**0.36**
5	[–100; –400]	–0.167	0.05	**0.43**

Significant tests are indicated by p_{emp} values written in bold.

EVENT-RELATED POTENTIALS

The response-locked grand average ERPs are illustrated in **Figure 5**. A clear negativity with a peak latency at approximately 50 ms and a mediofrontal distribution (see **Figure 6**) akin the ERN component emerged in the lottery task for those responses which entailed a divergence from the optimal behavior [i.e. choosing the lottery in bin (100; 390) and choosing money for bin (−100; −400)]. By contrast, obviously incorrect responses to the same bins in the bisection task (**Figure 5D**, "center yes" responses in the bins [−100; −400] and [100; 390]) resulted in much smaller negativities. The amplitude of these negativities was similar to the responses in the indifference range (**Figure 5C**). Statistical analysis of response-locked ERPs resulted in significant interactions condition × response × bin [$F(4,48) = 7.18$, $p = 0.002$], condition × response [$F(1,12) = 5.29$, $p = 0.03$], condition × bin [$F(4,48) = 3.23$, $p = 0.03$] and response × bin [$F(4,48) = 3.91$, $p = 0.03$]. With regard to main effects, only the 'condition' factor became significant [$F(1,12) = 5.78$, $p = 0.03$; for the remaining main effects response and bin $F < 0.9$, $p > 0.4$]. *Post-hoc* comparisons contrasting decisions between bins within and between conditions are illustrated in **Table 2.**

Table 2 | Ordered *p*-values for the comparison of the mean ERN amplitudes between conditions.

Order	Contrast	*t*-value	p_{crit}	p_{emp}
1	[–40; 40] choice lottery vs. center no	–1.437	0.008	0.085
2	[–40; 40] choice lottery vs. center yes	–1.09	0.016	0.15
3	[–40; 40] choice lottery vs. choice money	–1.01	0.025	0.16
4	[–40; 40] choice money vs. center no	–0.481	0.033	0.34
5	[–40; 40] center no vs. center yes	0.377	0.041	0.36
6	[–40; 40] choice money vs. center yes	0.015	0.05	0.49
1	[100; 390] choice lottery vs. [100; 390] choice money	–4.55	0.007	**<0.001**
2	[100; 390] choice lottery vs. [–400;–100] choice money	–3.58	0.014	**0.002**
3	[100; 390] choice lottery vs. [100; 390] center yes	–3.05	0.021	**0.005**
4	[–400; –100] choice money vs. [–400; –100] choice lottery	–2.6	0.028	**0.011**
5	[100; 390] center no vs. [100; 390] center yes	–2.53	0.035	**0.011**
6	[–400; –100] center no vs. [–400; –100] center yes	–2.42	0.042	**0.014**
7	[–400; –100] choice money vs. [–400; –100] center yes	–1.7	0.05	**0.05**

Depicted are all performed contrasts; significant tests are indicated by p_{emp} values written in bold. The ordering of the critical p-values was performed for the indifference point and both endpoints respectively.

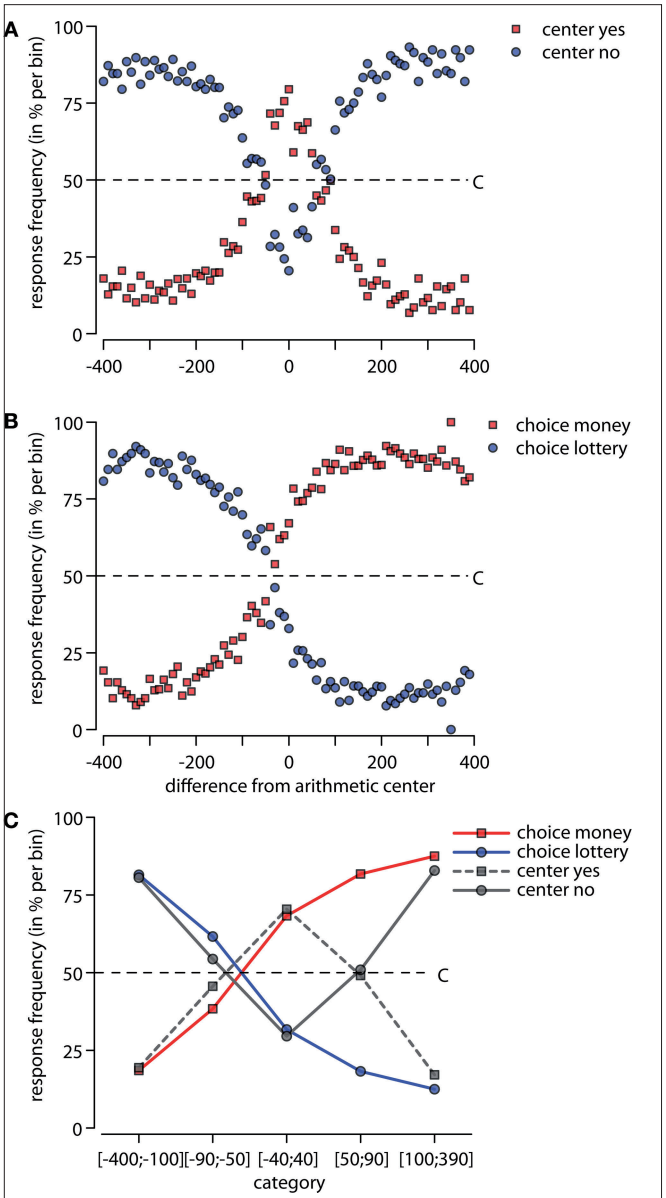


FIGURE 4 | Behavioral data. Mean percentage of choices are shown for the binary lottery **(A)** and the bisection task **(B)**. The center is indicated by a dashed line and refers to the indifference point in the bisection task. **(C)** depicts the cumulated choices per bin for both task. Circles referring to "YES"-, squares to "NO"-responses. Indifference point is indicated by a dotted line. Please note, that statistical comparisons were only calculated for YES responses.

Additionally, we checked whether differences in the ERPs seen between the unfavorable choices in the lottery task might simply be due to a different distance between the indifference point at −30 (see **Figure 4A**) and the respective bins [100;390] and [−100; −400]. Therefore, two new bins were obtained according to the procedure described in section "EEG-Recording and Analysis", but with new intervals (see **Figure 5B**), which are equally distant from the indifference point. Compared to the original bins the ERN amplitude related to the unfavorable lottery choices decreases while the ERN for unfavorable money choices increases. Nevertheless, the difference between both conditions remains significant (one-tailed t-test, $t = -2.08$, df = 12, $p = 0.03$).

DISCUSSION

In the present study, action monitoring processes as indexed by the ERN component were differentially engaged in two decision making paradigms that have been frequently used in economic sciences to determine the utility of money. This suggests fundamental differences between the lottery and bisection methods at the cognitive and neural level. More specifically, in the binary lottery task a pronounced ERN characterized decisions in favor of the less advantageous options at either endpoint of the range of possible decisions. That is, an ERN was found for "lottery"-decisions in trials that would have yielded a sure win greater than the mean outcome of the lottery. Similarly, an ERN was present for "sure win"-decisions in trials in which the mean outcome of the lottery would have exceeded the sure win. The occurrence of an ERN in the lottery task can be interpreted as evidence for the subjects' motivational participation even though the lottery's payoffs were only hypothetical, as it is known from previous research that the appearance of an ERN depends on the subjective significance in a given task (Hajcak et al., 2005; Holroyd and Coles, 2008). The difference between the trial bins associated with the two most prominent ERNs indicates different degrees in action monitoring and is likely caused by the interaction of two factors: the difference between expected value and sure payoff on the one hand and the risk to sustain a potential loss on the other. For both kinds of decisions associated with an ERN the difference between the expected value of the lottery and the sure win was similar. Nevertheless, the ERN's amplitude was larger for disadvantageous lottery decisions compared to the disadvantageous selection of a sure payoff. This suggests that the anticipation of a risky decision's potentially negative outcome, namely to win nothing instead of getting a small but sure payoff in the unfavorable sure win selection, leads to increased activation of monitoring mechanisms despite similar expected values. One might argue, that this difference is simply caused by the dissimilar distances between the indifference point at −30 and the end-bins respectively (see **Figure 4B**). However, the described effect persists (albeit somewhat smaller) even when trials were rearranged to yield bins that are equidistant to the indifference point. The change in ERN amplitude for these rearranged bins is due to the fact that, for lottery choices, the bin [70; 340] comprises less disadvantageous selections, whereas in the bin [−130; −400] the proportion of unfavorable choices increases. It is to be expected that the corresponding ERN amplitudes decrease and increase, respectively.

In contrast to the binary lottery task the bisection task's incorrect responses at the end-bins are associated with a very small negative deflection, i.e. a rudimentary ERN. Although the amplitudes of these negativities are significantly different from the end-bins' correct choices, the parameter values are indicating a much smaller degree of action monitoring involvement (Nieuwenhuis et al., 2001; Burle et al., 2008; Heldmann et al., 2008) and are similar to the amplitudes observed for the indifference point. Based on our previous argumentation the rudimentary ERNs can be seen as an indicator for the absence of risk perception in these conditions.

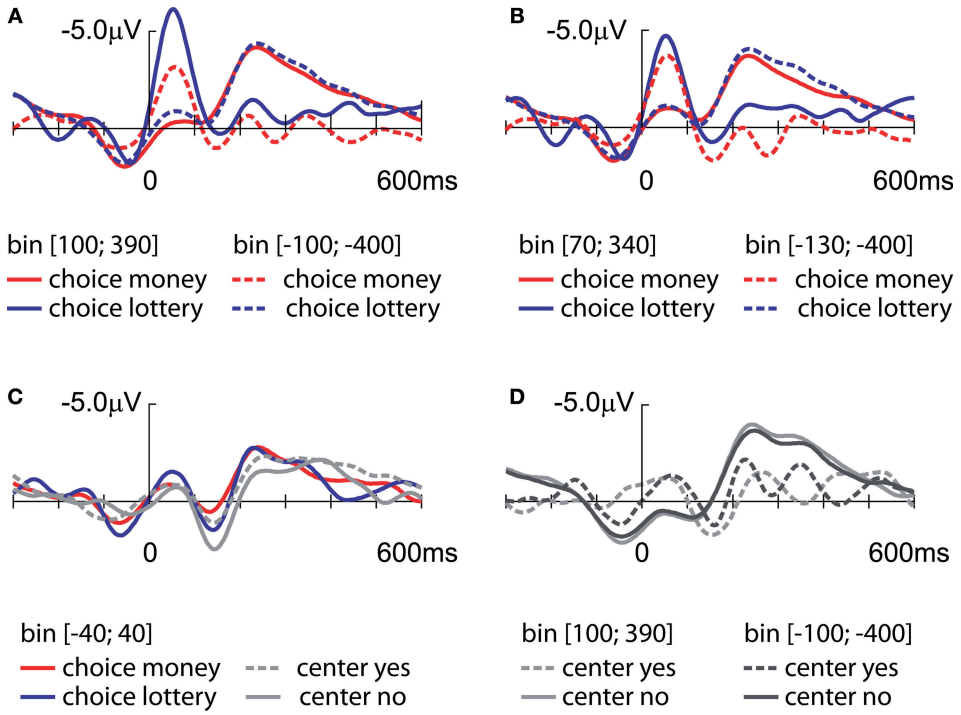


FIGURE 5 | Response-locked event related potentials for the lottery's task endpoints (A), the bisection's task endpoint (C) and the indifference point for all conditions (D) at Cz are shown. (B) Depicts the lottery's task ERPs which are corresponding to the bins equidistant to the indifference point [bin 70;340] and choice money [–130;–400].

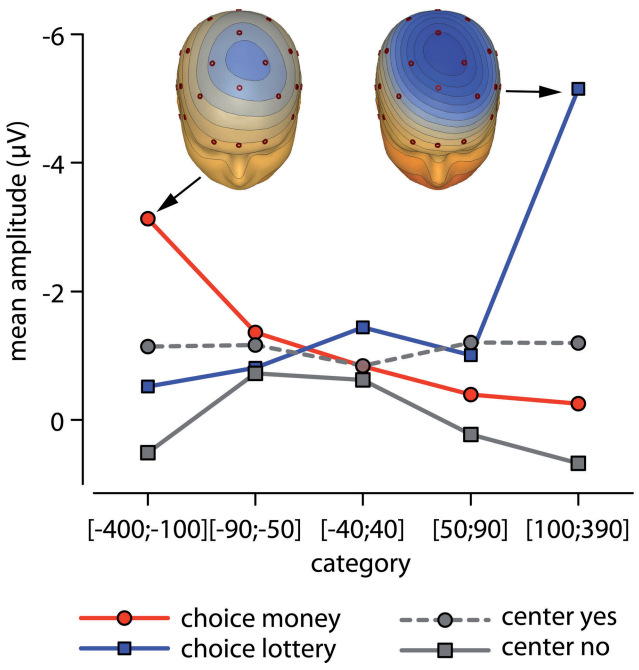


FIGURE 6 | ERN topography. Topographies for the lottery conditions left) choice money [–100; –400] and right) choice lottery [bin 100; 390] are shown. The line graph depicts the mean amplitudes of the response-locked ERPs at CZ per category. Circles referring to "YES"-, squares to "NO"-responses.

In which way are the ERP results able to inform economic reasoning? Looking at the economic starting point of our analysis shown in **Figure 1** we discussed the differences and similarities between the bisection and the binary lottery method. The bisection method determines the mean point in utility between the two monetary amounts only by the utility function of money with the following formula: $u(M) = 0.5^{*}u(x_1)$. In the binary lottery method the CE is determined by the utility function of money and other factors connected with the risk of the lottery, like probability weighting in Prospect Theory. According to Prospect Theory the CE is: $u(\mathrm{CE}) = \Pi(0.5)^{*}u(x_1)$. The difference between both formulas is obvious: the weighting of probabilities Π(prob). Using the formula of Prospect Theory we would expect the CE and the mean point in the bisection task M to be equal if we assume $\Pi(0.5) = 0.5$. Here, only the utility function of money determines the CE.

In our investigation behavioral as well as ERP data suggest that both methods used are resulting in similar utility functions for money. Although there are visible differences between the bisection and the binary lottery task's indifference point, indicating risk avoidance, the statistical analyses revealed that these differences are absolutely not significant. The corresponding ERPs for decisions around the indifference point of the bisection and the lottery task (see **Figure 5C**) are also indicating, that action monitoring processes are not differentially engaged in the related decision making processes. Postulating the bisection method captures utility of money itself and, as argued previously, the binary lottery the combination of utility and risk, an implication of our finding is

that utility function and probability weighting can be separated by initially determining the utility function with the bisection method and afterwards using the obtained function as input in the lottery method to get the lottery's probability weighting experimentally. This procedure could also be applied to lotteries with probabilities different from 0.5 and would allow for a more precise discrimination between effects of the probability weighting and utility function for money. The same procedure allows for the separation of other effects related to risk and not to money evaluation as well. For example, the implications of Regret or Disappointment Theory can be tested more easily by a combination of these two methods, since the result of the determination of the utility function can be used in the analysis of the lottery method. In general, using both methods and looking at the differences helps to characterize situations connected with risk (money and probabilities) in comparison to situations connected with certainty (money only).

It is important to note that ERNs were observed in the lottery condition even though money was not paid. As we stated previously, this indicates that despite the absence of real payoffs the choices made in the lottery task were subjectively more engaging compared to the bisection task and impelled participants to control their behavior to a larger extend. Assuming that hypothetical payoffs are less engaging then real payoffs, one would expect an increase in subjective relevance for real payoffs and therefore increasing ERN amplitudes. In line with the argumentation by Hewig et al. (2007) subjective relevance of an intended behavioral outcome is one factor that drives sensitivity for risk related choices. Therefore, we believe that it is justified to ascribe the hypothesized ERN amplitude increase for lotteries with real payoffs partially to an altered perception of risk.

According to standard economic theory of risky decision making the choice of an individual is solely determined by consequences and probabilities entailed in a decision. The influence of time between the presentation of the problem, the decision and the realisation of the outcome is very often neglected. In the present study we showed that the temporal resolution of the ERP method and its ability to reveal cognitive processes that are not directly linked to perceivable behavior make it possible to identify the point in time at which post-decision evaluation processes takes place. That is, if the performed choice fits the subject's response strategies and finally their long-term goals (Albers et al., 2000).

In summary, we have shown that the combination of the bisection and the binary lottery task allows to separate probability weighting and the perception of risk in the determination of utility functions for money. We characterized common properties and differences of these two methods. Our ERP results are also indicating, that disadvantageous choices in the risky decision making task are processed differentially by cognitive action monitoring processes. Since we found no evidence for differences around the indifference point at the behavioral or the neural level the use of both methods to determine the utility of money will result under the given conditions in similar estimates.

ACKNOWLEDGMENTS

This work was supported by the DFG (SFB 779 projects A3 and A5) awarded to Thomas F. Münte and Hans-Jochen Heinze. Also supported by a special grant of the Center for Behavioral Brain Sciences, Magdeburg, to Marcus Heldmann.

REFERENCES

Albers, W., Pope, R., Selten, R., and Vogt, B. (2000). Experimental Evidence for Attractions to Chance. *German Economic Review.* 1, 113–130.

Bell, D. E. (1985). Disappointment in Decision Making under Uncertainty. *Oper. Res.* 33, 1–27.

Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn. Affect. Behav. Neurosci.* 7, 356–366.

Burle, B., Roger, C., Allain, S., Vidal, F., and Hasbroucq, T. (2008). Error negativity does not reflect conflict: a reappraisal of conflict monitoring and anterior cingulate cortex activity. *J. Cogn. Neurosci.* 20, 1637–1655.

Cohen, M. X., Elger, C. E., and Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage* 35, 968–978.

Debener, S., Ullsperger, M., Siegel, M., and Engel, A. K. (2006). Single-trial EEG-fMRI reveals the dynamics of cognitive function. *Trends Cogn. Sci. (Regul. Ed.)* 10, 558–563.

Dehaene, S., Posner, M. I., and Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychol. Sci.* 5, 303–305.

Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalogr. Clin Neurophysiol.* 78, 447–455.

Galanter, E. (1962). The Direct Measurement of Utility and Subjective Probability. *Am. J. Psychol.* 75, 208–220.

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., and Donchin, E. (1993). A neural system for error detection and compensation. *Psychol. Sci.* 4, 385–390.

Gehring, W. J., and Knight, R. T. (2000). Prefrontal-cingulate interactions in action monitoring. *Nat. Neurosci.* 3, 516–520.

Hajcak, G., Moser, J. S., Yeung, N., and Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology* 42, 151–160.

Heldmann, M., Russeler, J., and Munte, T. F. (2008). Internal and external information in error processing. *BMC Neurosci.* 9, 33.

Hewig, J., Trippe, R., Hecht, H., Coles, M. G., Holroyd, C. B., and Miltner, W. H. (2007). Decision-making in Blackjack: an electrophysiological analysis. *Cereb. Cortex* 17, 865–877.

Holroyd, C. B., and Coles, M. G. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709.

Holroyd, C. B., and Coles, M. G. (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex* 44, 548–559.

Joyce, C. A., Gorodnitsky, I. F., and Kutas, M. (2004). Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. *Psychophysiology* 41, 313–325.

Kahneman, D., and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47, 263–291.

Kierkels, J. J., van Boxtel, G. J., and Vogten, L. L. (2006). A model-based objective evaluation of eye movement correction in EEG recordings. *IEEE Trans. Biomed. Eng.* 53, 246–253.

Loomes, G., and Sugden, R. (1982). Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty. *Econ. J.* 92, 805–824.

Luu, P., Collins, P., and Tucker, D. M. (2000). Mood, personality, and self-monitoring: negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *J. Exp. Psychol. Gen.* 129, 43–60.

Mathalon, D. H., Whitfield, S. L., and Ford, J. M. (2003). Anatomy of an error: ERP and fMRI. *Biol. Psychol.* 64, 119–141.

Munte, T. F., Heldmann, M., Hinrichs, H., Marco-Pallares, J., Kramer, U. M., Sturm, V., and Heinze, H. J. (2007). Nucleus Accumbens is Involved in Human Action Monitoring: Evidence from Invasive Electrophysiological Recordings. *Front. Hum. Neurosci.* 1, 11. doi:10.3389/neuro.09.011.2007.

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., and Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: evidence from an antisaccade task. *Psychophysiology* 38, 752–760.

Paus, T. (2001). Primate anterior cingulate cortex: where motor control, drive and

cognition interface. *Nat. Rev. Neurosci.* 2, 417–424.
Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447.
Samuel-Cahn, E. (1996). Is the Simes improved Bonferroni procedure conservative? *Biometrika.* 83, 928–933.
Sen, P. K. (1999). Some remarks on Simes-type multiple tests of significance. *J. Stat. Plan. Inference.* 82, 139–145.
Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika* 73, 751–754.
Tucker, D. M., Hartry-Speiser, A., McDougal, L., Luu, P., and deGrandpre, D. (1999). Mood and spatial memory: emotion and right hemisphere contribution to spatial cognition. *Biol. Psychol.* 50, 103–125.
Wendt, M., Heldmann, M., Munte, T. F., and Kluwe, R. H. (2007). Disentangling sequential effects of stimulus- and response-related conflict and stimulus-response repetition using brain potentials. *J. Cogn. Neurosci.* 19, 1104–1112.
Wiswede, D., Munte, T. F., Goschke, T., and Ruesseler, J. (2009). Modulation of the error-related negativity by induction of short-term negative affect. *Neuropsychologia* 47, 83–90.
Yeung, N., and Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *J. Neurosci.* 24, 6258–6264.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 May 2009; paper pending published: 03 June 2009; accepted: 13 October 2009; published online: 30 October 2009.
Citation: Heldmann M, Vogt B, Heinze H-J, and Münte TF (2009) Different methods to define utility functions yield similar results but engage different neural processes. Front. Behav. Neurosci. ***3****:43. doi: 10.3389/neuro.08.043.2009*

frontiers in
BEHAVIORAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 11 June 2009
doi: 10.3389/neuro.08.009.2009

Temporal discounting and inter-temporal choice in rhesus monkeys

Jaewon Hwang[1†], Soyoun Kim[2†] and Daeyeol Lee[2*]

[1] *Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY, USA*
[2] *Department of Neurobiology, Yale University School of Medicine, New Haven, CT, USA*

Edited by:
Jeansok J. Kim, University of Washington, USA

Reviewed by:
Ben Seymour, Wellcome Trust Centre for Neuroimaging, UK
Veit Stuphorn, Johns Hopkins University, USA

***Correspondence:**
Daeyeol Lee, Department of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510, USA.
e-mail: daeyeol.lee@yale.edu

†*Jaewon Hwang and Soyoun Kim contributed equally to this study.*

Humans and animals are more likely to take an action leading to an immediate reward than actions with delayed rewards of similar magnitudes. Although such devaluation of delayed rewards has been almost universally described by hyperbolic discount functions, the rate of this temporal discounting varies substantially among different animal species. This might be in part due to the differences in how the information about reward is presented to decision makers. In previous animal studies, reward delays or magnitudes were gradually adjusted across trials, so the animals learned the properties of future rewards from the rewards they waited for and consumed previously. In contrast, verbal cues have been used commonly in human studies. In the present study, rhesus monkeys were trained in a novel inter-temporal choice task in which the magnitude and delay of reward were indicated symbolically using visual cues and varied randomly across trials. We found that monkeys could extract the information about reward delays from visual symbols regardless of the number of symbols used to indicate the delay. The rate of temporal discounting observed in the present study was comparable to the previous estimates in other mammals, and the animal's choice behavior was largely consistent with hyperbolic discounting. Our results also suggest that the rate of temporal discounting might be influenced by contextual factors, such as the novelty of the task. The flexibility furnished by this new inter-temporal choice task might be useful for future neurobiological investigations on inter-temporal choice in non-human primates.

Keywords: reward, neuroeconomics, decision making, prefrontal cortex

INTRODUCTION

The rewards that humans and animals seek to obtain are often not delivered immediately after the required actions are completed. In such cases, the subjective desirability or utility of the expected reward decreases with its delay, and this is referred to as temporal discounting. Consequently, during inter-temporal choice in which the decision makers choose between rewards delivered after unequal delays, they might in some cases prefer a small but immediate reward to a larger but more delayed reward. Such impulsive choices can be often parsimoniously accounted for by a discount function, which is defined as the fraction of the subjective value of a delayed reward relative to that of the same reward delivered immediately. The value of a delayed reward multiplied by the discount function is referred to as the temporally discounted value. In addition, denoting the discount function as $F(D)$, in which D refers to the delay of a reward, the ratio $F'(D)/F(D)$ is referred to as the discount rate and indicates how rapidly the discount function decreases with delay. Abnormally high discount rate underlies a number of psychiatric disorders, including substance abuse and pathological gambling (see Reynolds, 2006).

Regardless of the absolute value of discount rate, if the discount rate is constant and does not change with the reward delay, the discount function is exponential (Samuelson, 1937). This implies that the relative preference for two different rewards available at time t_1 and t_2 would not be affected when their delays are altered by the same amount and become available at time $t_1 + \Delta t$ and $t_2 + \Delta t$, respectively. The fact that the preference between the two delayed rewards does not change with the elapse of time is referred to as time-consistency, but this assumption is commonly violated (Ainslie and Herrnstein, 1981; Green et al., 1981, 1994; Rachlin and Green, 1972; Strotz, 1955–1956). In addition, a large number of empirical studies have found that behaviors of humans and animals during inter-temporal choice are better described by hyperbolic discount functions that violate time consistency (Frederick et al., 2002; Green and Myerson, 2004; Kalenscher and Pennartz, 2008). A decision maker with a hyperbolic discount function might prefer a larger and more delayed reward when both rewards have relatively large delays, but his or her preference might change when their delays are reduced by the same amount.

Although hyperbolic discount functions have successfully described behaviors for many different animal species, including humans, the overall rate of temporal discounting varied tremendously between humans and other animals. The reasons for this discrepancy are not fully understood, but might be related to the differences in the methods to measure the discount functions for humans and animals. In human studies, choices are typically presented using verbal cues, and the subjects are often allowed to engage in other activities while waiting for the delivery of rewards. In contrast, animals are tested in a more controlled environment and consume their chosen rewards after experiencing the corresponding delays. Moreover, in previous animal studies, reward delays and magnitudes are either fixed or adjusted gradually across

successive trials so that they must be estimated from the animal's experience. In the present study, we trained rhesus monkeys in a new inter-temporal choice task in which the information about the magnitude and delay of each reward is delivered symbolically and as a result could be manipulated independently across trials. We found that the animal's behaviors were largely better accounted for by hyperbolic discount functions, whereas the form and rate of temporal discounting might be influenced by the novelty of the task.

MATERIALS AND METHODS

ANIMAL PREPARATION AND APPARATUS

Two male rhesus monkeys (monkeys D and J; body weight = 9.5 and 9.0 kg) were tested. During an aseptic surgery, a set of four titanium head posts were attached to the animal's skull for the purpose of fixing the animal's head during the experiment. The animals were seated in a primate chair and faced a 17-inch computer monitor located 57 cm away. A custom-designed software was used to control the task and coordinate data acquisition. Eye movements were monitored using a video eye tracking system with 225 Hz sampling rate (ET-49, Thomas Recording, Germany). All the procedures used in the present study were in accordance with the guidelines of the National Institutes of Health and were approved by the University of Rochester Committee on Animal Research.

INTER-TEMPORAL CHOICE TASK

General

Each trial began when the animal fixated a white square (0.9° × 0.9°) presented at the center of the monitor (**Figure 1**). After a 1-s fore-period during which the animal was required to maintain its fixation of the central square within a 2°-radius window, two targets (1° disk in diameter) were presented 8° to the left and right of fixation. The animal was required to continue its central fixation until the white square was extinguished 1 s later. At the end of this cue period, the animal was then required to shift its gaze towards one of the two targets. One of the targets (TS) was green and delivered a small reward when it was chosen by the animal, whereas the other target (TL) was red and delivered a large reward. The delay between the fixation of the chosen target and the reward delivery was indicated by a variable number of small disks (0.9° in diameter) presented around each target. When the target was presented without any disks, the animal was rewarded after a 0.5 delay (Experiment I) or immediately (Experiments II and III) upon fixation of its chosen target. Otherwise, disks were extinguished one at a time according to a specific schedule described below, and the animal was rewarded after all the disks were extinguished. Yellow disks were extinguished at the rate of 0.5 s/disk (Experiment I) or 1.0 s/disk (Experiments II and III). In Experiment III, a mixture of yellow (1.0 s/disk) and cyan (4.0 s/disk) disks were used in some trials. The brightness of a yellow disk was fixed until it was extinguished, whereas a cyan disk dimmed gradually during the 4-s period before it was extinguished. The target that was not chosen by the animal and its clock were extinguished immediately after the animal fixated its chosen target. If the animal chose the large reward, the central white square for the next trial was presented following a 2-s inter-trial interval after the reward delivery. If the animal chose the small reward, the inter-trial interval was increased by the difference in the reward delays for the small and large reward targets. Therefore, the onset of the next trial was not affected by the animal's choice.

The animal was required to maintain its fixation of the chosen target during the reward delay, but was allowed to re-fixate the target without any penalty if the target was re-fixated within 0.3 s after breaking the fixation. This also allowed the animals to blink without any penalty during the fixation on its chosen target. Throughout the experiment, the proportion of the trials that were aborted due to the animal's failure to maintain its fixation during the reward delay was relatively low and never exceeded 2% of the trials. This always corresponded to a relatively small proportion fixation breaks during the entire trials, never exceeding 17% of all fixation breaks (mean = 1.6% and 6.9% for monkeys D and J, respectively). Moreover, extensive training was not necessary for fixation during the reward delays, and the animals frequently made saccades among the small disks. Although we could not quantify the additional efforts necessary for the fixation of the chosen target

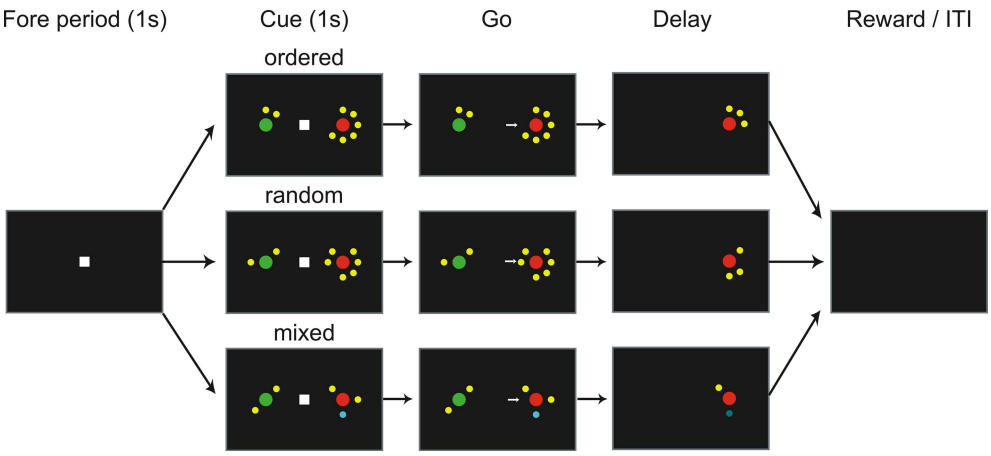


FIGURE 1 | Spatiotemporal sequences of the inter-temporal choice task. Three different types of clocks are referred to as ordered, random, and mixed. For both ordered and random clocks, the reward delay was indicated by the number of yellow disks that disappeared in a fixed or random order, respectively. Each yellow and cyan disk in mixed clocks corresponds to 1 and 4 s added to the reward delay, respectively.

during the reward delays, these observations indicate that such efforts are likely to be relatively minor.

Reward delays and clocks

All the disks in the clock for a given target were presented on the circumference of an imaginary circle (4.0° in diameter) concentric with the target. In the following, the position of a disk in a given clock is described by its clockwise angular deviation from the position directly above the target. Disks were presented only at multiples of 45° (**Figure 1**). In the present study, three different types of clocks were used, and referred to as ordered, random, and mixed, respectively. For ordered and random clocks, only yellow disks were used, whereas mixed clocks included both yellow and cyan disks. In an ordered clock with n yellow disks, disks were presented at the positions corresponding to 0°, 45°,…, $(n-1)\times 45°$, and were extinguished counter-clockwise during the reward delay so that the disk at 0° position was always extinguished at the end of the reward delay (**Figure 1**, top). In random and mixed clocks, the positions of disks were determined randomly, and they were extinguished in a random order during the delay period (**Figure 1** middle and bottom).

Preliminary training

Each animal was initially trained to fixate the central white square. Next, it was trained to choose between the green small-reward target and the red large-reward target, while the delay for the small reward was always 0.5 s. Within a few days, both animals were gradually exposed to various reward delays and started to choose the large-reward target less frequently as its reward delay increased. No rewards were omitted during this training period, as long as the animal performed the task correctly. Before the data collection began for Experiment I, monkeys D and J were trained for this inter-temporal choice task for 9 and 12 days, respectively.

Experiment I

During the trials of Experiment I, only the ordered clocks were used and all disks in the clocks were yellow. The reward delay for the clock with n yellow disks was $(n+1)/2$ s, where $n = 0, 1,\ldots 8$, corresponding to the delays ranging from 0.5 to 4.5 s. Among the 64 possible combinations of reward delays for the two targets, only those in which the reward delay for the large-reward target was equal to or longer than the delay for the small-reward target were used. This resulted in 45 different combinations of the reward delays. The positions of the large-reward and small-reward targets were counter-balanced across trials, resulting in 90 trials in a block. In Experiment I-A, the animal received 0.2 and 0.4 ml of apple juice for small and large rewards, respectively. The size of the small reward was increased to 0.27 ml in Experiment I-B, in order to encourage the animals to choose the small-reward target more frequently. Each animal performed 10 blocks (900 trials) each day (**Table 1**). Monkey D was tested in Experiment I-A for 5 days and then in Experiment I-B for 5 days, whereas the order of these two experiments was reversed for Monkey J.

Experiment II

In Experiment II, the clock with n yellow disks indicated that the reward delay was n seconds ($n = 0, 1,\ldots, 8$). Thus, reward delays ranged between 0 and 8 s. During Experiment II, the small and large rewards were 0.27 and 0.4 ml of juice. As in Experiment I, all possible combinations of reward delays were used as long as the delay for the large reward was equal to or larger than the delay for the small reward. Each animal performed 10 blocks (900 trials) daily. Only the random clocks were used in Experiment II-A, whereas for Experiment II-B, only the ordered clocks were used (**Table 1**). After Experiment I, both animals were tested in neurophysiological experiments in which a subset of conditions included in Experiment II-A was used (Kim et al., 2008). Accordingly, Experiment II was conducted approximately 6 and 8 months after Experiment I for monkeys D and J, respectively. Both animals were tested for 5 days in Experiment II-A, and then for 5 days in Experiment II-B.

Table 1 | Summary of conditions tested in each experiment.

Experiment	Clock type	Reward delays (s)	Reward magnitude	*N* trials/ animal
I-A	Ordered	0.5–4.5	1:2	4,500
I-B	Ordered	0.5–4.5	2:3	4,500
II-A	Random	0–8	2:3	4,500
II-B	Ordered	0–8	2:3	4,500
III-A	Mixed	0–8	2:3	3,840
III-B	Random	0–8	2:3 or 1:2	4,500

Experiment III

In Experiment III-A, mixed clocks were introduced to test whether the animals could extract the information about the reward delays independently of the number of disks in the clock. During Experiment III, a clock that includes n_Y yellow disks and n_C cyan disks indicated the reward delay of $(n_Y + 4\ n_C)$ s. Therefore, clocks did not include any cyan disks ($n_C = 0$) if the reward delay was less than 4 s. In addition, when the reward delay was 4, 5, 6, or 7 s, a given delay was indicated by one of two different types of clocks ($n_C = 0$ or 1). For example, the delay of 4 s could be indicated by $(n_Y, n_C) = (4, 0)$ or (0,1), and the delay of 5 s by (5, 0) or (1, 1). Finally, three different types of clocks were used to indicate the 8-s reward delay, namely, $(n_Y, n_C) = (8, 0)$, (4, 1), or (0, 2). Accordingly, 15 different types of clocks were available to indicate the reward delay ranging from 0 to 8 s. To limit the number of different combinations of clocks, the reward delays for the small-reward target were restricted to 0, 2, 4, and 6 s. Excluding the cases in which the delay for the small reward is longer than the delay for the large reward, therefore, a total of 64 different combinations of clocks were used in Experiment III-A. The positions of the large-reward and small-reward targets were counter-balanced, and this resulted in 128 trials in a given block. Both monkeys were tested for 5 days in Experiment III-A and completed six blocks (768 trials) each day. In Experiment III-A, the animal was rewarded by 0.27 and 0.4 ml of juice for choosing the small-reward and large-reward target, respectively.

Prior to Experiment III-A, both animals were trained with mixed clocks for several weeks. This preliminary training began approximately 5 and 3 months after Experiment II for monkeys D and J, respectively. During this preliminary training, each animal was trained for 17 days (monkey D) or 13 days (monkey J) with a subset of reward delays used in Experiment III-A in which the

delay for the small reward was either 0 or 2 s. Each animal was then trained for another day (day 18 and day 14 for monkeys D and J, respectively) with all the conditions described above for Experiment III-A before collecting the data described in the Results. After Experiment III-A, one of the monkeys (monkey J) was tested using the mixed clocks in a neurophysiological experiment (Kim et al., 2008). During this period, only a subset of reward delays in Experiment III-A was used (0 and 2 s for small reward and 0, 2, 5, and 8 s for large reward). Both animals were then tested in Experiment III-B in order to investigate whether exposure to mixed clocks influenced the animal's discount function. Experiment III-B was identical to Experiment II-A, except that the magnitude of small reward was reduced to 0.2 ml for monkey J.

DATA ANALYSIS

In the following, the symbol Ω is used to denote a set of variables corresponding to the magnitudes and delays of small and large rewards. Namely, $\Omega = \{A_{TS}, A_{TL}, D_{TS}, D_{TL}\}$, in which A_{TS} (A_{TL}) and D_{TS} (D_{TL}) refer to the magnitude and delay of small (large) reward, respectively. To estimate the animal's discount function from its choices, we assumed that the probability of choosing TS given Ω, $P(\mathrm{TS}|\Omega)$, was determined by the difference in the temporally discounted values for the two targets. In other words, denoting the temporally discounted value of a given target x as $\mathrm{DV}(A_x, D_x)$,

$$\log\mathrm{it}\, p(\mathrm{TS}\,|\,\Omega) \equiv \log\frac{p(\mathrm{TS}\,|\,\Omega)}{1-p(\mathrm{TS}\,|\,\Omega)} = \beta\left[\mathrm{DV}\left(A_{TS},\, D_{TS}\right) - \mathrm{DV}\left(A_{TL},\, D_{TL}\right)\right].$$

This is also known as softmax transformation, and is equivalent to the Boltzmann distribution given by the following:

$$p(\mathrm{TS}\,|\,\Omega) = \frac{\exp\beta\,\mathrm{DV}\left(A_{TS}, D_{TS}\right)}{\exp\beta\,\mathrm{DV}\left(A_{TS}, D_{TS}\right) + \exp\beta\,\mathrm{DV}\left(A_{TL}, D_{TL}\right)},$$

where β denotes the inverse temperature controlling the randomness of the animal's choices. In addition, $p(\mathrm{TL}|\Omega) = 1 - p(\mathrm{TS}|\Omega)$. Therefore, $p(\mathrm{TS}|\Omega) = p(\mathrm{TL}|\Omega) = 0.5$, if the temporally discounted values are equal for both targets, and $p(\mathrm{TS})$ approaches 1, as the temporally discounted value of TS increases. The temporally discounted value of the reward with the magnitude A and delay D is determined by the following:

$$\mathrm{DV}(A, D) = A\,F(D),$$

where $F(D)$ refers to a discount function. An exponential discount function corresponds to the following:

$$F_E(D) = \exp(-k_E D),$$

where k_E denotes the discount rate (s^{-1}). A hyperbolic discount function can be given by the following:

$$F_H(D) = \frac{1}{1+k_H D},$$

where the parameter k_H controls the steepness of discounting. We have also tested three additional discount functions. One of them is a variant of hyperbolic discount function in which the more immediate reward is not discounted and the more delayed reward is discounted according to the hyperbolic discount function based on the difference in the delays of the two rewards (Green et al., 2005). In addition, the general hyperbolic discount function (Green and Myerson, 2004; Takahashi et al., 2008), F_G, and the β-δ discount function (Phelps and Pollak, 1968), $F_{\beta\text{-}\delta}$, are given by the following:

$$F_G(D) = \frac{1}{\left(1+k_G D\right)^g},$$

$$F_{\beta-\delta}(D) = 1 \quad \text{if } D = 0, \quad \text{and} \quad \omega_\beta \exp(-k_\beta D) \quad \text{if } D > 0.$$

It should be noted that the general hyperbolic discount function shown above is mathematically equivalent to the so-called q-exponential discount function (Cajueiro, 2006; Takahashi et al., 2008), which is given by the following:

$$F_q(D) = \frac{1}{[1+(1-q)\,k_q D]^{1/(1-q)}}.$$

The parameters of the general hyperbolic discount function and q-exponential discount function are related by the following;

$q = (g-1)/g$, and $k_q = k_G\, g$.

Denoting the animal's choice in trial t as c_t (=TS or TL), the likelihood of the animal's choices was given by,

$$L = \Pi_t\; p(c_t\,|\,\Omega_t) = p(c_1\,|\,\Omega_1)\; p(c_2\,|\,\Omega_2)\ldots\; p(c_N\,|\,\Omega_N),$$

where Ω_t denotes the magnitudes and delays for the rewards in trial t, and N the number of trials. For each discount function, model parameters were chosen to maximize the log likelihood (Pawitan, 2001), using a function minimization procedure in Matlab (Mathworks, Natick, MA, USA). Since the models with exponential and hyperbolic discount functions both include two parameters (β and k), these two models were compared using their log likelihood. This was carried out for the entire data from a given experiment as well as separately for each daily session. The general hyperbolic and β-δ discount functions included an additional parameter. Therefore, the Bayesian information criterion (BIC) was used to compare the performance of models with different numbers of parameters. BIC was computed as follows:

$$\mathrm{BIC} = -2\,\log L + m\log N,$$

where N is the number of trials and m the number of model parameters (e.g., 2 for the model with exponential or hyperbolic discount function). For the results obtained from monkey D in Experiments I-B and III-A, the process of parameter search failed to converge for the general hyperbolic discount function. In these two cases, the values of the parameters in the general hyperbolic discount functions were computed by estimating the parameters of the q-exponential discount function instead and converting them as described above. Since the general hyperbolic discount function and q-exponential discount function are mathematically equivalent, the log likelihood for the best parameters of these two models should be the same.

During Experiment III-A, the physical reward delay was given by $(n_Y + 4\,n_C)$ s, in which n_Y and n_C indicate the numbers of yellow and cyan disks, respectively. Temporally discounted values of rewards

associated with mixed clocks were computed without assuming that the animal accurately estimated the value of n_C. This was done by using the subjective delays for cyan disks, which were estimated as a free parameter in the maximum likelihood procedure described above. In other words, the subjective reward delays used to compute temporally discounted values were given by $(n_Y + D_C\, n_C)$ s, in which D_C refers to the subjective delay for one cyan disk.

RESULTS

EXPERIMENT I

In Experiment I, the reward delays ranged from 0.5 to 4.5 s, and the disks were always removed in a counter-clock direction (referred to as "ordered" clocks; **Figure 1**). In Experiment I-A, the ratio for the small and large reward was 1:2, whereas this ratio was 2:3 in Experiment I-B (**Table 1**). In both Experiments I-A and I-B, the animals almost always chose the large reward when the reward delays were 0.5 s for both targets. Monkey D never chose the small-reward target, whereas monkey J chose the small-reward target in 1% and 3% of the trials when the reward delays were both 0.5 s during Experiments I-A and I-B, respectively. Therefore, both animals displayed a clear preference for the large reward when both large and small rewards were immediately available. In contrast, collapsed across all possible reward delays, the probability that the animal chose the small-reward target through the entire Experiment I-A was 0.37 and 0.38 for monkeys D and J (**Table 2**). Therefore, both animals chose the small-reward target much more frequently, when the large reward was delayed. The corresponding values for Experiment I-B were 0.46 and 0.48, indicating that the animals were more likely to choose the small reward when its magnitude was more similar to that of the large reward. This difference is unlikely to reflect the difference in the animal's experience with the task, since the two animals were tested for Experiments I-A and I-B in different orders. Most importantly, both animals were increasingly more likely to choose the small-reward target as the delay for the small reward decreased and the delay for the large reward increased (**Figure 2**), and this was true for both Experiments I-A and I-B (data not shown). Therefore, the animal's choice between two different rewards was systematically affected by both the magnitudes and delays of rewards. This suggests that the animal's preference for a given reward might be parsimoniously described by its temporally discounted value.

To test whether the animal's behavior during the inter-temporal choice task was better accounted for by an exponential or hyperbolic discount function, we compared the log likelihood of the choice models based on these two discount functions (see Data Analysis). When the analysis was applied to the entire data set, the exponential discount function provided a better fit to the data for both animals (**Figure 2; Table 3**). This was true for both Experiments I-A and I-B. The results were similar, even when the same analysis was applied separately to the data from each daily session (**Figure 3**). The data from both animals were fit better by an exponential discount function, except for the 2 days in Experiment I-B in monkey J (**Figure 3**). When the model with the exponential discount function was fit to the entire data set from Experiment I-A, the maximum likelihood estimates of the discount rate were 0.39 s^{-1} for both animals. This value decreased to 0.29 and 0.32 for monkeys D and J in Experiment I-B (**Table 2**), although the results from individual daily sessions were somewhat more variable

Table 2 | Probability of choosing small reward target (TS) and the value of *k* parameters for the exponential (k_E) and hyperbolic (k_H) discount functions.

Experiment	Monkey D			Monkey J		
	p(TS)	k_E	k_H	*p*(TS)	k_E	k_H
I-A	0.37	0.39	1.24	0.38	0.39	1.18
I-B	0.46	0.29	0.61	0.48	0.32	0.74
I-A	0.40	0.13	0.23	0.41	0.13	0.23
II-B	0.42	0.14	0.25	0.39	0.12	0.21
III-A	0.66	0.27	0.57	0.66	0.49	1.32
III-A*	0.65	0.25	0.44	0.69	0.61	1.00
III-B	0.27	0.09	0.14	0.45	0.26	0.64

**indicates the results obtained from the Experiment III-A after excluding the trials with cyan disks.*

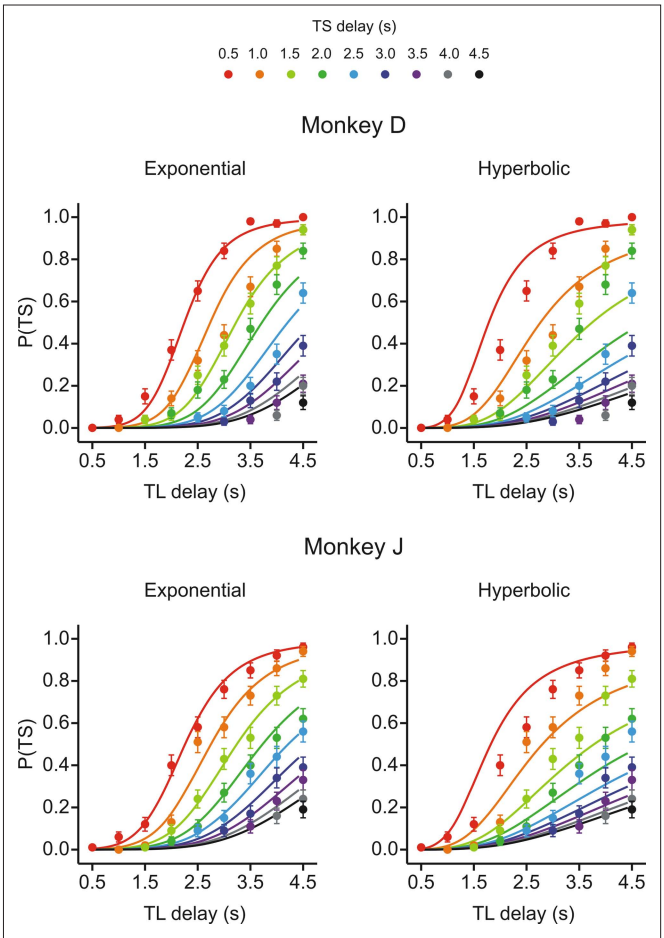


FIGURE 2 | Choice behaviors in Experiment I-A. Plots show the probability that the animal would choose the small-reward target as a function of the delays for the large-reward (TL) and small-reward (TS) targets, which are indicated in the abscissa and by different colored symbols, respectively. Lines indicate the predictions from the exponential (left) or hyperbolic (right) discount functions. Error bars, SEM.

Table 3 | Log likelihood (Bayesian information criterion, BIC) for exponential and hyperbolic discount functions.

Experiment	Monkey D		Monkey J	
	Exponential	Hyperbolic	Exponential	Hyperbolic
I-A	–1866.8 (**3750.4**)	–1990.5 (3997.8)	–2098.9 (**4214.7**)	–2197.0 (4410.9)
I-B	–2030.6 (**4078.0**)	–2138.6 (4294.0)	–1803.8 (**3624.3**)	–1887.4 (3791.7)
II-A	–1731.5 (3479.8)	–1566.7 (**3150.2**)	–2447.4 (4911.6)	–2319.9 (**4656.6**)
II-B	–1907.2 (3831.2)	–1742.9 (**3502.6**)	–2097.9 (4212.6)	–1927.8 (**3872.5**)
III-A	–1680.0 (**3384.8**)	–1895.6 (3816.0)	–1660.8 (**3346.3**)	–1700.1 (3424.9)
III-A*	–509.2 (**1033.0**)	–588.03 (1190.6)	–522.8 (1060.0)	–520.1 (**1054.7**)
III-B	–1464.6 (2946.0)	–1396.4 (**2809.6**)	–1720.4 (3457.6)	–1681.3 (**3379.4**)

*The discount functions with the better fit to the data are indicated by the BIC values in bold typeface. *indicates the results obtained from the Experiment III-A after excluding the trials with cyan disks.*

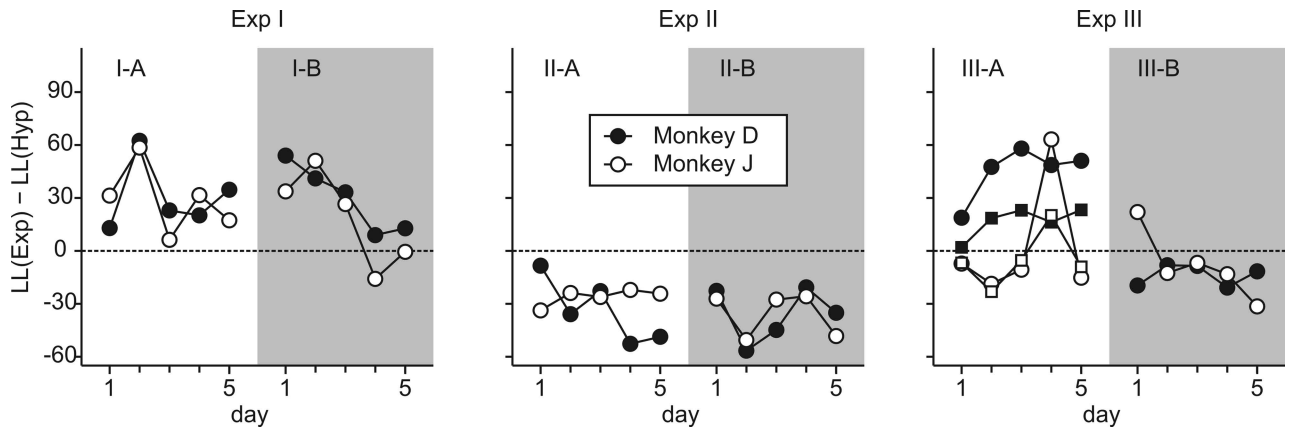


FIGURE 3 | Daily changes in the log likelihood ratio between the exponential and hyperbolic discount function. Positive (negative) values indicate that the exponential (hyperbolic) discount function accounted for the behavioral data better.

(**Figure 4**). For an exponential discount function, the temporally discounted value would be reduced by 50% for the delay equal to $-(1/k_E)\log 0.5$. Therefore, the approximate half-life for the subjective value of a reward was 2.2–2.4 s.

EXPERIMENT II

In Experiment II-A, the maximum reward delay was increased to 8 s. In addition, the positions of yellow disks in the clocks were randomized in Experiment II-A (referred to as "random" clocks; **Figure 1**). In Experiment II-B, only the ordered clocks were used to test whether the animal's behavior was affected by the manner in which the clocks represent the reward delays. As in Experiment I, the percentage of trials in which the animal chose the small-reward target was relatively small (<6%) when the reward delays were 0 s for both targets. In contrast, the overall probability that the animal would choose the small reward across all the reward delays used in Experiment II-A was 0.40 and 0.41 for monkeys D and J, respectively. The corresponding values for Experiment II-B were 0.42 and 0.39. Therefore, both animals chose the small reward targets much more frequently when the large reward was not available immediately. In addition, similar to the results in Experiment I, the animals chose the small reward increasingly more often as the delay for the large reward increased and as the delay for the small reward decreased in both Experiment II-A (**Figure 5**) and II-B (not shown).

In contrast to the results in Experiment I, the data from Experiment II were fit better by a hyperbolic discount function than by an exponential discount function. This was true for both Experiments II-A and II-B (**Table 3**). The slope and discount rate of a hyperbolic discount function decrease with delay. Consistent with this feature of hyperbolic discounting, the comparison between the data and the predictions from the best-fitting exponential discount function shows that the animals were particularly more likely to choose the small reward available without any delays than predicted by the exponential discount function (**Figure 5,** left). For Experiment II-A, the value of parameter k_H in the hyperbolic discount function was 0.23 for both animals. The corresponding values for Experiment II-B were 0.25 and 0.21. For hyperbolic discount function, the temporally discounted value is reduced by half when the reward delay is $1/k_H$. This implies that the half-life for the subjective value of reward was approximately 4.0 to 4.8 s. Moreover, the overall results from Experiments II-A and II-B were relatively similar (**Figure 4**). Therefore, the animals reliably extracted the information about reward delays from the visual displays regardless

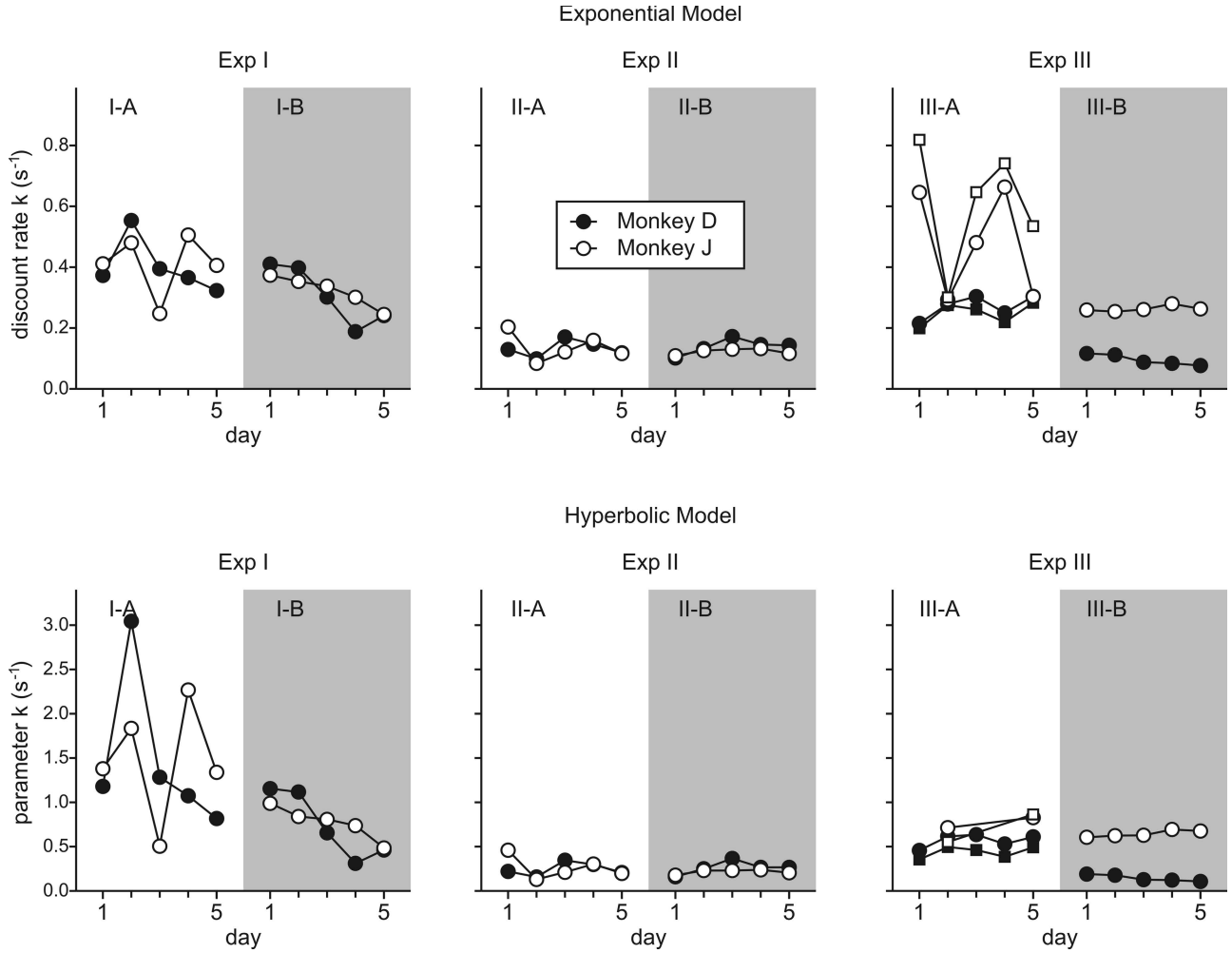


FIGURE 4 | Daily changes in the parameter *k* for the exponential (top) and hyperbolic (bottom) discount function. For Experiment III-A, squares indicate the values obtained from the trials in which the clocks did not include any cyan disks. For some sessions during Experiment III-A (monkey J), the model parameters did not converge for the hyperbolic discount function and therefore omitted.

of the manner in which the disks were arranged and removed in the clocks.

EXPERIMENT III

To test whether the animals can reliably estimate reward delays from the clocks without relying entirely on the number of disks, clocks used in Experiment III-A sometimes included a combination of yellow and cyan disks. Yellow and cyan disks increased the reward delay for a given target by 1 and 4 s/disk, respectively. Not surprisingly, when the animals were first exposed to mixed clocks, their choices were largely determined by the number of disks in each clock, regardless of their colors. For example, when the animals chose between a small reward with a 2-s delay and a large reward with a 5-s delay, they were at first more likely to choose the small reward if the 5-s delay was indicated by five yellow disks compared to when the same delay was indicated by a mixed clock with one yellow disk and one cyan disk (**Figure 6A**). This difference was gradually diminished during the preliminary training, especially for monkey D, whereas it was not completely eliminated for monkey J. We have also estimated the subjective delay associated with each cyan disk using a maximum likelihood procedure (see Data Analysis) for the data obtained during the preliminary training. Consistent with the changes in the choice probabilities, the subjective delays for cyan disks were initially relatively close to the delay for yellow disks (1 s) and gradually increased towards the correct value (4 s; **Figure 6B**). This was true regardless of whether the subjective delays were estimated using exponential or hyperbolic discount functions.

During Experiment III-A, the probability of choosing the small reward was 0.66 for both monkeys. To examine how the animal's choice was influenced by the delays for small and large rewards, we assumed that the subjective delay for a mixed clock was given by $(n_Y + D_C\, n_C)$ s, in which n_Y and n_C refer to the numbers of yellow and cyan disks and D_C was the subjective delay for a cyan disk. For exponential discount functions, the maximum likelihood estimate of D_C was 3.82 and 2.34 s for monkeys D and J, whereas corresponding values for hyperbolic discount functions were 4.17 and 2.46 s, respectively. This analysis showed that the animals

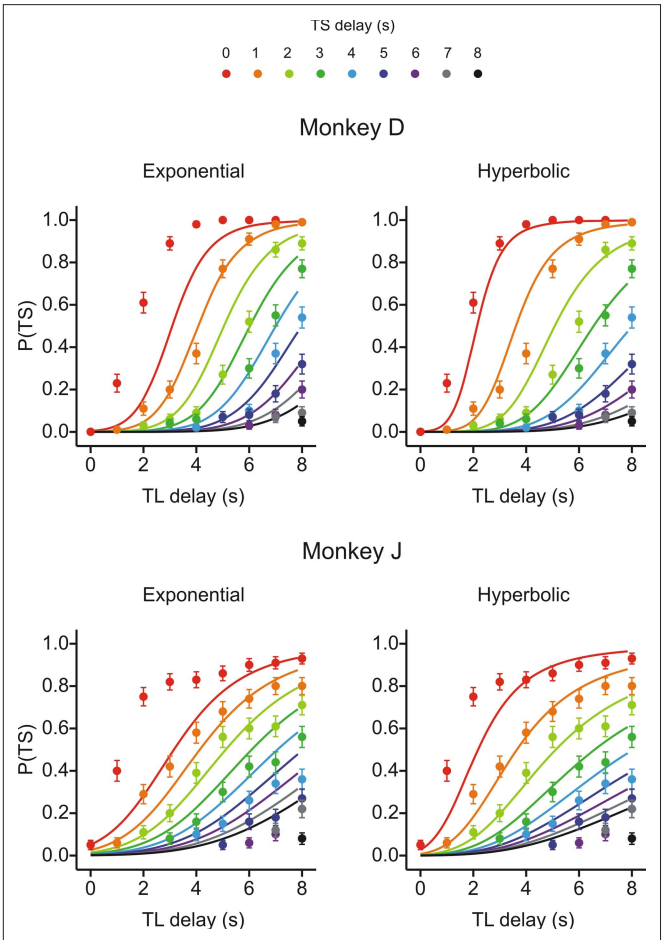


FIGURE 5 | Choice behaviors in Experiment II-A. Same format as in **Figure 2**.

tended to choose the small-reward target more frequently as the subjective delays for the large reward increased, and that this was relatively unaffected by the number of cyan disks used to indicate the delay for the large reward (**Figure 7**). In contrast to the results from Experiment II-A, however, the results from Experiment III-A were better fit by an exponential discount function. This was true, even when physical delays were used instead of subjective delays (not shown). Moreover, the exponential discount functions fit the results from monkey D better, even when the analysis was applied after excluding the trials with mixed clocks (**Table 3**). For monkey J, the hyperbolic discount function provided the better fit to the data when the trials with mixed clocks were excluded, but the difference in the log likelihood for the two discount functions was relatively small. For Experiment III-A, the discount rate estimated for the best-fitting exponential discount function was 0.27 and 0.49 s^{-1} for moneys D and J, respectively (**Table 2**).

After Experiment III-A, monkey J was tested for several months in a neurophysiological experiment using a subset of conditions included in Experiment III-A. The choice behavior of this animal during this period was better accounted for by a hyperbolic discount function than by an exponential discount function (61 of 69 sessions, 88.4%). To test whether the animal's discount function was irreversibly modified by the exposure to the mixed clocks, we have also re-tested both animals using only the clocks with yellow disks. During this experiment (III-B), the choice behaviors of both animals were better accounted for by hyperbolic discount functions (**Table 3; Figure 3**). These results suggest that the exponential discounting found in Experiment III-A was specifically related to the introduction of mixed clocks. Finally, we have fit the exponential and hyperbolic discount functions to the entire dataset collected from all the experiments described above. The results showed that the hyperbolic discount function provided a better fit to the data. The log likelihood ratio between the hyperbolic and exponential discount functions was 419.2 and 574.4 for monkeys D and J, respectively.

OTHER DISCOUNT FUNCTIONS

Both exponential and hyperbolic discount functions include only one free parameter, making it possible to compare their performance using the log likelihood directly (**Table 2**). When the number of parameters differs for different models, the likelihood tends to improve with the use of additional parameters. Therefore, we used the Bayesian information criterion to compare the performance of two additional discount functions, referred to as a general hyperbolic discount function (Mazur, 1987) and a β-δ discount function (Laibson, 1997). For the results obtained in Experiment I-A, an exponential discount function remained as the best model even when these additional discount functions were considered (**Table 4**). Exponential discount functions also best accounted for the behaviors of monkey D in Experiment I-B and Experiment III-A, whereas the results from monkey J in these two experiments were best accounted for by a general hyperbolic discount function. The data from monkey D in Experiments II-A was also most consistent with a β-δ discount function (**Table 5**), whereas a hyperbolic discount function still accounted for the data from monkey D in Experiment III-B. In all the remaining cases, the results were best accounted for by the general hyperbolic discount functions (**Table 4**), including four out of six cases in which the data were better accounted for by hyperbolic discount functions than by exponential discount functions. We have also tested a variant of hyperbolic discount function in which only the more delayed reward is discounted according to the difference in the delays for the two alternative rewards (Green et al., 2005), but found that this model did not account for the data better than the exponential or hyperbolic discount functions in any of the experiments.

DISCUSSION

MODELS OF TEMPORAL DISCOUNTING

Reward resulting from a particular action is often delayed in real life. In addition, a large number of laboratory studies have demonstrated that decision makers tend to choose an action leading to a more immediate reward delivery, when the difference in the reward magnitude is relatively small. This pattern of choice behavior can be parsimoniously accounted for by the concept of temporal discounting. Despite the methodological differences that often existed in various studies, the results from previous studies have been quite consistent and largely favored a hyperbolic discount function over an exponential discount function (Kalenscher and Pennartz, 2008; Kirby, 1997; Kirby and Maraković, 1995; Madden et al., 2003; Mazur,

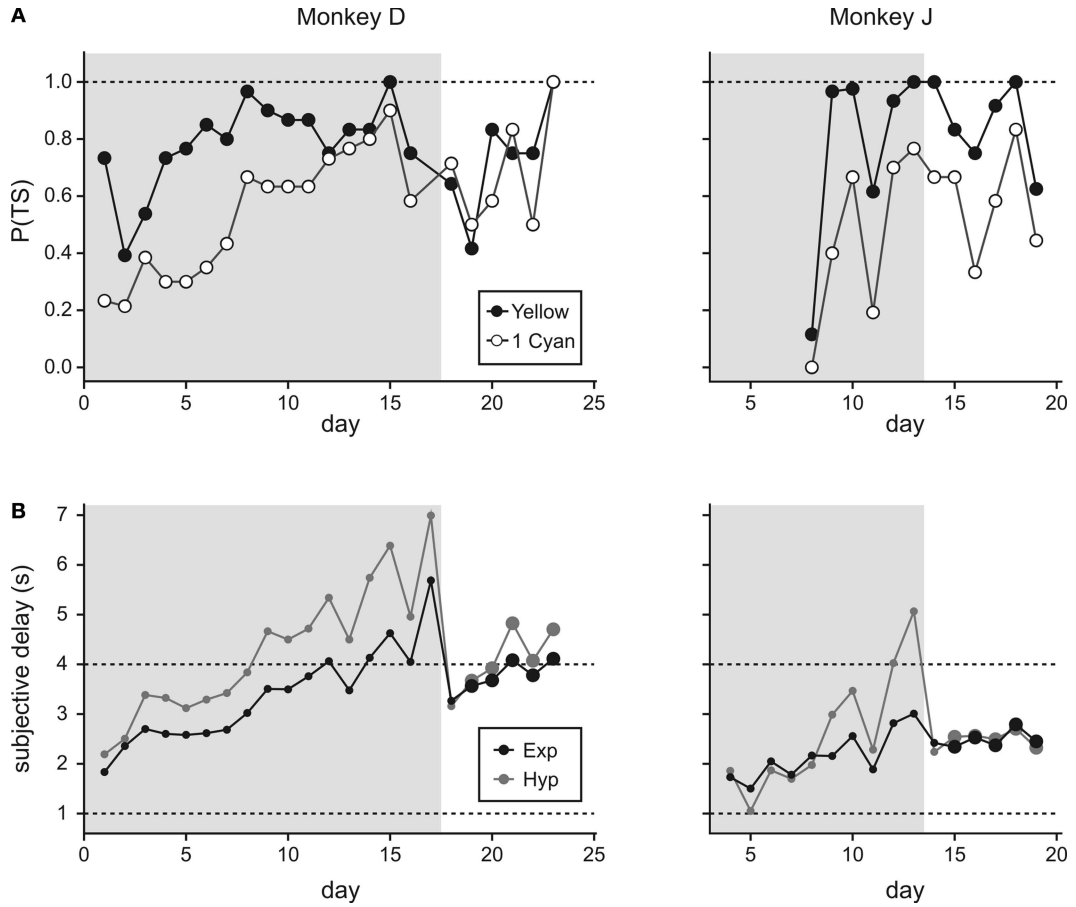


FIGURE 6 | Time course for the learning of delay information from the mixed clocks. (A) The plot shows daily changes in the probability that the animal would choose the small-reward target with a 2-s delay instead of the large-reward target with a 5-s delay. The delay for the large reward was indicated by either five yellow disks (filled circles) or by a combination of a yellow disk and a cyan disk (empty circles). **(B)** Daily changes in the subjective delay attributed to a single cyan disk. This was determined separately for exponential and hyperbolic discount functions. The actual delays corresponding to the yellow (1 s) and cyan (4 s) disks are indicated by the dotted lines. Large symbols show the results from the last 5 days that were included in the main analysis. Gray background indicates the period in which only a subset of conditions tested in Experiment III-A were used for the purpose of training. The results for monkey J during the first several days are missing, because cyan dots were introduced more gradually for this animal.

1987; Murphy et al., 2001; Myerson and Green, 1995; Rachlin et al., 1991; Simpson and Vuchinich, 2000; Woolverton et al., 2007).

For exponential discount function, the discount rate is constant, whereas for hyperbolic discount functions, discount rate decreases with reward delay. This hyperbolic discount function might arise due to the uncertainty in hazard rates (Luhmann et al., 2008; Sozou, 1998) or in the discount rate itself (Azfar, 1999). Alternatively, hyperbolic discounting may result from logarithmic time perception (Takahashi, 2005), since it has been shown that the individual variability in delay discounting might be related to time perception (Barkley et al., 2001; Reynolds and Schiffbauer, 2004; Wittmann et al., 2007). The logarithmic time perception implies that the subjective delay, τ, is given by the following function of physical delay, D.

$$\tau = a \log(1 + b\ D).$$

When a constant discount rate is applied to this subjective duration, then the resulting discount function for the physical delay for a particular reward would be a general hyperbolic discount function of the following form.

$$F(D) = \exp(-k\tau) = \frac{1}{(1 + k_G D)^g},$$

where k is the discount rate in the exponential discount function and $g = k\ a$. It has been shown that the general hyperbolic discount function tends to account for the behaviors of human decision makers better than the original hyperbolic discount function (Green and Myerson, 2004; Myerson and Green, 1995; Takahashi et al., 2008). Therefore, logarithmic time perception might provide a parsimonious explanation for the shape of discount function commonly observed in human decision makers.

In the present study, we have examined the choice behaviors of two rhesus monkeys during a novel inter-temporal choice task, and found that the results were consistent with exponential discount functions only in a minority of cases. First, the animals showed exponential discounting when the range of reward delays was

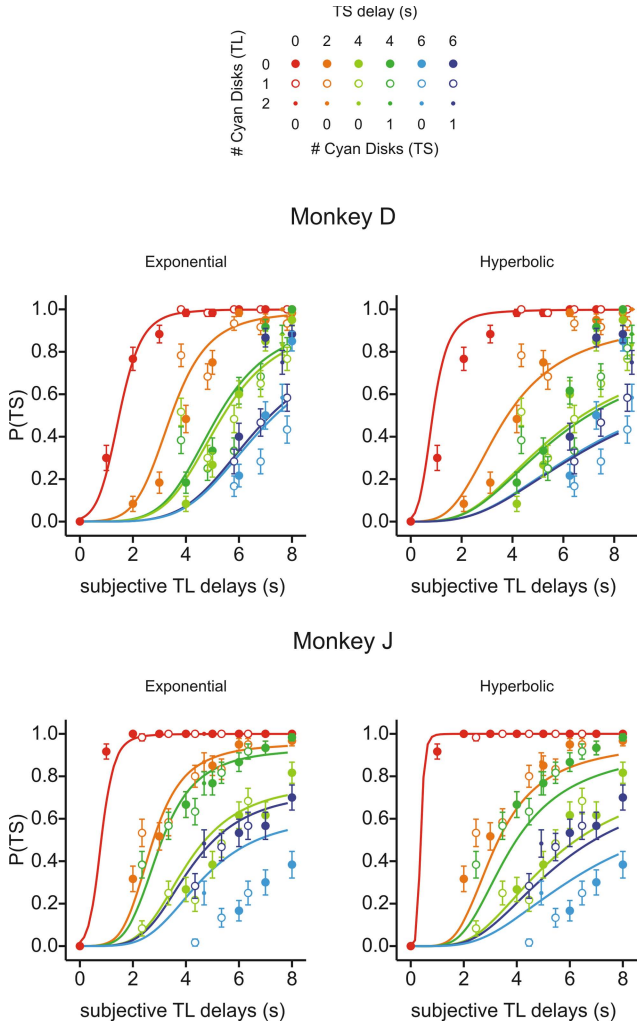


FIGURE 7 | Choice behaviors in Experiment III-A. The delays for large reward were calculated using the subjective delay for the cyan disk, whereas the physical reward delays and the number of cyan disks used for the small-reward and large-reward targets are indicated by the colors and sizes of the symbols. Lines indicate the predictions from the exponential (left) or hyperbolic (right) discount functions. Error bars, SEM.

Table 4 | Bayesian information criterion (BIC) for the general hyperbolic discount function and their parameters (k_G and g) and the best parameters for the q-exponential discount function (k_q and q).

Experiment	BIC	k_G	g	k_q	q
MONKEY D					
I-A	3757.6	0.02	16.92	0.41	0.94
I-B	4075.0	(–0.06)	(–3.78)	0.24	1.26
II-A	3128.0	0.40	0.73	0.29	–0.38
II-B	**3482.3**	0.42	0.73	0.31	–0.36
III-A	3377.2	(–0.05)	(–4.31)	0.22	1.23
III-B	2814.7	0.18	0.85	0.15	–0.18
MONKEY J					
I-A	4222.2	0.02	17.39	0.41	0.94
I-B	**3622.5**	0.07	5.36	0.36	0.81
II-A	**4576.9**	0.75	0.52	0.39	–0.93
II-B	**3794.5**	0.57	0.57	0.33	–0.74
III-A	**3294.8**	0.29	2.35	0.69	0.57
III-B	**3203.3**	0.23	1.88	0.43	0.47

The values in the parentheses were estimated indirectly from the q-exponential discount function. The bold typeface indicates that the data were best fit by this model.

Table 5 | Bayesian information criterion (BIC) for the β-δ discount function and their parameters (k_β and ω_β).

Experiment	BIC	k_β	ω_β
MONKEY D			
I-A	3758.8	0.68	0.65
I-B	40.86.4	0.75	0.48
II-A	**3079.3**	0.89	0.81
II-B	3531.1	0.89	0.81
III-A	3392.7	0.77	0.99
III-B	2821.8	0.92	0.90
MONKEY J			
I-A	4223.1	0.75	0.68
I-B	3632.8	0.51	0.73
II-A	4679.0	0.90	0.78
II-B	3921.4	0.90	0.80
III-A	3295.4	0.71	0.71
III-B	3304.5	0.79	0.79

The bold typeface indicates that the data were best fit by this model.

relatively small and did not include rewards without any delays, as in Experiment I. The range of reward delays during Experiment I was between 0.5 and 4.5 s, which was smaller than those used in the remaining experiments, and might not have been sufficient to observe a detectable change in the discount rate. Second, although both animals devalued delayed rewards hyperbolically during Experiment II, they returned to exponential discounting when the mixed clocks were introduced in Experiment III-A. For one animal (monkey D), the results from Experiment III-A still strongly favored an exponential discount functions even when the analysis was restricted to the trials including the clocks that were already familiar to the animals, namely, the clocks that included only yellow disks. For the other animal (monkey J), the results for the same subset of trials could not clearly distinguish between these two discount functions, although the hyperbolic discount function was slightly favored. Therefore, these results suggest that the exposure to a novel context might bias the animal to devaluate delayed rewards according to an exponential discount function. Indeed, when one of the animals was further tested using the mixed clocks during the subsequent neurophysiological experiment, its behavior was largely consistent with hyperbolic discounting (Kim et al., 2008). Therefore, it is also possible that the animals showed exponential discounting during Experiment I due to the lack of sufficient experience with the task used in the present study. Although the neural mechanisms involved in switching between exponential and hyperbolic discount function are unknown, it is possible that extensive

experience with a particular type of inter-temporal choice makes the process of decision making more habitual. Therefore, it would be important for future research to test whether the contributions of the prefrontal cortex and basal ganglia during inter-temporal choice change with experience.

TEMPORAL DISCOUNTING IN HUMANS AND ANIMALS

Although temporal discounting in both humans and other animals are well accounted for by hyperbolic discount functions, the value of the parameter k that controls the rate of discounting varies substantially across different animal species. For example, pigeons tend to discount the value of a delayed reward more steeply than rats and monkeys. The values of the parameter k_H in the hyperbolic discount function ranged from 0.3 to 2.24 s^{-1} for pigeons (Green et al., 2004, 2007; Mazur, 2000). If the subjective value of a delayed reward is given by a hyperbolic discount function, its half-life would be $1/k_H$. In other words, the value of a particular reward would be halved after the interval of $1/k_H$. Accordingly, pigeons would be roughly indifferent between an immediate reward and another reward which is twice as large but delayed by 0.4–3.3 s. The value of k_H parameter for rats ranged from 0.07 to 0.36 s^{-1} (Green et al., 2004; Richards et al., 1997), corresponding to the half-life of 2.8–14.3 s. In the present study, although the exact value of k_H varied according to the range of reward delays and the type of clocks used to signal reward delays, it was relatively stable and remained close to 0.2 s^{-1} during the course of Experiment II. This is comparable to the results obtained for the rats in previous studies. Similar results have been found in new world monkeys. For example, tamarins and marmosets are willing to wait on average for 7.9 and 14.4 s to choose the reward three times as large as the immediately available reward (Stevens et al., 2005). Assuming that they discount the value of delayed rewards hyperbolically, these results correspond to the k_H-values of 0.25 and 0.14 s, respectively. However, other studies have found substantially less steep discounting in rhesus monkeys. For example, when rhesus monkeys were trained to choose between different doses of cocaine injections, the value of k_H parameter was 0.008 s^{-1}, corresponding to the half-life of 125 s (Woolverton et al., 2007). In addition, rhesus monkeys become less risk-seeking as inter-trial intervals increase, when they choose between a small but certain reward and a large but uncertain reward (Hayden and Platt, 2007). It has been suggested that the animal's choice during this task might be determined by the temporally discounted value of a delayed reward expected in subsequent trials (Hayden and Platt, 2007). Under this assumption, the value of k_H parameter in the hyperbolic discount function that best fit the animal's choice behaviors was 0.033 s^{-1}. Thus, although the value of k_H parameter estimated in the present study was comparable to the previous estimates of other non-human primates, it was smaller than the values from the previous studies on rhesus monkeys.

Compared to the values of k_H obtained for non-human animals, the values of k_H estimated for the hyperbolic discount function in humans is substantially smaller, ranging from 4.0×10^{-4} to 0.027 $days^{-1}$ (Johnson and Bickel, 2002; Madden et al., 1997, 2003; Murphy et al., 2001; Takahashi et al., 2008), corresponding to the half-life of 37 to 2,500 days. Therefore, the half-life for the subjective value of delayed reward is many orders of magnitude larger in humans than in other animals. The difference in the rate of discounting between humans and animals may arise from a number of factors. For example, animal studies have always used the primary rewards, such as food or water, whereas human studies have largely relied on conditioned reinforcements, such as money. Indeed, human subjects show steeper discounting when tested with primary rewards compared to when they are tested with money (Estle et al., 2007; McClure et al., 2004, 2007). In addition, children and adolescents tend to show steeper discounting than in adults (Green et al., 1994; Olson et al., 2007; Scheres et al., 2006). This might be mediated at least in part by the gradual maturation of the prefrontal cortex (Kim et al., 2008; McClure et al., 2004, 2007). Indeed, apes and humans show similar rate of temporal discounting when tested under similar conditions (Rosati et al., 2007).

NEURAL CORRELATES OF TEMPORAL DISCOUNTING

An essential feature of inter-temporal choice is that the decision makers combine the information about the magnitude and delay of reward. Single-neuron recording studies in monkeys have found that the information about the magnitude of expected reward is distributed in a large number of cortical and subcortical areas, including the prefrontal cortex (Leon and Shadlen, 1999), posterior parietal cortex (Dorris and Glimcher, 2004; Platt and Glimcher, 1999; Sugrue et al., 2004), and basal ganglia (Hollerman et al., 1998; Kawagoe et al., 1998). In addition, the information about the immediacy of reward is also found in the prefrontal cortex (Sohn and Lee, 2007; Tsujimoto and Sawaguchi, 2005). Some neurons in the dorsolateral and orbitofrontal cortex encode the information about both the magnitude and delay of expected reward (Roesch and Olson, 2005a,b; Roesch et al., 2006). In most previous studies, however, the effects of reward magnitude and delay on neural activity were examined separately. In addition, many of these studies have examined the changes in neural activity related to the magnitude and delay of reward during the task in which the animals were instructed to produce a particular behavioral response in each trial. Accordingly, it was not necessary for the animals to compute the temporally discounted values of alternative rewards. In contrast, single-neuron recordings during the same inter-temporal choice used in the present study showed that the individual neurons in the dorsolateral prefrontal cortex encode the temporally discounted value of the reward expected from a particular target by combining the information about its magnitude and delay (Kim et al., 2008). Similarly, neuroimaging studies in human subjects have suggested that the dorsolateral prefrontal cortex might play an important role in evaluating the value of delayed reward (Luhmann et al., 2008; McClure et al., 2004, 2007; Tanaka et al., 2004). Whereas comparing the values of immediate and delayed rewards is likely to engage multiple brain areas, including the basal ganglia, amygdala, orbitofrontal cortex, insula, and posterior cingulate cortex (Cardinal et al., 2001; Kable and Glimcher, 2007; Luhmann et al., 2008; Roesch et al., 2006; Winstanley et al., 2004; Wittmann et al., 2007), how each of these multiple areas contributes to inter-temporal choice remains poorly understood. For example, whether the information about the magnitude and delay of reward is processed separately before these two different types of information are integrated in such areas as the prefrontal cortex is currently known. The behavioral task

used in the present study provides means to manipulate the delays of different rewards independently across trials, and therefore might be useful in elucidating the neural basis of temporal discounting and inter-temporal choice in animals.

ACKNOWLEDGEMENTS

We are grateful to Lindsay Carr and Bradley Davis for their help with the experiments. This research was supported by the grants from the NIH (RL1 DA024855, P01 NS048328, and P30 EY000785).

REFERENCES

Ainslie, G., and Herrnstein, R. J. (1981). Preference reversal and delayed reinforcement. *Anim. Learn. Behav.* 9, 476–482.

Azfar, O. (1999). Rationalizing hyperbolic discounting. *J. Econ. Behav. Organ.* 38, 245–252.

Barkley, R. A., Edwards, G., Laneri, M., Fletcher, K., and Metevia, L. (2001). Executive functioning, temporal discounting, and sense of time in adolescents with attention deficit hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). *J. Abnorm. Child Psychol.* 29, 541–556.

Cajueiro, D. O. (2006). A note on the relevance of the q-exponential function in the context of intertemporal choices. *Physica A* 364, 385–388.

Cardinal, R. N., Pennicott, D. R., Sugathapala, C. L., Robbins, T. W., and Everitt, B. J. (2001). Impulsive choice induced in rats by lesions of the nucleus accumbens core. *Science* 292, 2499–2501.

Dorris, M. C., and Glimcher, P. W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.

Estle, S. J., Green, L., Myerson, J., and Holt, D. D. (2007). Discounting of monetary and directly consumable rewards. *Psychol. Sci.* 18, 58–63.

Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preference: a critical review. *J. Econ. Lit.* 40, 351–401.

Green, L., Fisher, E. B. Jr, Perlow, S., and Sherman, L. (1981). Preference reversal and self control: choice as a function of reward amount and delay. *Behav. Anal. Lett.* 1, 43–51.

Green, L., Fry, A. F., and Myerson, J. (1994). Discounting of delayed rewards: a life-span comparison. *Psychol. Sci.* 5, 33–36.

Green, L., and Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychol. Bull.* 130, 769–792.

Green, L., Myerson, J., Holt, D. D., Slevin, J. R., and Estle, S. J. (2004). Discounting of delayed food rewards in pigeons and rats: is there a magnitude effect? *J. Exp. Anal. Behav.* 81, 39–50.

Green, L., Myerson, J., and Macaux, E. W. (2005). Temporal discounting when the choice is between two delayed rewards. *J. Exp. Psychol. Learn. Mem. Cogn.* 31, 1121–1133.

Green, L., Myerson, J., Shah, A. K., Estle, S. J., and Holt, D. D. (2007). Do adjusting-amount and adjusting-delay procedures produce equivalent estimates of subjective value in pigeons? *J. Exp. Anal. Behav.* 87, 337–347.

Hayden, B. Y., and Platt, M. L. (2007). Temporal discounting predicts risk sensitivity in rhesus macaques. *Curr. Biol.* 17, 49–53.

Hollerman, J. R., Tremblay, L., and Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J. Neurophysiol.* 80, 947–963.

Johnson, M. W., and Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *J. Exp. Anal. Behav.* 77, 129–146.

Kable, J. W., and Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.

Kalenscher, T., and Pennartz, C. M. A. (2008). Is a bird in the hand worth two in the future? The neuroeconomics of intertemporal decision-making. *Prog. Neurobiol.* 84, 284–315.

Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416.

Kim, S., Hwang, J., and Lee, D. (2008). Prefrontal coding of temporally discounted values during inter-temporal choice. *Neuron* 59, 161–172.

Kirby, K. N. (1997). Bidding on the future: evidence against normative discounting of delayed rewards. *J. Exp. Psychol. Gen.* 126, 54–70.

Kirby, K. N., and Maraković, N. N. (1995). Modeling myopic decisions: evidence for hyperbolic delay-discounting within subjects and amounts. *Organ. Behav. Hum. Decis. Process.* 64, 22–30.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Q. J. Econ.* 112, 443–477.

Leon, M. I., and Shadlen, M. N. (1999). Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24, 415–425.

Luhmann, C. C., Chun, M. M., Yi, D.-J., Lee, D., and Wang, X.-J. (2008). Neural dissociation of delay and uncertainty in inter-temporal choice. *J. Neurosci.* 28, 14459–14466.

Madden, G. J., Begotka, A. M., Raiff, B. R., and Kastern, L. L. (2003). Delay discounting of real and hypothetical rewards. *Exp. Clin. Psychopharmacol.* 11, 139–145.

Madden, G. J., Petry, N. M., Badger, G. J., and Bickel, W. K. (1997). Impulsive and self-control choices in opioid-dependent patients and non-drug-using control participants: drug and monetary rewards. *Exp. Clin. Psychopharmacol.* 5, 256–262.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In Quantitative Analyses of Behaviors: the Effect of Delay and of Intervening Events on Reinforcement Value, Vol. 5, M. L. Commons, J. E. Mazur, J. A. Nevin and H. Rachlin, eds (Hillsdale, Erlbaum), pp. 55–73.

Mazur, J. E. (2000). Tradeoffs among delay, rate, and amount of reinforcement. *Behav. Processes* 49, 1–10.

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2007). Time discounting for primary rewards. *J. Neurosci.* 27, 5796–5804.

McClure, S. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507.

Murphy, J. G., Vuchinich, R. E., and Simpson, C. A. (2001). Delayed reward and cost discounting. *Psychol. Rec.* 51, 571–588.

Myerson, J., and Green, L. (1995). Discounting of delayed rewards: models of individual choice. *J. Exp. Anal. Behav.* 64, 263–276.

Olson, E. A., Hooper, C. J., Collins, P., and Luciana, M. (2007). Adolescents' performance on delay and probability discounting tasks: contributions of age, intelligence, executive functioning, and self-reported externalizing behavior. *Pers. Individ. Dif.* 43, 1886–1897.

Pawitan, Y. (2001). In All Likelihood: Statistical Modelling and Inference Using Likelihood. New York, Oxford University Press.

Phelps, E. S., and Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *Rev. Econ. Stud.* 35, 185–199.

Platt, M. L., and Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.

Rachlin, H., and Green, L. (1972). Commitment, choice and self-control. *J. Exp. Anal. Behav.* 17, 15–22.

Rachlin, H., Raineri, A., and Cross, D. (1991). Subjective probability and delay. *J. Exp. Anal. Behav.* 55, 233–244.

Reynolds, B. (2006). A review of delay-discounting research with humans: relations to drug use and gambling. *Behav. Pharmacol.* 17, 651–667.

Reynolds, B., and Schiffbauer, R. (2004). Measuring state changes in human delay discounting: an experiential discounting task. *Behav. Processes* 67, 343–356.

Richards, J. B., Mitchell, S. H., de Wit, H., and Seiden, L. S. (1997). Determination of discount functions in rats with an adjusting-amount procedure. *J. Exp. Anal. Behav.* 67, 353–366.

Roesch, M. R., and Olson, C. R. (2005a). Neuronal activity dependent on anticipated and elapsed delay in macaque prefrontal cortex, frontal and supplementary eye fields, and premotor cortex. *J. Neurophysiol.* 94, 1469–1497.

Roesch, M. R., and Olson, C. R. (2005b). Neuronal activity in primate orbitofrontal cortex reflects the value of time. *J. Neurophysiol.* 94, 2457–2471.

Roesch, M. R., Taylor, A. R., and Schoenbaum, G. (2006). Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* 51, 509–520.

Rosati, A. G., Stevens, J. R., Hare, B., and Hauser, M. D. (2007). The evolutionary origins of human patience: temporal preferences in chimpanzees, bonobos, and human adults. *Curr. Biol.* 17, 1663–1668.

Samuelson, P. A. (1937). A note on measurement of utility. *Rev. Econ. Stud.* 4, 155–161.

Scheres, A., Dijkstra, M., Ainslie, E., Balkan, J., Reynolds, B., Sonuga-Barke, E., and Castellanos, F. X. (2006). Temporal and probabilistic discounting of rewards in children and adolescents: effects of age and ADHD symptoms. *Neuropsychologia* 44, 2092–2103.

Simpson, C. A., and Vuchinich, R. E. (2000). Reliability of a measure of temporal discounting. *Psychol. Rec.* 50, 3–16.

Sohn, J.-W., and Lee, D. (2007). Order-dependent modulation of directional signals in the supplementary and pre-supplementary motor areas. *J. Neurosci.* 27, 13655–13666.

Sozou, P. D. (1998). On hyperbolic discounting and uncertain hazard rates. *Proc. R. Soc. Lond., B, Biol. Sci.* 265, 2015–2020.

Stevens, J. R., Hallinan, E. V., and Hauser, M. D. (2005). The ecology and evolution of patience in two new world monkeys. *Biol. Lett.* 1, 223–226.

Strotz, R. H. (1955–1956). Myopia and inconsistency in dynamic utility maximization. *Rev. Econ. Stud.* 23, 165–180.

Sugrue, L. P., Corrado, G. S., and Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.

Takahashi, T. (2005). Loss of self-control in intertemporal choice may be attributable to logarithmic time-perception. *Med. Hypotheses* 65, 691–693.

Takahashi, T., Oono, H., and Radford, M. H. B. (2008). Psychophysics of time perception and intertemporal choice models. *Physica A* 387, 2066–2074.

Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., and Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* 7, 887–893.

Tsujimoto, S., and Sawaguchi, T. (2005). Neuronal activity representing temporal prediction of reward in the primate prefrontal cortex. *J. Neurophysiol.* 93, 3687–3692.

Winstanley, C. A., Theobald, D. E. H., Cardinal, R. N., and Robbins, T. W. (2004). Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J. Neurosci.* 24, 4718–4722.

Wittmann, M., Leland, D. S., Churan, J., and Paulus, M. P. (2007). Impaired time perception and motor timing in stimulant-dependent subjects. *Drug Alcohol Depend.* 90, 183–192.

Woolverton, W. L., Myerson, J., and Green, L. (2007). Delay discounting of cocaine by rhesus monkeys. *Exp. Clin. Psychopharmacol.* 15, 238–244.

Conflict of Interest Statement: Authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 April 2009; paper pending published: 17 May 2009; accepted: 01 June 2009; published online: 11 June 2009.
*Citation: Hwang J, Kim S and Lee D (2009) Temporal discounting and inter-temporal choice in rhesus monkeys. Front. Behav. Neurosci. (2009) **3**:9. doi:10.3389/neuro.08.009.2009*

frontiers in
BEHAVIORAL NEUROSCIENCE

REVIEW ARTICLE
published: 23 October 2009
doi: 10.3389/neuro.08.039.2009

Temporal decision-making: insights from cognitive neuroscience

Christian C. Luhmann*

Department of Psychology, Stony Brook University, Stony Brook, NY, USA

Edited by:
Daeyeol Lee, Yale University School of Medicine, USA

Reviewed by:
Soyoun Kim, Yale University School of Medicine, USA
Paul Phillips, University of Washington, USA

****Correspondence:***
Christian C. Luhmann, Department of Psychology, Stony Brook University, Stony Brook, NY 11794-25001, USA. e-mail: christian.luhmann@stonybrook.edu

Decisions frequently have consequences that play out over time and these temporal factors can exert strong influences on behavior. For example, decision-makers exhibit delay discounting, behaving as though immediately consumable goods are more valuable than those available only after some delay. With the use of functional magnetic resonance imaging, we are now beginning to characterize the physiological bases of such behavior in humans and to link work on this topic from neuroscience, psychology, and economics. Here we review recent neurocognitive investigations of temporal decision-making and outline the theoretical picture that is beginning to take shape. Taken as a whole, this body of work illustrates the progress made in understanding temporal choice behavior. However, we also note several questions that remain unresolved and areas where future work is needed.

Keywords: reward, neuroeconomics, decision making, intertemporal choice

INTRODUCTION

Decision-makers are frequently faced with choices that differ in the timing of their consequences. Such intertemporal choices require shrewd decision-makers to consider, not only what they want, but when they want it. For example, when asked to deliver a guest lecture, your response is likely to depend strongly on whether the lecture is to be delivered relatively soon or in the more distant future. More gravely, decisions about whether to refinance one's mortgage (Harding, 2000) and about whether governments should spend money to protect the environment (Dasgupta, 2008; Hardisty and Weber, 2009) can be characterized as intertemporal choices. Furthermore, abnormalities in intertemporal choice behavior have been associated with an array of undesirable behavior including drug addiction (Kirby and Petry, 2004; Rossow, 2008). Given the relevance of intertemporal choice, it is clear that we have much to gain by understanding how intertemporal choices are made, what factors influence intertemporal choices, and what is responsible for aberrations in intertemporal choice in some patient populations.

Like much of decision-making, intertemporal choice has long been the province of economics. Work from this field has provided both normative guidelines for intertemporal choice and the theoretical tools to evaluate observed behavior. Empirical support has come primarily from psychology and has, as it often does, focused on decision-makers' deviations from the prescriptions of economics. With the recent interest in utilizing neurobiological techniques to understand decision making behavior (Glimcher, 2003; Glimcher et al., 2008), particularly functional magnetic resonance imaging (fMRI), we are in the position to observe the operation of the processes responsible for intertemporal decisions, processes that are extremely difficult to evaluate using behavior alone. Here, we review recent work on intertemporal choice with a focus on studies involving humans, the majority of which have utilized a combination of behavior, fMRI, and quantitative economic theory.

DELAY DISCOUNTING

The majority of intertemporal choice studies have been designed to explore delay discounting, the robust finding that animals, including humans, behave as though immediately consumable goods are more valuable than those only available after some delay. This phenomenon is so powerful that decision-makers frequently forgo delayed rewards in favor of immediate rewards even when the delayed rewards are objectively more valuable. For example, a decision-maker might choose $100 delivered immediately over $200 to be delivered in 3 years. Such a choice is said reflect the subjective value of the $200 option, discounted according to the associated 3-year delay. The sway of negative events is similarly blunted by delay. The idea of working on your taxes next month seems less unpleasant than the prospect of working on them tonight.

Economics has viewed delay discounting from within the framework of discounted utility theory (Samuelson, 1937) according to which the subjective value of goods drops by a fixed percentage (frequently referred to as the discount rate) for each unit of time that those goods are delayed. If a decision-maker discounts the future at a rate of 10% annually, then $100 available in a year is only worth $90 right now. That same reward offered in 5 years is only worth $59. If this drop in subjective value is plotted over time, the resulting discounting curve is exponential in shape.

Behavioral work on delay discounting has primarily focused on two major facets of the phenomenon. First, it appears that animals, including humans, do not discount exponentially. Given that such behavior is arguably non-normative, this possibility has generated a large body of behavioral data (Ainslie and Herrnstein, 1981; Loewenstein and Thaler, 1989; Ainslie, 1992; Green et al., 1994a; Kirby and Herrnstein, 1995; Rachlin, 1995; Kirby, 1997) nearly all of which demonstrates that decision-makers behave as though their discount rate declines as rewards are pushed further into the future. Waiting 2 years for a reward might be worth 10% less than waiting 1 year, but waiting 4 years for a reward might be worth only 5% less than waiting 3 years. Such discounting is

referred to as hyperbolic or quasi-hyperbolic and is blamed for a variety of unwanted behavior (Ainslie, 2001) all stemming from the fact that hyperbolic discounters make one set of choices about rewards in the distant future only to reverse their preferences as those same rewards draw near.

Second, work has focused on the rate of discounting itself and has found that discounting rates vary across individuals and contexts, and are sometimes unreasonably extreme. For example, in two of the more well-cited studies (Hausman, 1979; Gately, 1980), discount rates were estimated based on the purchase price and operating costs of home appliances. The estimated rates were shown to be significantly greater than typically assumed by economists (anywhere from 25 to 300% per year which is obviously well above the rates at which consumers borrow and invest). Thus, there has been a significant effort to characterize the rate at which various populations discount rewards. For example, children (Green et al., 1994b; Scheres et al., 2006), including those with attention deficit hyperactivity disorder (ADHD, Barkley et al., 2001), alcoholics (Vuchinich and Simpson, 1998; Petry, 2001), smokers (Bickel et al., 1999; Reynolds et al., 2004), cocaine and heroin addicts (Coffey et al., 2003; Kirby and Petry, 2004), and compulsive gamblers (Holt et al., 2003) all discount at a faster rate than healthy adults; they exhibit a relative inability to wait for rewards. In contrast, older adults (Green et al., 1994b) and those with a higher IQ (Shamosh and Gray, 2008) have been shown to discount at a slower rate; they exhibit relative patience.

Utilizing classic intertemporal choice tasks, recent work in cognitive neuroscience has begun to address the neural mechanisms associated with delay discounting. One basic question that this field is uniquely suited to address is what distinguishes those occasions on which decision-makers choose to wait from those occasions on which they choose immediate rewards. That is, what leads to patient and impatient choices? Wittmann et al. (2007) utilized fMRI and a standard intertemporal choice task (though with completely hypothetical rewards). Based on subjects' choices, the magnitude of the immediate option was adjusted incrementally to find the point at which that particular decision-maker would be indifferent between the immediate and delayed options. Trials on which the delayed option was chosen (patient choices) were then compared to trials on which the immediate option was chosen (impatient choices). This contrast yielded a network of brain regions that included bilateral posterior insular cortex, left posterior cingulate, as well as temporal and parietal regions. Interestingly, no regions appeared to exhibit greater activity when choosing the immediate reward. This study also observed higher levels of activity in the striatum when subjects were asked about rewards to be delivered in the near future (<1 year) than when they were asked about delayed rewards in the distant future (≥1 year).

These findings, particularly the involvement of the insula, extends previous work (Tanaka et al., 2004) on reward-based learning that has shown a delay-related gradient running from anterior to posterior insular cortex. When subjects learn to make sequences of actions to acquire monetary rewards, anterior and inferior portions of insular cortex appear to be differentially involved in producing reward-prediction error signals related to immediate rewards. In contrast, posterior and superior portions of insular cortex appear to serve this same function when learning about more delayed rewards. Taken together, these studies suggest that decisions involving increased delay are associated with activity in the posterior insula.

Though insular cortex has been implicated in a variety of sensory, cognitive, and emotional processes, there are intriguing intersections between these decision-related findings and previous work on pain. Mirroring the delay-related gradient in insular cortex, work on pain perception has found a similar differentiation along the anterior–posterior axis with more posterior portions associated with the more sensory aspects of pain processing and the more anterior portions associated with the more cognitive or emotional aspects of pain (Singer et al., 2004). For example, the anticipation of impending pain elicits activity in more anterior portions of the insula than the subsequent pain experience itself (Ploghaus et al., 1999). More generally, insular cortex has been associated with drug addiction, a condition marked by, and presumably maintained by, pronounced difficulties in weighing short-term gains (e.g., drug-use) against long-term outcomes (e.g., jail, health). For example, cocaine addicts exhibit structural abnormalities in insular cortex including white matter legions (Bartzokis et al., 1999) and a reduction in gray matter (Franklin et al., 2002). In particular, insular activity appears to be closely related to drug craving (Garavan et al., 2000; Kilts et al., 2001; Schneider et al., 2001; Bonson et al., 2002; Brody et al., 2002; Myrick et al., 2004; Wang et al., 2007) and relapse (Paulus et al., 2005; Naqvi et al., 2007). Abnormal insular activation has also been found in individuals with ADHD (Ernst et al., 2003; Rubia et al., 2009, but see Scheres et al., 2006) and conduct disorder (Rubia et al., 2009) who, like addicts, exhibit diminished patience in delay discounting tasks (Barkley et al., 2001).

Other work has sought to explore what neural features distinguish patient individuals from impatient individuals. Activity in the striatum has been shown (Hariri et al., 2006) to predict discounting rates across individuals such that larger but less discriminative reward prediction errors are associated with diminished patience. This same pattern of striatal activity was recently shown (Forbes et al., 2009) to be associated with genetic variation in genotypes thought to influence the release, availability, and signaling strength of dopamine (DAT1, DRD2, and DRD4). There also appears (Boettiger et al., 2007) to be a relationship between polymorphic variation of the catechol-*O*-methyltransferase (COMT) gene and delay discounting with the $158^{Val/Val}$ genotype being associated with diminished patience and hyperactivity in dorso-lateral prefrontal cortex (DLPFC) and posterior parietal cortex (with no apparent effects in the striatum). The $158^{Val/Val}$ genotype has also been linked to perseverative errors during reinforcement learning tasks which have been attributed to reduced levels of dopamine in prefrontal cortex (Egan et al., 2001; Frank et al., 2007). Lastly, there is intriguing evidence (Yacubian et al., 2007) that variation of COMT and DAT may interact to modulate complex patterns of activity in the striatum during reward processing.

More recent neurocognitive work has explored delay discounting using more fine-grained analytical methods. For example, an fMRI study by Kable and Glimcher (2007) utilized what they refer to as a "neurometric" approach in order to explore brain regions whose activity varied with the subjective value of various monetary rewards. In their study, decision-makers completed a standard intertemporal choice task in which they chose between pairs of rewards

that varied both in their magnitude and in when they would be delivered. For example, a subject might choose between $20 to be delivered that day and $40 to be delivered 30 days later. By observing how changes in delay and reward magnitude modulated behavioral choices, the discounting curves underlying subjects' choices could be reconstructed (Myerson and Green, 1995). These reconstructed curves could then be used to compute the idiosyncratic subjective value of any arbitrary reward–delay combination. To explore the neural representation of subjective value, these authors investigated what, if any, brain regions exhibited activity that corresponded to these subjective value functions. The results indicated that the ventral striatum, medial prefrontal cortex, and posterior cingulate cortex exhibited such a pattern of activity. Variation in these regions' activity was better predicted by subjective value than by several related quantities (e.g., delay, reward magnitude, choice) and closely mirrored individual differences in subjects' discounting rates.

One problem in relating the neurobiological work on human intertemporal choice with the currently larger literature on non-human animals (Cardinal, 2006) is that the delays typically utilized in human tasks (e.g., days, months, years) are significantly longer than those used with other animals (e.g., less than a minute). In an attempt to bridge this gap, recent work (Gregorios-Pippas et al., 2009) has investigated human delay discounting utilizing relatively short delays. Subjects completed a delay discounting task involving delays ranging from 4 to 14 s. Unlike other studies, subjects did not choose between rewards. Instead, subjects were presented with visual cues about impending, temporally delayed rewards with the identity of the cue reliably signaling the length of the delay (although rewards were only paid out at the conclusion of the study). The results reveal that the visual cues elicited graded increases in the ventral striatum (the focus of this study) such that cues associated with shorter delays (thus indicating more subjectively valuable rewards) elicited greater striatal activity. Furthermore, these neural responses mirrored individual subjects' patterns of choice in a separate behavioral choice task. Intriguingly, these cue-induced neural responses tended to decrease as subjects' total accumulated reward increased, suggesting a potential neural analog of diminishing marginal utility (Edwards, 1954). Taken together with the work of Kable and Glimcher (2007), these results suggest that activity in the ventral striatum, along with portions of anterior and posterior medial cortex, exhibits a graded signal that represents the subjective value of delayed rewards. This, along with related pharmacological work demonstrating the role of dopamine in delay discounting (Montague and Berns, 2002; Kheramin et al., 2004; Winstanley et al., 2004; Phillips et al., 2007; Moustafa et al., 2008) suggests that striatal–cortical circuitry is likely to be a key player in the valuation of delayed rewards and a target for therapeutic work on disorders characterized by impulsive behavior (Rahman et al., 2001).

MECHANISMS UNDERLYING DELAY DISCOUNTING

Despite the large and growing literature describing the neural signals that represent the idiosyncratic, subjective value of delayed rewards, we ultimately wish to understand the origin of these value signals, their variation across healthy individuals, and their aberrations in clinical populations. If the subjective value of delayed rewards underlies impatient choices occur when, it seems reasonable to ask why they are not valued more strongly. With a better understanding of how subjective value is computed, we would be in a much better position to design both diagnostic instruments and treatments.

Theorizing in psychology has emphasized the idea that choices between delayed rewards (as well as other types of choices) involve a competition between "the passions" and reason (Ainslie, 1975, 2001; Schelling, 1984; Loewenstein, 1996; Soman et al., 2005). Some (Metcalfe and Mischel, 1999) suggested that this competition is between rational, cognitive processes and irrational, emotional processes. Others (Thaler and Sheffrin, 1981; Ainslie, 1992; McClure et al., 2004) have suggested a competition between a prudent, far-sighted process concerned with overall welfare and a greedy, myopic process more concerned with immediate gains. Regardless of the details, what is common across these accounts is the belief that the relative value of waiting and immediate gratification results from a struggle between mutually incompatible drives. If the prudent, rational, cognitive system is able to suppress the greedy, myopic, emotional system, then the decision-maker will see the wisdom of waiting and exhibit relative patience. Otherwise, the emotional system will dominate, producing a strong aversion to waiting and relative impatience.

Several broad literatures have yielded data in support of this general proposal, though it is predominantly indirect in nature. For example, there appear to be large inter-species differences in delay discounting, though the comparison is plagued by methodological differences which make interpretation difficult. Compared to humans, non-human animals exhibit greater impatience for delayed rewards (Logue et al., 1986). Even monkeys, which exhibit relative cognitive sophistication, will choose immediate rewards over significantly large delayed rewards even when the delay is only several seconds (Kim et al., 2008; Hwang et al., 2009). For pigeons, the situation is even more dramatic, with immediate rewards losing approximately 50% of their value when delayed by a single second (Mazur, 1984). To the extent that one associates prudent, rational control of behavior with frontal lobe function (and to the extent that species differences are not a methodological artifact), these differences across species suggestively mirror the phylogenetic development of frontal cortex (Fuster, 2002). Similarly, delay discounting behavior appears to follow a systematic trajectory over the course of the human lifespan (Green et al., 1994b, 1999b). Relative to young adults, children exhibit significantly less patience for delayed rewards. Here again, this developmental trend is generally consistent with the ontogenetic changes taking place in frontal cortex (Sowell et al., 1999; Fuster, 2002). A related and growing literature has also demonstrated a strong relationship between overall intellectual ability and patience (Mischel et al., 1989; Burks et al., 2009). Indeed, a recent meta-analysis of 24 relevant delay discounting studies ultimately concluded that higher IQ is reliably associated with greater patience (Shamosh and Gray, 2008).

Two related studies by McClure and colleagues (McClure et al., 2004; McClure et al., 2007) provide the first neural evidence to support the idea that delay discounting involves a dual-process competition. Specifically, this group tested Laibson's beta-delta account of discounting (Laibson, 1997) which posits two components: one concerned with immediate rewards (beta) and one concerned with delayed rewards (delta). Using a traditional delay discounting task, decision-makers were asked to choose between

pairs of rewards of varying sizes to be delivered at various points in the future. To isolate neural activity associated with the beta component, trials involving an immediate reward were compared with trials that involved only delayed rewards. This comparison revealed several brain regions that exhibited greater activity when faced with an immediate reward. These regions included ventral striatum, medial prefrontal cortex, and posterior cingulate cortex. To isolate the delta component, brain regions that were activated by the task, but that did not distinguish between the different delays were selected. This resulted in a broad network of regions including dorsolateral and ventrolateral portions of prefrontal cortex as well as lateral orbital frontal cortex.

In isolation, these contrasts are relatively coarse, especially given how well-specified the theory being tested is. Critically, however, further analyses demonstrated that the relative activity in these two networks was predictive of subjects' choices. When faced with a choice between an immediate reward and a delayed reward, choosing the immediate reward was associated with increased activity in the beta network and decreased activity in the delta network. Choices for the delayed reward were associated with the opposite pattern. A recent replication of this study generalized these findings to decisions involving primary rewards (juice and water) and shorter delays (up to 20 min). Contrasts revealed similar networks of brain regions associated with the beta and delta components. Furthermore, choices were again found to be predicted by the relative activity in the two networks, this time utilizing more rigorous regression analyses.

It is interesting to note that the anatomical details of the beta and delta networks grossly mirror the psychologist's conceptualization. The greedy, irrational, myopic drive is embodied by portions of the evolutionarily older limbic system whereas the rational, patient drive is embodied by the relatively recent frontal cortex (particularly DLPFC). The relationship between activity in these networks and choice behavior also matches the expected competition. To the extent that DLPFC can suppress the relatively insolent limbic system, the decision-maker will make choices that are beneficial in the long-run. If the passionate limbic system can overcome the DLPFC's control, the decision-maker makes impatient choices.

Along with a fairly well-entrenched theoretical story, investigations into the mechanisms underlying delay discounting face another hurdle; such investigations are simply difficult to conduct. The above investigation of the beta and delta networks is illustrative. Though these findings are consistent with the theoretical framework proposed by its authors, this interpretation has been criticized (Kable and Glimcher, 2007) as being consistent with alternative formulations. Recall the investigations into the neural representation of subjective value reviewed above. These studies found that activity in a highly similar set of regions was related to both delay and reward magnitude (and their combination, subjective value). Thus, it is possible that the ostensible beta network exhibited greater activity for immediate rewards simply because the immediate reward represented an option with a large subjective value. Furthermore, if choices are made on the basis of subjective value, then it is not surprising that activity in the beta should be related to choice behavior (see regression analyses, McClure et al., 2007). Below we outline other obstacles.

DELAY AND IMPLIED RISK

The work reviewed above illustrates that neuroscientists have done much to shed light on what distinguishes patient from impatient choices and individuals and have even begun to gain insight to the cognitive and neural processes that govern decisions about delayed rewards. However, there is an even more basic issue that has been largely ignored. Why are delayed rewards discounted at all? Why are small, immediate rewards ever tempting enough to eclipse larger, delayed rewards? Why would rational decision-makers not always wait for larger rewards, regardless of the associated delay?

Again, one likely explanation may come from a long history of theorizing in economics (Yaari, 1965; Benzion et al., 1989; Prelec and Loewenstein, 1991; Sozou, 1998; Dasgupta and Maskin, 2005), ecology (Kacelnik, 2003), and psychology (Mischel, 1966; Stevenson, 1986; Mazur, 1989; Rachlin et al., 1991; Mazur, 1995, 1997) which suggests that delay exerts its influence on choices via the perceived risk associated with waiting; a suggestion that has been referred to as the implicit risk hypothesis (Benzion et al., 1989). If a decision-maker believes that the probability of acquiring a promised reward is uncertain simply by virtue of being delayed, then that decision-maker is justified in reducing the subjective value of the reward. For example, a bird waiting for fruit to ripen might choose to eat some immediately if it believed the fruit's future availability was not guaranteed (i.e., it could be eaten by a competitor, it might rot, etc.). Furthermore, decision-makers might believe that the probability of receiving a promised reward generally decreases with time which would give rise to the monotonic decreases in subjective value that occur with increases in delay.

In one sense, the implicit risk hypothesis is attractive because it has the potential to entirely eliminate the phenomenon of delay discounting by translating time, the processing of which we are just beginning to grapple with (Mauk and Buonomano, 2004), into probability and uncertainty, concepts that are relatively well understood. In another sense, however, this hypothesis creates ambiguity when attempting to interpret previous delay discounting results (both behavioral and neural). For example, according to the implicit risk hypothesis, comparisons between immediate and delayed rewards are actually comparisons between high and low probability rewards. Thus, any results from such comparisons (e.g., contrasts in fMRI analyses) could reflect temporal processing or the processing of implicit probability or both. Similarly, one can reconceptualize the computation of subjective value as reflecting implicit probability instead of delay and the same can be done for dual-process accounts of choice.

Because of this potential ambiguity, it is instructive to briefly compare the temporal decision-making results reviewed above with work on choice under risk and uncertainty. Just as with the delay discounting work reviewed above, insular cortex has been implicated in risky decision-making. For example, insular cortex exhibits greater activity when decisions-makers chose a low probability reward than when decisions-makers chose a high probability reward (Paulus et al., 2003). Furthermore, insular activity predicts the likelihood of choosing low probability rewards. Reward probability also modulates activity in orbitofrontal and ventromedial frontal cortices (Critchley et al., 2001; Smith et al., 2002; Clark et al., 2008; Xue et al., 2009) as well as in the striatum (Hsu et al., 2005; Xue et al., 2009) and activity in the striatum is correlated with the

subjective value of risky rewards (Hsu et al., 2005; Knutson et al., 2005; Yacubian et al., 2006; Tobler et al., 2007; Yacubian et al., 2007). The overlap between these findings and those from investigations of ostensibly temporal decision-making suggest that there is at least a reasonable possibility that the implicit risk hypothesis is correct. To be clear, reducing temporal decision-making to risky choice in no way trivializes the work on temporal decision-making. Indeed, substantiating the neural equivalence of delay and reward probability would be a major step forward, helping to unify two, currently separate, processes and to validate long-standing theory.

Unfortunately, not all of the empirical evidence for the implicit risk hypothesis is as straightforward. For example, manipulations of probability and delay appear to elicit different patterns of choices (Ostaszewski et al., 1998; Holt et al., 2003; Green and Myerson, 2004; Chapman and Weber, 2006). For example, as reward magnitudes increase, probability appears to have more influence on behavior, whereas delay appears to have less influence (Green et al., 1999a). Temporal decisions appear to depend on whether the relevant rewards are immediately consumable (e.g., candy) or not (e.g., money) whereas discounting over probability does not (Estle et al., 2007). Furthermore, some authors (Green et al., 1999a) have noted that the lay concept of "impulsivity" seems to best describe an increased preference for low probability rewards (e.g., the temptation to play the lottery) but a decreased preference for delayed rewards (e.g., the temptation to take out a payday loan). Indeed, even with large samples, choice behavior in delay and probability discounting tasks is only weakly correlated within individual subjects (Myerson et al., 2003). Lastly, there are results showing that delay can have behavioral consequences even when probability is held constant. Work on what is referred to as the temporal resolution of uncertainty (Chew and Ho, 1994; Arai, 1997) has found people exhibit strong preferences between gambles in which reward delivery time is fixed and only differ in when the outcome of the gamble is revealed.

The partial dissociation of risky and temporal decision-making implies that the neural basis of temporal decision-making is significantly less clear than it might appear. Without appropriate comparisons, it remains ambiguous as to whether delay discounting results are being driven by delay, the risk implied by delay, or both. Recent work has begun to tackle this issue directly.

DREAD, HOPE, AND THE TEMPORAL RESOLUTION OF UNCERTAINTY

The first direct test of the implicit risk hypothesis to utilize fMRI (or any other physiological measure) was recently carried out (Weber and Huettel, 2008). Subjects in this study were asked to make two sorts of choices. First, subjects performed a classic risky choice task, choosing between rewards that varied in both magnitude and probability (e.g., a 50% chance of $13.50 or a 100% chance of $7). Second, subjects performed a traditional delay discounting task, choosing between rewards that varied in both magnitude and delay (e.g., $6.25 today or $9.25 in 1 month). According to the implicit risk hypothesis, these two conditions should be essentially identical because the stated delays are only influencing choices via the risk they imply. In contrast, this study revealed a variety of brain regions that were differentially engaged by the two tasks. Risky choice elicited greater involvement of posterior portions of parietal cortex, anterior cingulate, and anterior portions of the insula whereas the delay discounting task elicited greater involvement of DLPFC, posterior cingulate, and the caudate. Unfortunately, this comparison was complicated by the fact that subjects exhibited strikingly different patterns of choice in the risky and delayed choice tasks. Thus, it remains somewhat unclear whether the neural dissociation of risk and delay was driven by the task dimensions or other factors.

A recent fMRI study from our lab (Luhmann et al., 2008) has taken a slightly different approach to this same question. Rather than comparing risky and delay tasks, we instead choose to compare a risky task with a temporal resolution of uncertainty task that involved both risk and delay. Doing so allowed us to exert considerable control over the decision variables and to thus isolate behavioral and physiological results specifically tied to the temporal dimension. Subjects choose between pairs small rewards (10 and 20 cents) that were delivered with varying probabilities (39–100%). In the immediate condition, the uncertainty associated with subjects' choices was resolved immediately; subjects' learned whether they would or would not be receiving their chosen reward as soon as they made their choice. In the delay condition, the uncertainty associated with subjects' choices was resolved only after some variable delay. The delays were constructed such that lower probability rewards were resolved after a longer delay and higher probability rewards were resolved after a shorter delay. Specifically, the probabilities were such that the delay period embodied a constant hazard rate, a pattern that has been theorized to underlie normative delay discounting (Sozou, 1998; Dasgupta and Maskin, 2005). Comparing the two conditions, we found that both risk and delay exerted influence on subjects' choices. Subjects were significantly more likely to choose the larger, less probable reward when the outcomes were revealed immediately, despite the fact that the probabilities were identical. Neurally, the delay condition elicited greater activity in the posterior cingulate than did the immediate condition. Furthermore, we observed parametric effects in the parahippocampal gyri, the anterior cingulate and the portions of superior parietal cortex such that activity in these regions increased as the delay associated with chosen rewards increased. Lastly, we found that differences in individuals' attitudes toward the delay component of our task were mirrored by activity in a region of frontopolar cortex.

These two studies appear to contradict the implicit risk hypothesis and may begin to shed some light on how delay and risk exert dissociable influences on choice. We have noted that the specific brain regions implicated in temporal processing in our study have also been implicated in the process of prospection, the imagining of events in one's future (Okuda et al., 2003; Addis et al., 2007; Buckner and Carroll, 2007; Hassabis et al., 2007; Szpunar et al., 2007; Addis and Schacter, 2008). Thus, we suggested that one way in which temporal decisions might differ from similar, risky decisions, is that deliberation about temporal choices is likely to involve evaluating both the reward and its value, but also the experience of waiting itself. Our subjects' choice behavior implied that delaying the resolution of uncertainty decreased the subjective value of options (much like when reward delivery is delayed). This very well may be due to the fact that the delay interval itself evoked a negative subjective experience. Decision-makers able to foresee such experiences as they deliberated their choices would be in a much better position to make superior choices.

This possibility highlights the true complexity of temporal decision-making. Risky choice involves presenting a choice to the decision-maker, allowing a choice to be made, and resolving the outcome, all of which can and usually does happen rather quickly. Temporal decisions, on the other hand, are made at one point in time, but produce consequences that are subsequently stretched out over time. Decision-makers have subjective experiences as they attempt to make decision, while waiting, when uncertainty is resolved, and when receiving (or not receiving) the reward itself. To the extent that any or all of these experiences can be forecast before choices are made (Wilson and Gilbert, 2005), they can presumably exert some influence on decision-makers' behavior (Loewenstein, 1987; Loewenstein et al., 2001).

Indeed, we already know something about the physiology of anticipation itself. For example, in a particularly elegant fMRI study (Berns et al., 2006), human subjects were shown cues followed by a variable delay and then an electric shock. The identity of the cue signaled the duration of the delay interval, so subjects knew in advance how long they would have to wait. As subjects waited for the shock delivery, a network of brain regions exhibited a complex and theoretically interesting pattern of activity. Regions that respond to pain, including somatosensory cortex, insular cortex, and the anterior cingulate, exhibited activity that reflected both the anticipation of the impending shock as well as dread, the negative, subjective experience associated with the waiting itself. Furthermore, the neural patterns exhibited during this delay period were associated with preferences in a behavioral decision-making task performed separately. Those subjects that exhibited the strongest neural effects of dread were more likely to choose stronger, immediate shocks over weaker, but delayed shocks.

We also know that delay period activity is modulated by factors such as risk. In one study (Critchley et al., 2001), decision-makers made risky choices and had the outcomes of their choices withheld for 8 s. Across trials, the probability of winning was varied to investigate anticipatory processes. Several regions, including the anterior cingulate and orbital frontal cortex, and anterior insula exhibited delay period activity that reflected the amount of uncertainty subjects were facing.

These findings suggest that temporal decisions pose a formidable challenge for the savvy decision-maker. Despite the relative simple descriptions temporal outcomes can take (e.g., $100 in 12 months), they actual embody a complex sequence of events including emotional and cognitive events, each of which can potentially influence the subjective value of a choice. Delay periods can elicit negative emotional reactions (e.g., dread) and thus decrease the value of delayed outcomes, but waiting can also elicit positive emotional reaction (Loewenstein, 1987; Chew and Ho, 1994) and thus act to increase value. This lability, coupled with the finding that people are not necessarily adept at predicting their future emotional reactions (Wilson and Gilbert, 2005), begins to make temporal choice look even more difficult. Furthermore, not only do these factors complicate temporal decision-making itself, they also complicate our attempts to study it; attempts to fully control and dissociate each of the relevant influences are unlikely to be feasible. Nonetheless, studies that acknowledge and take these factors into account certainly have the potential to help overcome some of the ambiguities noted above and to paint a much richer picture of temporal decision-making.

CONCLUSIONS

The work reviewed above suggests that we are just beginning to gain insight into the nature temporal decision-making. Much of this work has sought to explore the phenomenon of delay discounting; the finding that waiting for rewards decreases their attractiveness. As a first step, this work has begun to characterize the difference between patient and impatient choices within a single individual as well as between patient individuals and impatient individuals. While critically informative, we ultimately need to understand the processes that operate to produce temporal decisions. Fortunately, the study of decision-making has a wealth of theoretical tools available from economics. More recent work has attempted to leverage economic theories to better understand the neural patterns observed during decision-making. The results have been provocative, but there are many questions left unanswered.

We have pointed to several places where there is currently ambiguity in the treatment of temporal decision-making: correlates of patience vs. subjective value signals, delay vs. risk, etc. Furthermore, we have tried to point to a small number of relatively unexplored dimensions that are likely to be relevant for temporal decision-making: affective forecasting, anticipation, etc. This is certainly not an exhaustive list, as others (Berns et al., 2007; Wittmann and Paulus, 2008) have noted additional relevant factors. At this point, it appears that, despite complicating our experimental designs, investigating the interactions between these factors would be most valuable in illuminating the subtleties of temporal decision-making. Given the clinical and public policy implications of temporal decision-making as well as the sheer scientific potential in psychology, neuroscience, and economics, the benefits of such an approach seem to outweigh the costs.

ACKNOWLEDGMENTS

We are grateful to Heather Moore and Marla Krukowski for comments and feedback on drafts of this manuscript.

REFERENCES

Addis, D. R., and Schacter, D. L. (2008). Constructive episodic simulation: temporal distance and detail of past and future events modulate hippocampal engagement. *Hippocampus* 18, 227–237.

Addis, D. R., Wong, A. T., and Schacter, D. L. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia* 45, 1363–1377.

Ainslie, G. (1975). Specious reward: a behavioral theory of impulsiveness and impulse control. *Psychol. Bull.* 82, 463–496.

Ainslie, G. (1992). Picoeconomics: The Strategic Interaction of Successive Motivational States within the Person. New York, Cambridge University Press.

Ainslie, G. (2001). Breakdown of Will. Cambridge, Cambridge University Press.

Ainslie, G., and Herrnstein, R. J. (1981). Preference reversal and delayed reinforcement. *Anim. Learn. Behav.* 9, 476–482.

Arai, D. (1997). Temporal resolution of uncertainty in risky choices. *Acta Psychol.* 96, 15–26.

Barkley, R. A., Edwards, G., Laneri, M., Fletcher, K., and Metevia, L. (2001). Executive functioning, temporal discounting, and sense of time in adolescents with attention deficit hyperactivity disorder (ADHD) and oppositional defiant disorder (ODD). *J. Abnorm. Child. Psychol.* 29, 541–556.

Bartzokis, G., Goldstein, I. B., Hance, D. B., Beckson, M., Shapiro, D., Lu, P. H., Edwards, N., Mintz, J., and Bridge, P. (1999). The incidence of T2-weighted

MR imaging signal abnormalities in the brain of cocaine-dependent patients is age-related and region-specific. *Am. J. Neuroradiol.* 20, 1628–1635.

Benzion, U., Rapoport, A., and Yagil, J. (1989). Discount rates inferred from decisions: an experimental study. *Manage. Sci.* 35, 270–284.

Berns, G. S., Chappelow, J., Cekic, M., Zink, C. F., Pagnoni, G., and Martin-Skurski, M. E. (2006). Neurobiological substrates of dread. *Science* 312, 754–758.

Berns, G. S., Laibson, D., and Loewenstein, G. (2007). Intertemporal choice–toward an integrative framework. *Trends Cogn. Sci.* 11, 482–488.

Bickel, W. K., Odum, A. L., and Madden, G. J. (1999). Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology* 146, 447–454.

Boettiger, C. A., Mitchell, J. M., Tavares, V. C., Robertson, M., Joslyn, G., D'Esposito, M., and Fields, H. L. (2007). Immediate reward bias in humans: fronto-parietal networks and a role for the catechol-*O*-methyltransferase 158Val/Val genotype. *J. Neurosci.* 27, 14383–14391.

Bonson, K. R., Grant, S. J., Contoreggi, C. S., Links, J. M., Metcalfe, J., Weyl, H. L., Kurian, V., Ernst, M., and London, E. D. (2002). Neural systems and cue-induced cocaine craving. *Neuropsychopharmacology* 26, 376–386.

Brody, A. L., Mandelkern, M. A., London, E. D., Childress, A. R., Lee, G. S., Bota, R. G., Ho, M. L., Saxena, S., Baxter, L. R. Jr., Madsen, D., and Jarvik, M. E. (2002). Brain metabolic changes during cigarette craving. *Arch. Gen. Psychiatry* 59, 1162–1172.

Buckner, R. L., and Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57.

Burks, S. V., Carpenter, J. P., Goette, L., and Rustichini, A. (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7745–7750.

Cardinal, R. N. (2006). Neural systems implicated in delayed and probabilistic reinforcement. *Neural Netw.* 19, 1277–1301.

Chapman, G. B., and Weber, B. J. (2006). Decision biases in intertemporal choice and choice under uncertainty: testing a common account. *Mem. Cogn.* 34, 589–602.

Chew, S. H., and Ho, J. L. (1994). Hope: an empirical study of attitude toward the timing of uncertainty resolution. *J. Risk Uncertain.* 8, 267–288.

Clark, L., Bechara, A., Damasio, H., Aitken, M. R. F., Sahakian, B. J., and Robbins, T. W. (2008). Differential effects of insular and ventromedial prefrontal cortex lesions on risky decision-making. *Brain* 131, 1311–1322.

Coffey, S. F., Gudleski, G. D., Saladin, M. E., and Brady, K. T. (2003). Impulsivity and rapid discounting of delayed hypothetical rewards in cocaine-dependent individuals. *Exp. Clin. Psychopharmacol.* 11, 18–25.

Critchley, H. D., Mathias, C. J., and Dolan, R. J. (2001). Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron* 29, 537–545.

Dasgupta, P. (2008). Discounting climate change. *J. Risk Uncertain.* 37, 141–169.

Dasgupta, P., and Maskin, E. (2005). Uncertainty and hyperbolic discounting. *Am. Econ. Rev.* 95, 1290–1299.

Edwards, W. (1954). The theory of decision making. *Psychol. Bull.* 51, 380–417.

Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E., Goldman, D., and Weinberger, D. R. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proc. Natl. Acad. Sci. U.S.A.* 98, 6917–6922.

Ernst, M., Kimes, A. S., London, E. D., Matochik, J. A., Eldreth, D., Tata, S., Contoreggi, C., Leff, M., and Bolla, K. (2003). Neural substrates of decision making in adults with attention deficit hyperactivity disorder. *Am. J. Psychiatry* 160, 1061–1070.

Estle, S. J., Green, L., Myerson, J., and Holt, D. D. (2007). Discounting of monetary and directly consumable rewards. *Psychol. Sci.* 18, 58–63.

Forbes, E. E., Brown, S. M., Kimak, M., Ferrell, R. E., Manuck, S. B., and Hariri, A. R. (2009). Genetic variation in components of dopamine neurotransmission impacts ventral striatal reactivity associated with impulsivity. *Mol. Psychiatry* 14, 60–70.

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., and Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc. Natl. Acad. Sci. U.S.A.* 104, 16311–16316.

Franklin, T. R., Acton, P. D., Maldjian, J. A., Gray, J. D., Croft, J. R., Dackis, C. A., O'Brien, C. P., and Childress, A. R. (2002). Decreased gray matter concentration in the insular, orbitofrontal, cingulate, and temporal cortices of cocaine patients. *Biol. Psychiatry* 51, 134–142.

Fuster, J. M. (2002). Frontal lobe and cognitive development. *J. Neurocytol.* 31, 373–385.

Garavan, H., Pankiewicz, J., Bloom, A., Cho, J. K., Sperry, L., Ross, T. J., Salmeron, B. J., Risinger, R., Kelley, D., and Stein, E. A. (2000). Cue-induced cocaine craving: neuroanatomical specificity for drug users and drug stimuli. *Am. J. Psychiatry* 157, 1789–1798.

Gately, D. (1980). Individual discount rates and the purchase and utilization of energy-using durables: comment. *Bell J. Econ.* 11, 373–374.

Glimcher, P. (2003). Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics (Bradford Books). Cambridge, The MIT Press.

Glimcher, P., Camerer, C., Poldrack, R., and Fehr, E. (2008). Neuroeconomics: Decision Making and the Brain. London, Academic Press.

Green, L., Fristoe, N., and Myerson, J. (1994a). Temporal discounting and preference reversals in choice between delayed outcomes. *Psychon. Bull. Rev.* 1, 383–389.

Green, L., Fry, A. F., and Myerson, J. (1994b). Discounting of delayed rewards: a life-span comparison. *Psychol. Sci.* 5, 33–36.

Green, L., and Myerson, J. (2004). A discounting framework for choice with delayed and probabilistic rewards. *Psychol. Bull.* 130, 769–792.

Green, L., Myerson, J., and Ostaszewski, P. (1999a). Amount of reward has opposite effects on the discounting of delayed and probabilistic outcomes. *J. Exp. Psychol. Learn. Mem. Cogn.* 25, 418–427.

Green, L., Myerson, J., and Ostaszewski, P. (1999b). Discounting of delayed rewards across the life span: age differences in individual discounting functions. *Behav. Process.* 46, 89–96.

Gregorios-Pippas, L., Tobler, P. N., and Schultz, W. (2009). Short-term temporal discounting of reward value in human ventral striatum. *J. Neurophysiol.* 101, 1507–1523.

Harding, J. P. (2000). Mortgage valuation with optimal intertemporal refinancing strategies. *J. Hous. Econ.* 9, 233–266.

Hardisty, D. J., and Weber, E. U. (2009). Discounting future green: money versus the environment. *J. Exp. Psychol. Gen.* 138, 329–340.

Hariri, A. R., Brown, S. M., Williamson, D. E., Flory, J. D., de Wit, H., and Manuck, S. B. (2006). Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. *J. Neurosci.* 26, 13213.

Hassabis, D., Kumaran, D., Vann, S. D., and Maguire, E. A. (2007). Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1726.

Hausman, J. A. (1979). Individual discount rates and the purchase and utilization of energy-using durables. *Bell J. Econ.* 10, 33–54.

Holt, D. D., Green, L., and Myerson, J. (2003). Is discounting impulsive? Evidence from temporal and probability discounting in gambling and non-gambling college students. *Behav. Process.* 64, 355–367.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., and Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science* 310, 1680–1683.

Hwang, J., Kim, S., and Lee, D. (2009). Temporal discounting and intertemporal choice in rhesus monkeys. *Front. Behav. Neurosci.* 3:9. doi:10.3389/neuro.08.009.2009.

Kable, J. W., and Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.

Kacelnik, A. (2003). The evolution of patience. In: Time and Decision: Economic and Psychological Perspectives on Intertemporal Choice, G. Loewenstein, D. Read and R. F. Baumeister, eds (New York, Russell Sage Foundation), pp. 115–138.

Kheramin, S., Body, S., Ho, M., Velazquez-Martinez, D. N., Bradshaw, C. M., Szabadi, E., Deakin, J. F. W., and Anderson, I. M. (2004). Effects of orbital prefrontal cortex dopamine depletion on inter-temporal choice: a quantitative analysis. *Psychopharmacology* 175, 206–214.

Kilts, C. D., Schweitzer, J. B., Quinn, C. K., Gross, R. E., Faber, T. L., Muhammad, F., Ely, T. D., Hoffman, J. M., and Drexler, K. P. G. (2001). Neural activity related to drug craving in cocaine addiction. *Arch. Gen. Psychiatry* 58, 334–341.

Kim, S., Hwang, J., and Lee, D. (2008). Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron* 59, 161–172.

Kirby, K. N. (1997). Bidding on the future: evidence against normative discounting of delayed rewards. *J. Exp. Psychol. Gen.* 126, 54–70.

Kirby, K. N., and Herrnstein, R. J. (1995). Preference reversals due to myopic discounting of delayed reward. *Psychol. Sci.* 6, 83–89.

Kirby, K. N., and Petry, N. M. (2004). Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction* 99, 461–471.

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., and Glover, G. (2005).

Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Q. J. Econ.* 112, 443–477.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *Econ. J.* 97, 666–684.

Loewenstein, G. (1996). Out of control: visceral influences on behavior. *Organ. Behav. Hum. Decis. Process.* 65, 272–292.

Loewenstein, G., and Thaler, R. H. (1989). Anomalies: intertemporal choice. *J. Econ. Perspect.* 3, 181–193.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., and Welch, N. (2001). Risk as feelings. *Psychol. Bull.* 127, 267–286.

Logue, A. W., Peña-Correal, T. E., Rodriguez, M. L., and Kabela, E. (1986). Self-control in adult humans: variation in positive reinforcer amount and delay. *J. Exp. Anal. Behav.* 46, 159–173.

Luhmann, C. C., Chun, M. M., Yi, D. J., Lee, D., and Wang, X. J. (2008). Neural dissociation of delay and uncertainty in intertemporal choice. *J. Neurosci.* 28, 14459–14466.

Mauk, M., and Buonomano, D. V. (2004). The neural basis of temporal processing. *Annu. Rev. Neurosci.* 27, 307–340.

Mazur, J. E. (1984). Tests of an equivalence rule for fixed and variable reinforcer delays. *J. Exp. Psychol. Anim. Behav. Process.* 10, 426–436.

Mazur, J. E. (1989). Theories of probabilistic reinforcement. *J. Exp. Anal. Behav.* 51, 87–99.

Mazur, J. E. (1995). Conditioned reinforcement and choice with delayed and uncertain primary reinforcers. *J. Exp. Anal. Behav.* 63, 139–150.

Mazur, J. E. (1997). Choice, delay, probability, and conditioned reinforcement. *Learn. Behav.* 25, 131–147.

McClure, S. M., Ericson, K. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2007). Time discounting for primary rewards. *J. Neurosci.* 27, 5796–5804.

McClure, S. M., Laibson, D. I., Loewenstein, G., and Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science* 306, 503–507.

Metcalfe, J., and Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. *Psychol. Rev.* 106, 3–19.

Mischel, W. (1966). Theory and research on the antecedents of self-imposed delay of reward. *Prog. Exp. Pers. Res.* 3, 85–132.

Mischel, W., Shoda, Y., and Rodriguez, M. I. (1989). Delay of gratification in children. *Science* 244, 933–938.

Montague, P. R., and Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron* 36, 265–284.

Moustafa, A. A., Cohen, M. X., Sherman, S. J., and Frank, M. J. (2008). A role for dopamine in temporal decision making and reward maximization in Parkinsonism. *J. Neurosci.* 28, 12294–12304.

Myerson, J., and Green, L. (1995). Discounting of delayed rewards: models of individual choice. *J. Exp. Anal. Behav.* 64, 263–276.

Myerson, J., Green, L., Scott Hanson, J., Holt, D. D., and Estle, S. J. (2003). Discounting delayed and probabilistic rewards: processes and traits. *J. Econ. Psychol.* 24, 619–635.

Myrick, H., Anton, R. F., Li, X., Henderson, S., Drobes, D., Voronin, K., and George, M. S. (2004). Differential brain activity in alcoholics and social drinkers to alcohol cues: relationship to craving. *Neuropsychopharmacology* 29, 393–402.

Naqvi, N. H., Rudrauf, D., Damasio, H., and Bechara, A. (2007). Damage to the insula disrupts addiction to cigarette smoking. *Science* 315, 531–534.

Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Tanji, K., Suzuki, K., Kawashima, R., Fukuda, H., Itoh, M., and Yamadori, A. (2003). Thinking of the future and past: the roles of the frontal pole and the medial temporal lobes. *Neuroimage* 19, 1369–1380.

Ostaszewski, P., Green, L., and Myerson, J. (1998). Effects of inflation on the subjective value of delayed and probabilistic rewards. *Psychon. Bull. Rev.* 5, 324–333.

Paulus, M. P., Rogalsky, C., Simmons, A., Feinstein, J. S., and Stein, M. B. (2003). Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *Neuroimage* 19, 1439–1448.

Paulus, M. P., Tapert, S. F., and Schuckit, M. A. (2005). Neural activation patterns of methamphetamine-dependent subjects during decision making predict relapse. *Arch. Gen. Psychiatry* 62, 761–768.

Petry, N. M. (2001). Delay discounting of money and alcohol in actively using alcoholics, currently abstinent alcoholics, and controls. *Psychopharmacology* 154, 243–250.

Phillips, P. E. M., Walton, M. E., and Jhou, T. C. (2007). Calculating utility: preclinical evidence for cost–benefit analysis by mesolimbic dopamine. *Psychopharmacology* 191, 483–495.

Ploghaus, A., Tracey, I., Gati, J. S., Clare, S., Menon, R. S., Matthews, P. M., and Rawlins, J. N. P. (1999). Dissociating pain from its anticipation in the human brain. *Science* 284, 1979–1981.

Prelec, D., and Loewenstein, G. (1991). Decision making over time and under uncertainty: a common approach. *Manage. Sci.* 37, 770–786.

Rachlin, H. (1995). Self-control: Beyond commitment. *Behav. Brain Sci.* 18, 109–159.

Rachlin, H., Raineri, A., and Cross, D. (1991). Subjective probability and delay. *J. Exp. Anal. Behav.* 55, 233–244.

Rahman, S., Sahakian, B. J., Cardinal, R. N., Rogers, R. D., and Robbins, T. W. (2001). Decision making and neuropsychiatry. *Trends Cogn. Sci.* 5, 271–277.

Reynolds, B., Richards, J. B., Horn, K., and Karraker, K. (2004). Delay discounting and probability discounting as related to cigarette smoking status in adults. *Behav. Process.* 65, 35–42.

Rossow, I. (2008). Alcohol consumption and discounting. *Addict. Res. Theory* 16, 572–584.

Rubia, K., Smith, A. B., Halari, R., Matsukura, F., Mohammad, M., Taylor, E., and Brammer, M. J. (2009). Disorder-specific dissociation of orbitofrontal dysfunction in boys with pure conduct disorder during reward and ventrolateral prefrontal dysfunction in boys with pure ADHD during sustained attention. *Am. J. Psychiatry* 166, 83–94.

Samuelson, P. A. (1937). A note on measurement of utility. *Rev. Econ. Stud.* 4, 155–161.

Schelling, T. C. (1984). Self-command in practice, in policy, and in a theory of rational choice. *Am. Econ. Rev.* 74, 1–11.

Scheres, A., Dijkstra, M., Ainslie, E., Balkan, J., Reynolds, B., Sonuga-Barke, E., and Castellanos, F. X. (2006). Temporal and probabilistic discounting of rewards in children and adolescents: effects of age and ADHD symptoms. *Neuropsychologia* 44, 2092–2103.

Schneider, F., Habel, U., Wagner, M., Franke, P., Salloum, J. B., Shah, N. J., Toni, I., Sulzbach, C., Honig, K., and Maier, W. (2001). Subcortical correlates of craving in recently abstinent alcoholic patients. *Am. J. Psychiatry* 158, 1075–1083.

Shamosh, N. A., and Gray, J. R. (2008). Delay discounting and intelligence: a meta-analysis. *Intelligence* 4, 289–305.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., and Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162.

Smith, K., Dickhaut, J., McCabe, K., and Pardo, J. V. (2002). Neuronal substrates for choice under ambiguity, risk, gains, and losses. *Manage. Sci.* 48, 711–718.

Soman, D., Ainslie, G., Frederick, S., Li, X., Lynch, J., Moreau, P., Mitchell, A., Read, D., Sawyer, A., and Trope, Y. (2005). The psychology of intertemporal discounting: why are distant events valued differently from proximal ones? *Mark. Lett.* 16, 347–360.

Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L., and Toga, A. W. (1999). In vivo evidence for postadolescent brain maturation in frontal and striatal regions. *Nat. Neurosci.* 2, 859–860.

Sozou, P. D. (1998). On hyperbolic discounting and uncertain hazard rates. *Proc. Biol Sci.* 265, 2015–2020.

Stevenson, M. K. (1986). A discounting model for decisions with delayed positive or negative outcomes. *J. Exp. Psychol. Gen.* 115, 131–154.

Szpunar, K. K., Watson, J. M., and McDermott, K. B. (2007). Neural substrates of envisioning the future. *Proc. Natl. Acad. Sci. U.S.A.* 104, 642.

Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., and Yamawaki, S. (2004). Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* 7, 887–893.

Thaler, R. H., and Sheffrin, H. M. (1981). A theory of self-control. *J. Polit. Econ.* 89, 392–406.

Tobler, P. N., O'Doherty, J. P., Dolan, R. J., and Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* 97, 1621.

Vuchinich, R. E., and Simpson, C. A. (1998). Hyperbolic temporal discounting in social drinkers and problem drinkers. *Exp. Clin. Psychopharmacol.* 6, 292–305.

Wang, Z., Faith, M., Patterson, F., Tang, K., Kerrin, K., Wileyto, E. P., Detre, J. A., and Lerman, C. (2007). Neural substrates of abstinence-induced cigarette cravings in chronic smokers. *J. Neurosci.* 27, 14035–14040.

Weber, B., and Huettel, S. (2008). The neural substrates of probabilistic and intertemporal decision making. *Brain Res.* 1234, 104–115.

Wilson, T. D., and Gilbert, D. T. (2005). Affective forecasting: Knowing what to want. *Curr. Dir. Psychol. Sci.* 14, 131–134.

Winstanley, C. A., Theobald, D. E. H., Cardinal, R. N., and Robbins, T. W. (2004). Contrasting roles of basolateral amygdala and orbitofrontal cortex in impulsive choice. *J. Neurosci.* 24, 4718–4722.

Wittmann, M., Leland, D. S., and Paulus, M. P. (2007). Time and decision making: differential contribution of the posterior insular cortex and the striatum during a delay discounting task. *Exp. Brain Res.* 179, 643–653.

Wittmann, M., and Paulus, M. P. (2008). Decision making, impulsivity and time perception. *Trends Cogn. Sci.* 12, 7–12.

Xue, G., Lu, Z., Levin, I. P., Weller, J. A., Li, X., and Bechara, A. (2009). Functional dissociations of risk and reward processing in the dedial prefrontal cortex. *Cereb. Cortex* 19, 1019–1027.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *Rev. Econ. Stud.* 137–150.

Yacubian, J., Glascher, J., Schroeder, K., Sommer, T., Braus, D. F., and Buchel, C. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *J. Neurosci.* 26, 9530–9537.

Yacubian, J., Sommer, T., Schroeder, K., Gläscher, J., Kalisch, R., Leuenberger, B., Braus, D. F., and Büchel, C. (2007). Gene–gene interaction associated with neural reward sensitivity. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8125–8130.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 13 June 2009; paper pending published: 13 July 2009; accepted: 06 October 2009; published online: 23 October 2009.

Citation: Luhmann CC (2009) Temporal decision-making: insights from cognitive neuroscience. Front. Behav. Neurosci. ***3****:39. doi: 10.3389/neuro.08.039.2009*

frontiers in
BEHAVIORAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 27 January 2010
doi: 10.3389/neuro.08.001.2010

A physiologically-inspired model of numerical classification based on graded stimulus coding

John Pearson[1,2,3]*, J.D. Roitman[4], E.M. Brannon[3,5], M.L. Platt[1,2,3,6] and Sridhar Raghavachari[1]

[1] *Department of Neurobiology, Duke University School of Medicine, Durham, NC, USA*
[2] *Center for Neuroeconomic Studies, Duke University, Durham, NC, USA*
[3] *Center for Cognitive Neuroscience, Duke University, Durham, NC, USA*
[4] *Department of Psychology, University of Illinois, Chicago, IL, USA*
[5] *Department of Psychology and Neuroscience, Duke University, Durham, NC, USA*
[6] *Department of Evolutionary Anthropology, Duke University, Durham, NC, USA*

Edited by:
Daeyeol Lee, Yale University School of Medicine, USA

Reviewed by:
Bruno Averbeck, University College London, UK
Jacqueline Gottlieb, Yale University School of Medicine, USA
Daeyeol Lee, Yale University School of Medicine, USA

***Correspondence:**
John Pearson, Duke University Medical Center, Box 3209, Durham, NC 27710, USA.
e-mail: pearson@neuro.duke.edu

In most natural decision contexts, the process of selecting among competing actions takes place in the presence of informative, but potentially ambiguous, stimuli. Decisions about magnitudes – quantities like time, length, and brightness that are linearly ordered – constitute an important subclass of such decisions. It has long been known that perceptual judgments about such quantities obey Weber's Law, wherein the just-noticeable difference in a magnitude is proportional to the magnitude itself. Current physiologically inspired models of numerical classification assume discriminations are made via a labeled line code of neurons selectively tuned for numerosity, a pattern observed in the firing rates of neurons in the ventral intraparietal area (VIP) of the macaque. By contrast, neurons in the contiguous lateral intraparietal area (LIP) signal numerosity in a graded fashion, suggesting the possibility that numerical classification could be achieved in the absence of neurons tuned for number. Here, we consider the performance of a decision model based on this analog coding scheme in a paradigmatic discrimination task – numerosity bisection. We demonstrate that a basic two-neuron classifier model, derived from experimentally measured monotonic responses of LIP neurons, is sufficient to reproduce the numerosity bisection behavior of monkeys, and that the threshold of the classifier can be set by reward maximization via a simple learning rule. In addition, our model predicts deviations from Weber Law scaling of choice behavior at high numerosity. Together, these results suggest both a generic neuronal framework for magnitude-based decisions and a role for reward contingency in the classification of such stimuli.

Keywords: LIP, number, signal detection, decision making, reinforcement learning, neuroeconomics, discrimination

INTRODUCTION

For one-dimensional quantities like number, time, length, and brightness that possess a natural linear order (Moyer and Landauer, 1967; Stevens, 1986), discrimination behavior is characterized by the distance and magnitude effects: discrimination improves as the difference in stimuli along the perceptual dimension increases, but suffers as the absolute magnitudes grow (Moyer and Landauer, 1967; Brannon and Terrace, 1998; Nieder and Miller, 2003). More generally, such quantities obey Weber's Law: the just-noticeable difference in a magnitude is proportional to the magnitude itself.

In the last several years, single unit recordings and fMRI studies have implicated neurons in the intraparietal sulcus in coding one of these quantities – number (Nieder et al., 2002; Nieder and Miller, 2003, 2004a,b; Nieder, 2005; Nieder and Merten, 2007; Roitman et al., 2007). Moreover, neurons in the ventral intraparietal area (VIP) show preferential firing to specific numerosities, with tuning curve widths scaling as the logarithm of the preferred number (Nieder et al., 2002; Nieder and Miller, 2003; Nieder and Merten, 2007). fMRI repetition suppression studies have largely confirmed these observations (Piazza et al., 2004). As a result, most recent theoretical work on numerical cognition has focused on models that represent number via pools of neurons preferentially activated by specific numerosities (Dehaene and Changeux, 1993; Zorzi and Butterworth, 1999; Grossberg and Repin, 2003; Verguts and Fias, 2004; Nieder, 2005; Verguts et al., 2005). Such models naturally account for performance in discrimination tasks of both dot arrays and Arabic numerals, and suggest that the logarithmic widening of VIP neuron tuning curves gives rise to Weber's Law for numerical discrimination.

By contrast, neurons in the lateral intraparietal area (LIP) were recently shown to represent the number of dots in a visual array in a graded, monotonic fashion (Roitman et al., 2007). Notably, there were separate populations of LIP neurons that increased firing with increasing numerosity and decreased firing with increasing numerosity. Similar reciprocal neuronal coding of somatosensory stimuli has previously been observed in somatosensory cortex for vibration frequency (Miller et al., 2003; Machens et al., 2005). These analog codes provide a physiological basis for alternative models of magnitude discrimination, including number, without the need to invoke explicit neuronal representations of specific values (Gibbon, 1977, 1981; Gibbon and Church, 1981; Gibbon and Fairhurst, 1994).

Such analog models, proposed initially for interval timing, typically rely on one of two underlying neural codes for magnitudes. In linear models, magnitudes are represented by linearly increasing firing rates, with noise that grows in proportion to the firing rate itself. Comparisons between magnitudes are performed by taking ratios of these linear representations, with the result that discriminations between magnitudes become easier as distances between them grow (the distance effect) and harder to distinguish (for fixed difference between them) as their absolute values increase (the magnitude effect). Moreover, the assumption of a noisy internal representation with standard deviation proportional to the mean (constant coefficient of variation), dubbed the "scalar property," gives rise to a discriminability parameter proportional to the difference in magnitudes divided by their absolute size, reproducing the Weber-Fechner discrimination law (Gibbon, 1977, 1981; Gibbon and Church, 1981; Brannon et al., 2001).

In the second class of models, magnitudes are represented by firing rates that scale with the logarithm of the underlying quantity (Gibbon, 1977, 1981; Gibbon and Church, 1981) (not to be confused with the logarithmically widening tuning curves of numerosity-selective neurons). Comparisons in these models are performed by subtraction, a linear operation equivalent to taking the ratio of the original magnitudes. In addition, constant variance in the logarithmically compressed internal representation corresponds to a log-normal variance in the original quantity, with a standard deviation proportional to that quantity, thus reproducing the scalar property from the linear models. Thus, in contrast to population code models, which represent numerosity via pools of neurons selective for each number ("cardinal codes"), these models represent numerosity in graded fashion in a single neuronal firing rate ("ordinal codes").

Given either of these noisy internal analog representations of magnitude, signal detection theory provides a principled framework for classification (Green and Swets, 1989; Gibbon, 1981). In signal detection theory, not only do the statistics of the underlying representation enter into the decision making process, but the costs and benefits of stimulus identification do so as well. Thus, if a "yes" response to a question about an ambiguous stimulus is rewarded twice as much as a "no," the optimal strategy (from a reward maximization standpoint) is to respond "yes" in all cases where the stimulus is equally likely to correspond to either, and even in many cases where it is more likely to correspond to "no." Typically, this prediction is tested in a bisection paradigm, in which subjects are asked to provide a binary classification of a quasi-continuous range of stimuli (Church and Deluty, 1977; Meck and Church, 1983; Platt and Davis, 1983; Meck et al., 1985; Roberts, 2005; Jordan and Brannon, 2006). Stimuli at either extreme of the range (the "anchors") are each paired with a unique rewarded response, but intermediate stimuli are classified freely. By measuring the resulting choice function, the underlying decision process can be characterized.

Measurement of psychometric curves in the bisection paradigm results in two primary empirical findings (Gibbon, 1981; Gibbon and Church, 1981; Gibbon and Fairhurst, 1994): First, points of subjective equality (PSE) for stimulus classification – stimulus magnitudes for which subjects are equally likely to produce either response – are located at the geometric mean of the two anchor values. Second, when plotted as a function of stimulus magnitude divided by PSE (a PSE-relative scale), psychometric curves for distinct pairs of anchors superimpose. The latter may be seen as a consequence of the scalar property (for either linear or logarithmic encoding schemes), since Weber's Law predicts that ratio-based discrimination should be invariant to magnitude rescaling.

Here, we demonstrate that a discrimination model based on observed numerosity tuning curves for LIP neurons, in the absence of explicit representation of numerical value, is sufficient to reproduce the choice performance of macaques in a separate behavioral study of numerical bisection. We further predict departures from Weber's Law at large numerosities that differentiate between linear and logarithmic encoding of numerosity. Furthermore, we show that simple reinforcement learning correctly sets the indifference point for numerical bisection in our model, without explicit knowledge of either reward history or underlying tuning functions, with important implications for classification performance in the case of unequally rewarded anchors. Together, our findings demonstrate that monotonic analog codes can support discrimination of abstract quantities like number, in addition to simple sensory stimuli like vibrotactile frequency (Machens et al., 2005).

MATERIALS AND METHODS

NEURONAL DATA: IMPLICIT DISCRIMINATION

We base our model on neurophysiological data published previously (Roitman et al., 2007). There, the authors characterized the responses of LIP neurons to arrays of dots in an implicit numerosity discrimination task (Roitman et al., 2007). Single units ($n = 53$) were isolated in area LIP, and their spatial receptive fields identified, using a standard delayed-saccade paradigm. During the task, the animal was required to hold fixation on a central cue. They were then presented with a saccade target in the hemifield opposite the receptive field of the neuron. After a variable delay, a dot array of numerosity 2, 4, 8, 16, or 32 (controlled for density, element size, and total pixels) was presented in the receptive field for 400 ms. After another variable delay, monkeys shifted gaze to the target opposite the receptive field. In each block, one numerosity was selected as standard and presented on 50% of trials. On the other half of trials, cue numerosities were randomly chosen from among the four remaining values. The animal received 100 ms of juice for successful saccades following standard cues, 150 ms for successful saccades following deviants. (Both saccades were to the same location.) Since every trial resulting in a saccade was rewarded, animals did not need to attend to numerosity to maximize reward, though decreased reaction times for trials with deviant cues argue that they did so.

BEHAVIORAL DATA: NUMERICAL BISECTION

To verify that our model produces psychometric curves of the form measured in behavioral studies of bisection, we compare its output to the previously-published work of Jordan and Brannon in a separate pair of monkeys (Jordan and Brannon, 2006) (note that the monkeys in the Roitman et al., 2007, study were numerically naïve and did not perform a bisection task). In Jordan and Brannon's experiment, adult rhesus monkeys were first trained to recognize the number of elements in a dot display in a delayed match-to-sample (DMS) paradigm using a touch-sensitive monitor. Upon trial initiation, a stimulus consisting of a yellow rectangle containing a

variable number of dot elements was presented, followed afterward by two choice stimuli (match and distractor) and the animal's response. Correct choices were rewarded by juice delivery, and several confounding dimensions of the dot arrays (cumulative area, dot size, density) were controlled, ensuring that only numerosity remained a reliable guide to behavior.

Once animals were able to recognize individual numerosities, two stimuli (block type 1: 2 and 8, block type 2: 3 and 12) were selected as anchors and presented (with equal probability) as the cue stimulus. As in the DMS paradigm, a match and distractor were subsequently presented, always equal in numerosity to the anchor values. Correct trials of this type were again rewarded. However, on 30% of trials, an intermediate numerosity appeared as the cue (block type 1: 3–7, block type 2: 4–11), followed by dot array choices corresponding to the two anchors. This required the animal to classify a non-matching stimulus with one or the other of the two anchors. Though these trials were never rewarded, the animals nevertheless displayed systematic responses to the intermediate numerosity cues, transitioning from responses corresponding to the small anchor value to responses favoring the large anchor value as cue numerosity increased (**Figure 1B**).

MODELING

The most common paradigm used to study magnitude estimation and number judgment in rats, pigeons, and non-human primates is numerical bisection, in which a subject is required to classify the numerosity of a cue as one or the other of a pair of "standard" values (Church and Deluty, 1977; Meck and Church, 1983; Platt and Davis, 1983; Meck et al., 1985; Roberts, 2005; Jordan and Brannon, 2006). Of tasks designed to quantify numerical capability, it remains the most direct, and gives the clearest demonstration of Weber's Law behavior. We asked whether the observed response functions of numerosity-sensitive neurons in area LIP (Roitman et al., 2007) might function as a code capable of reproducing choice behavior in a similar bisection task.

We modeled behavior in an oculomotor version of the numerical bisection task. The oculomotor paradigm has been widely used to study response properties of LIP neurons (Snyder et al., 1997; Andersen and Buneo, 2002) and to probe the neural correlates of decisions in a wide variety of cognitive tasks (Platt and Glimcher, 1999; Gold and Shadlen, 2000; Shadlen and Newsome, 2001; Roitman and Shadlen, 2002; Leon and Shadlen, 2003). Moreover, framing the experiment in this way allows us to make direct use of single-unit recordings from LIP in our model, as well as to make testable predictions about neuronal activity as task conditions are altered. In the task (**Figure 1A**), fixation on the central cue is followed by the presentation of two targets (red and green) in the hemifield opposite the receptive field of a recorded neuron. This is followed by a variable delay, after which a dot array cue is presented in the neuron's receptive field. This is followed by another variable delay, after which the animal is free to shift gaze to either the green ("small" response) or red ("large" response) target. **Figure 1B** presents behavioral data from a similar bisection paradigm (Jordan and Brannon, 2006), along with fits produced by our models (see below).

In order to extrapolate differences in model predictions to high numerosity, we fit neuronal responses of LIP during the 400 ms of stimulus presentation with both linear and logarithmic response

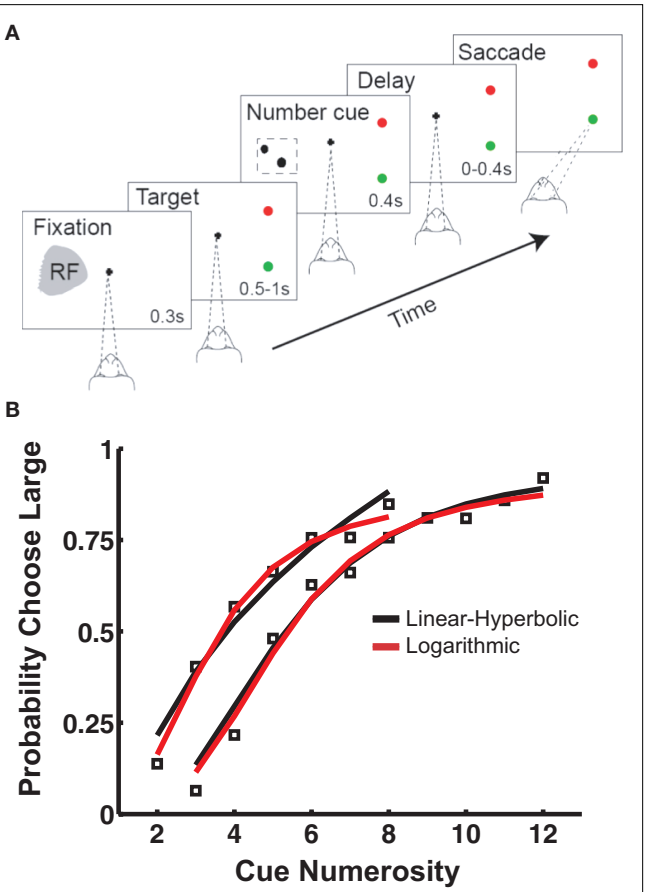


FIGURE 1 | Numerosity bisection task. (A) Schematic showing modeled oculomotor bisection task. Following fixation, two saccade targets appear: red for "small" and green for "large." After a variable delay, a dot array is briefly presented in the recorded neuron's response field. Following a second variable delay, the fixation target extinguishes, and the animal makes an eye movement to either choice target in the hemifield opposite the neuron's RF. **(B)** Choice behavior and model fits to a touch screen version of the numerosity bisection task (after Jordan and Brannon, 2006). Data points represent probability of choosing the response associated with the "large" anchor value. Red and black lines indicate fits based on families of choice curves derived from the linear-hyperbolic and logarithmic encoding models. Anchor values are 2 and 8 for the left set of curves, 3 and 12 for the right.

models, each of which contained neurons that increased and decreased firing in response to increasing numerosity. (Clearly, for numerosities within the range of the measured response curves, no fitting is necessary.) In the first model, these responses followed linear-hyperbolic tuning curves:

$$f_+ = an + b \tag{1a}$$

$$f_- = \frac{c}{n} + d \tag{1b}$$

while in the second, they followed logarithmic tuning curves:

$$f_+ = a \log n + b \tag{1c}$$

$$f_- = -c \log n + d \tag{1d}$$

Chi-squared values for fits to the measured mean response curves were calculated according to:

$$\tilde{\chi}^2 = \sum_i \frac{[f(n_i) - f_i]^2}{\nu\sigma_i^2} \quad (2)$$

where f_i are the measured firing rates, subscripts indicate positively (+) and negatively (−) monotonic responses, $f(n_i)$ are the model predictions, σ_i are the standard errors, and ν is the number of degrees of freedom. Noise was well fit by a Poisson process (**Figure 2D**) and, for simplicity, subsequently modeled as Gaussian with equivalent first and second moments:

$$f = \bar{f} + N(0, \bar{f}) \quad (3)$$

with $\bar{f}$ the tuning curve (either + or −) and N the normal distribution with variance $\bar{f}$. (Thus, although a real Poisson process will show deviations from this assumption, those deviations will only affect higher moments of the distribution.)

Choices were made by randomly sampling from both positive and negative tuning curves and taking differences in firing rates. As explained in the "Results" section, this firing rate difference was subsequently fed into a softmax choice model:

$$P_L(n) = \frac{ae^{\beta\delta f(n)}}{2a - 1 + e^{\beta\delta f(n)}} \quad (4)$$

with P_L the probability of choosing the option corresponding to the larger anchor, δf the difference in firing rates, β a parameter reflecting the variability of the animal's choice, and a a maximum choice preference for the option with higher firing. Indifference results when $\delta f = 0$.

We simulated distinct pairs of anchor values by shifting the relative baseline firing rates (bias input) of the positive and negative tuning curves, as detailed in the "Results" sections. This resulted in a one-parameter family of psychometric choice curves, differing in their points of subjective equality (PSE). For ease of computation, we parameterized this family of choice

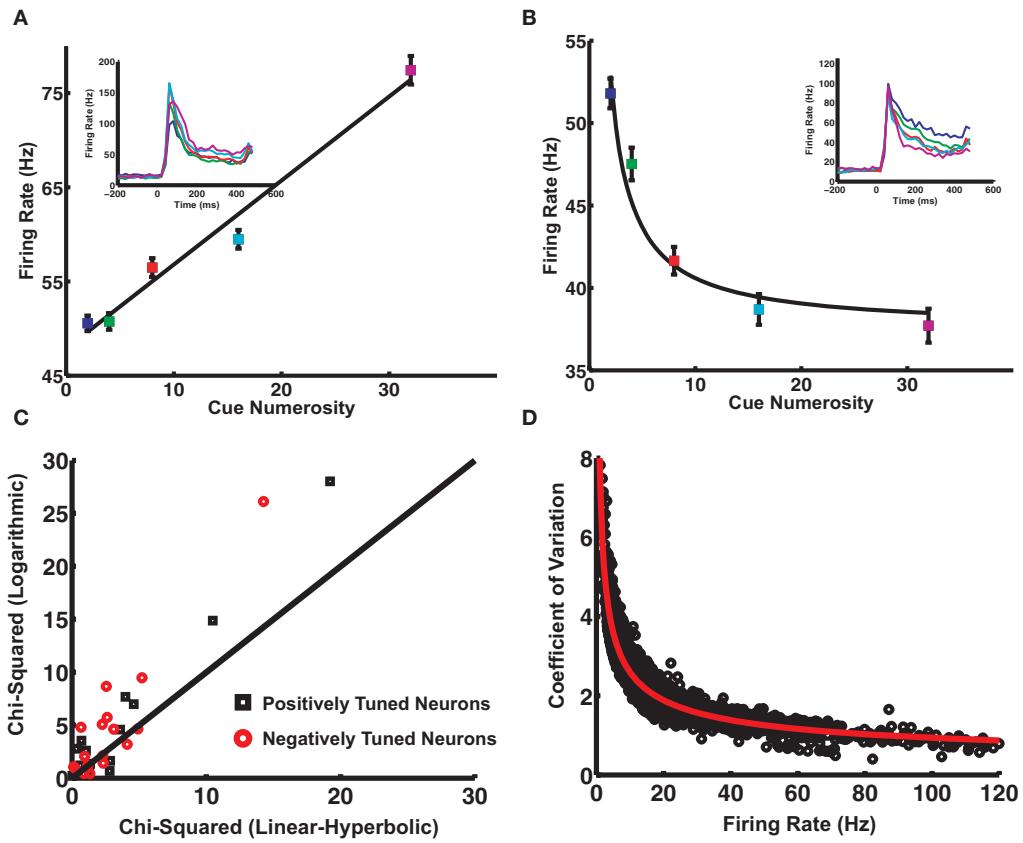


FIGURE 2 | Single neurons in LIP monotonically encode the number of visual objects. (A) Averaged single neuron (n = 15) firing rate during the cue presentation for five different numerosities (after Roitman et al., 2007; see Materials and Methods). Standard and deviant trials were averaged. Data are fit to a straight line. Inset shows PSTH of an example neuron for five different numerosities. Colors in the PSTH correspond to data points in the main plot. (B) Average firing rate (n = 17) for negatively monotonic neurons in LIP. Inset shows the PSTH for an example neuron. Data were fit to a hyperbolic decay. Colors are as in (A). (C) ^{2} goodness-of-fit comparison between linear-hyperbolic and logarithmic models for single neurons. Black line represents equally good fits. Black squares are positively monotonic neurons, red circles negatively monotonic. Smaller χ^2 values represent better fits. More symbols above the black line than below indicate the superiority of the linear-hyperbolic model as a fit to the data. (D) Coefficient of variation (mean/standard deviation) of firing rates as a function of firing rate during the 400-ms delay period following numerosity cue presentation in the implicit number task (see Materials and Methods). Each circle represents one trial for one neuron. Red line is a fit to the prediction of Poisson firing rates.

curves by two different functional forms (our results did not depend on the choice). We fit model-generated curves with both logistic:

$$P_L(n) = \frac{ae^{(n-n_*)/\sigma}}{(2a-1)+e^{(n-n_*)/\sigma}} \quad (5)$$

and Gompertz

$$P_L(n) = ae^{-\log(2a)e^{-(n-n_*)/\sigma}} \quad (6)$$

functions. As before, a represents a maximum preference level for the large option, while n_* represents the PSE and σ is a measure inversely related to discriminability. Once again, the fitting is a computational convenience, and the specific form of the parameterization does not matter. Results are unaffected if the direct outputs of the model are used instead. We fixed β and a in Eq. 4 by fitting our family of psychometric choice curves to the measured bisection behavior of monkeys in a separate experiment (Jordan and Brannon, 2006) (**Figure 1B**).

For both parameterizations of our family of choice curves, Weber's Law predicts:

$$\sigma \sim kn_* \quad (7)$$

with k a constant.

We modeled the process of learning the indifference point for bisection via a reinforcement learning algorithm that tracked the values of each of the two responses and updated these along with a "bias input" favoring either the "large" or "small" response. In this case, we parameterized our tuning functions as:

$$f_+ = a(n-n_*)+B \quad (8a)$$

$$f_- = c\left(\frac{1}{n}-\frac{1}{n_*}\right)+B \quad (8b)$$

for the linear-hyperbolic case and

$$f_+ = a\log\left(\frac{n}{n_*}\right)+B \quad (8c)$$

$$f_- = -c\log\left(\frac{n}{n_*}\right)+B \quad (8d)$$

in the logarithmic case, with n_* clearly equal to the point of subject equality (adjusted by the learning algorithm) and B a constant baseline firing rate common to both types of neurons. On each trial, either the large or small anchor was presented with equal probability, and the system made its response according to the output of the current decision model for the current value of n_*. As in the standard bisection task, only correct answers were rewarded. Subsequent to reward, the system performed the following updates for the action values corresponding to the two choices and the PSE:

$$Q_L \leftarrow Q_L + \lambda(R-Q_L) \quad \text{respond "large"} \quad (9a)$$

$$Q_S \leftarrow Q_S + \lambda(R-Q_S) \quad \text{respond "small"} \quad (9b)$$

$$n_* \leftarrow n_* + \alpha(Q_L-Q_S) \quad (9c)$$

with Q_L the action value of choosing "large," Q_S the action value of choosing "small," R the reward outcome (either 0 or 1, for incorrect or correct) and α and λ learning rates. Note that only the value corresponding to the chosen action is updated, though the PSE changes each trial. In this way, the PSE is adjusted upward (biasing toward the "large" response) if $Q_L > Q_S$ – in other words, in the direction of choosing the more profitable option. Clearly, equilibrium corresponds to equality of the two action values, at which point the animal should be indifferent, and reward is maximized. We report simulations performed for 15000 trials with both α and λ equal to 0.05. The initial value of the indifference point was set to the arithmetic mean of the anchors, though choosing either extreme worked equally well. Learning for most pairs of anchor values converges within 2000 trials, though mean PSEs and standard deviations were calculated over the last 4000 trials of simulation to ensure that learning had reached asymptote.

RESULTS

THE MODEL PREDICTS BISECTION BEHAVIOR IN THE ABSENCE OF EXPLICIT NUMEROSITY CODES

To model the response properties of neurons in LIP, we made use of single-unit neural activity recorded during an implicit numerosity discrimination task (see Materials and Methods). As shown in **Figures 2A,B**, firing rates in these neurons varied in both positively (n = 15/53) and negatively (n = 17/53) monotonic fashion with cue numerosity. Following previous theories of magnitude discrimination, we fit these neural response functions to two models (**Figures 2A,B**): In the first, the increasing response is modeled as linear ($f_+ = an + b$, a = 1.14, b = 45.2, χ^2 = 1.38), while the decreasing response is fit to a hyperbolic function ($f_- = c/n + d$, c = 30.7, d = 37.5, χ^2 = 1.18). This hyperbolic coding, not previously proposed for number, resembles that observed in primate superior colliculus when multiple, equally likely saccade targets are presented, and suggests, at least in part, an effective compression of one-half the internal representation of large numerosities (Basso and Wurtz, 1997).

In addition, we fit neuronal responses as logarithmically encoding numerosity ($f_+ = a\log n + b$, a = 9.01, b = 40.20, χ^2 = 7.40; $f_- = -c\log n + d$, c = 5.34, d = 54.6, χ^2 = 1.26). Clearly, both models reproduce the negatively monotonic curve well, though the logarithmic fit in the case of the increasing response function is somewhat less convincing (**Figure 2C**). However, since the logarithmic model possesses a number of interesting theoretical features and serves as an important contrast to the behavior of the linear-hyperbolic model, we report the results of our decision model in both cases. It is also important to note that such fits are only for computational convenience and the extrapolation of our predictions to high numerosity. Direct use of the empirical tuning curves produces equivalent behavior in our model over the range of numerosities tested. We also examined the variability of neuronal responses. As shown in **Figure 2D**, firing rates across the population were well fit by an assumption of Poisson noise (R^2 = 0.92), providing evidence against the scalar variance assumption of linear models for magnitude encoding.

A schematic for our decision model is presented in **Figure 3A**. We treat the decision process as a competition between two representations, one with positive response function, one with negative, to a dot array stimulus. Similar to models of interval discrimination

(Gibbon, 1977, 1981; Gibbon and Fairhurst, 1994), we sample from Poisson distributions of firing centered about these response curves, calculating the difference in firing between them. In this framework, high firing rates in positively monotonic neurons give evidence for a large encoded numerosity (and thus argue for a "large" response), whereas high firing in negatively monotonic neurons argues for a small numerosity in the stimulus. The difference between these two pieces of evidence then becomes the overall bias toward the "large" response. Mathematically, this is given by the difference in the two tuning curves (see Eqs. 8a,b in Materials and Methods):

$$\text{Bias} = \begin{cases} -an_* + \dfrac{c}{n_*} & \text{linear-hyperbolic} \\ (c-a)\log n_* & \text{logarithmic} \end{cases} \quad (10)$$

where n_* is the PSE. Clearly, this number is negative for many values of n_*, which case it represents either an inhibition of positively-tuned neurons or an upward shift of the negatively monotonic tuning curve. Clearly, this difference is unaffected if both firing rate responses are increased by the same amount, though such an overall shift does affect the amount of Poisson noise (and thus the variability of the signal).

THE MODEL PREDICTS DEVIATIONS FROM WEBER'S LAW BEHAVIOR AT LARGE NUMEROSITIES

This rudimentary firing rate difference model, using only two neurons, is capable of producing much less variable behavioral output than is typically observed in animals (Church and Deluty, 1977; Meck and Church, 1983; Platt and Davis, 1983; Meck et al., 1985; Roberts, 2005; Jordan and Brannon, 2006). That is, observed psychometric choice curves in bisection paradigms are much wider than those produced by our neurometric model, implying poorer classification performance than the LIP representation would necessitate.

Yet, it is not uncommon for animals to show much poorer asymptotic performance than discrimination models would predict. In fact, we argue that such noise is necessary to drive learning in the systems that are responsible for choice behavior (see below). As a result, we fed the results of the two-neuron comparison (the perceptual model) into a subsequent softmax action selection equation (the choice model) (Machens et al., 2005; Lo and Wang, 2006). This model incorporates both less-than-perfect asymptotic classifications of stimuli, as well as a substantial probability choices deviating from the underlying percept (for purposes of information-gathering about reward contingencies). This combined model is capable of producing excellent fits to behavioral data [PSE = 3.62, 5.37, R^2 = 0.98, 0.89 (linear-hyperbolic); PSE = 3.59, 5.43, R^2 = 0.99, 0.97 (logarithmic); **Figure 1B**].

As expected, the model is indifferent between responding "large" or "small" when firing rates for the two response curves are equal, that is, at their point of intersection. Clearly, this point may be shifted by adding a constant bias firing rate to either curve, resulting in a family of choice curves with increasing PSEs and broadening slopes (**Figures 3B,E**). These broadening curves represent decreased sensitivity to fixed differences in numerosity as PSE increases, with broader curves indicating a wider variance in task performance near the indifference point. That is, as the bias input in **Figure 3A** increases, discrimination between the presented numerosity and the classification threshold becomes poorer, as predicted by Weber's Law. In fact, the width of the curves in **Figures 3B,E** is inversely related

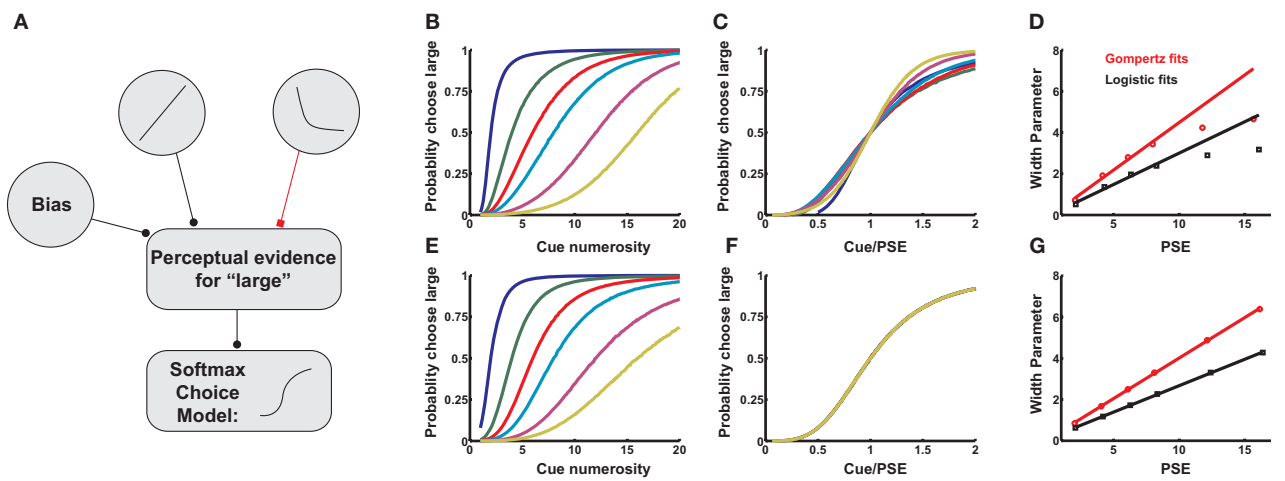


FIGURE 3 | Two models for numerosity bisection. (A) In both models, positive and negative neuronal responses from LIP neurons enter as evidence for "large" and "small" cue classifications, respectively. In addition, a bias input represents an additional propensity to choose one response over the other. In the figure, black lines represent excitation, red lines inhibition. The bias input augments the probability of perceiving the cue as "large," though in other cases it might favor a "small" outcome. Competition between the two tuning curves computes an effective difference in firing rates, which is passed on to a softmax choice model that apportions choices based on this difference (see Materials and Methods). **(B)** A family of choice curves generated by systematically decreasing the bias input in the model in **(A)**. As bias decreases, the point of subjective equality (PSE) shifts rightward, and the transition region interpolating between "small" and "large" classifications widens. **(C)** Choice curves from **(B)**, rescaled as a function of cue numerosity divided by PSE. Curves roughly overlap, indicating an approximate Weber's Law behavior. **(D)** Width Parameter (width of the transition region, inversely related to discriminability) as a function of PSE. Red data points indicate parameters derived from fitting choice curves with a Gompertz function, black points logistic fits (see Materials and Methods). Lines are fitted to the first four data points, and show a deviation from linearity at high PSE, predicting a departure from Weber's Law. Slopes of lines are equal to Weber fractions calculated in the models. **(E–G)** Same as **(B–D)**, except tuning curve inputs to **(A)** are modeled as positive and negative logarithms. As expected, Weber's Law holds to all orders.

to the discriminability of cue numerosity from the indifference point, and is expected to scale linearly with PSE. **Figures 3B,E** depict the resulting relationship between discriminability and PSE for a series of bias inputs to the network for both linear-hyperbolic and logarithmic response models. As predicted, the logarithmic model produces a precisely linear relationship between the two quantities, reproducing Weber's Law at all numerosities (**Figures 3F,G**). In the case of the linear-hyperbolic model, the relationship is approximately linear for small numerosities, but falls well short of linearity as the PSE increases (**Figure 3D**). This results from higher effective variance in the encoded numerosity in the linear tuning curve of the model (again, the variance in the logarithmically-encoded numerosity is constant), which results in a higher rate of misclassifications near the large anchor in **Figures 3B,C**. In principle, this violation of Weber's Law behavior would allow one to distinguish between the two models experimentally. However, since observed indifference points lie near the geometric means of anchor values, and since the largest measured PSEs to date are less than 8 (Jordan and Brannon, 2006), the anchor numerosities required to observe these predicted departures will necessarily be much higher than those thus far probed empirically. Most importantly, the model facilitates flexible classification behavior in the case of different anchor values by the adjustment of a single parameter, the PSE (see below).

REINFORCEMENT LEARNING DRIVES THE MODEL TO PSES AT THE GEOMETRIC MEAN, AS OBSERVED BEHAVIORALLY

To further investigate the adjusting bias model as a means of adapting to differing anchor values, we implemented a reinforcement learning algorithm designed to set the bias input (and thus the PSE) of the system based on maximizing reward. In our implementation, the animal learns three quantities: the values of both the "small" and "large" responses and the value of the bias input. The first two are updated by a traditional reward prediction error delta rule (see Materials and Methods), while the last is updated based on the difference in updated values of the two options. In addition, because only the anchors are rewarded, the algorithm never relies on an explicit knowledge of the full choice curve, only the values associated with choosing the "small" or "large" options. Thus, rather than treating the task as a perceptual discrimination, our algorithm seeks to maximize reward, which allows it to generalize to cases in which correct responses are only probabilistically rewarded or responses to the options are rewarded unequally. Indeed, for these latter cases, we predict that PSEs will not remain at the geometric mean, but will shift in order to maximize the reward harvested by our decision model's choice behavior.

Moreover, we note the importance of additional noise in our choice model for the convergence of the algorithm. Because we expect behavioral responses to anchor values to be dominated by the nonlinear "knees" of our choice curves, the convergence behavior of our learning model will exhibit high sensitivity to the slopes of the curves in these regions. If the curves are virtually noiseless, transitioning abruptly from "small" to "large" responses, learning will plateau rapidly, since any PSE located between the anchors will produce near-perfect classification of the extremes. Thus there is an inverse relationship between sensitivity of the choice curve (inversely proportional to its width parameter, σ, and proportional to the slope of its rise) and sensitivity of reward returns to PSE location: for a perfect classifier, performance at the anchors is all but insensitive to PSE location, while the performance of a noisy classifier depends heavily on the location of the indifference point. For this reason, and because choices in natural environments involve the classification of intermediate numerosities, learning favors the introduction of additional noise into the choice process beyond that inherent in the perceptual mechanism.

Figure 4 depicts the results of a series of simulations conducted for both the linear-hyperbolic and logarithmic models. **Figures 4A,B** show example learning curves for learning bisection with anchor values 3 and 12. After about 2000 trials, the first PSE (**Figure 4A**) converges to a mean value of around 5.3, just below the predicted value of 6, and in line with the slight deviation seen in **Figure 1B**. In **Figure 4B**, the PSE converges to the theoretical value of 6. In **Figures 4C,D**, we plot PSE values for a series of simulations performed for fixed values of the small anchor. If the PSE scales as the geometric mean of the anchor values, as theory predicts, the resultant curves should scale as the square root of the large anchor, which they do. However, the linear-hyperbolic model shows clear deviations from predicted behavior for large absolute differences between anchors, a reflection of the fact that choice curves are asymmetric, with more accurate classification of smaller numerosities. As a result, reward-maximizing PSEs systematically undershoot geometric means as the distance between anchors grows, a trend consistent with that seen in experimental studies (Jordan and Brannon, 2006) for anchor pairs (2,8) and (3,12) (**Figure 1B**). In a similar vein, **Figures 4E,F** show results of simulations with fixed ratio of small to large anchor values. In this case, theory predicts that the PSE should scale linearly as with the small anchor value, which approximately holds.

DISCUSSION

Our model of numerosity encoding in the bisection paradigm takes as its starting point the measured monotonic response functions and spiking statistics of neurons in LIP. Though these neurons conform to neither the linear/scalar variance nor logarithmic/constant variance models of graded numerosity encoding previously proposed, we are able, using a simple decision rule in conjunction with a hypothesized bias input, to reproduce observed bisection behavior. In addition, we are able to predict adherence to Weber's Law over a significant range of anchor value pairs. However, the differences between our model and previous proposals are illuminating, and offer predictions for future experiments. In the case of our linear/hyperbolic model (again, to be distinguished from the logarithmically widening preferred-numerosity responses in population coding models), we predict gradual deviations from Weber's Law behavior at very large numerosities, corresponding to PSEs of 10 or more. In our logarithmic model, we expect to see increasing nonlinearity in neuronal responses for very large numerosities, though we do not expect increasing Poisson noise to disrupt the Weber's Law property (see Supplementary Material). In both cases, we expect a constant relative shift in firing rates between the response curves for different pairs of task anchor values (and thus different PSEs), a key prediction of the model testable in future experiments.

In addition, we hypothesize that the disparity between measured task performance in animals and the classification behavior of an ideal observer using our neuronal data is due, at least in part, to additional noise added in the response selection process. We argue that this

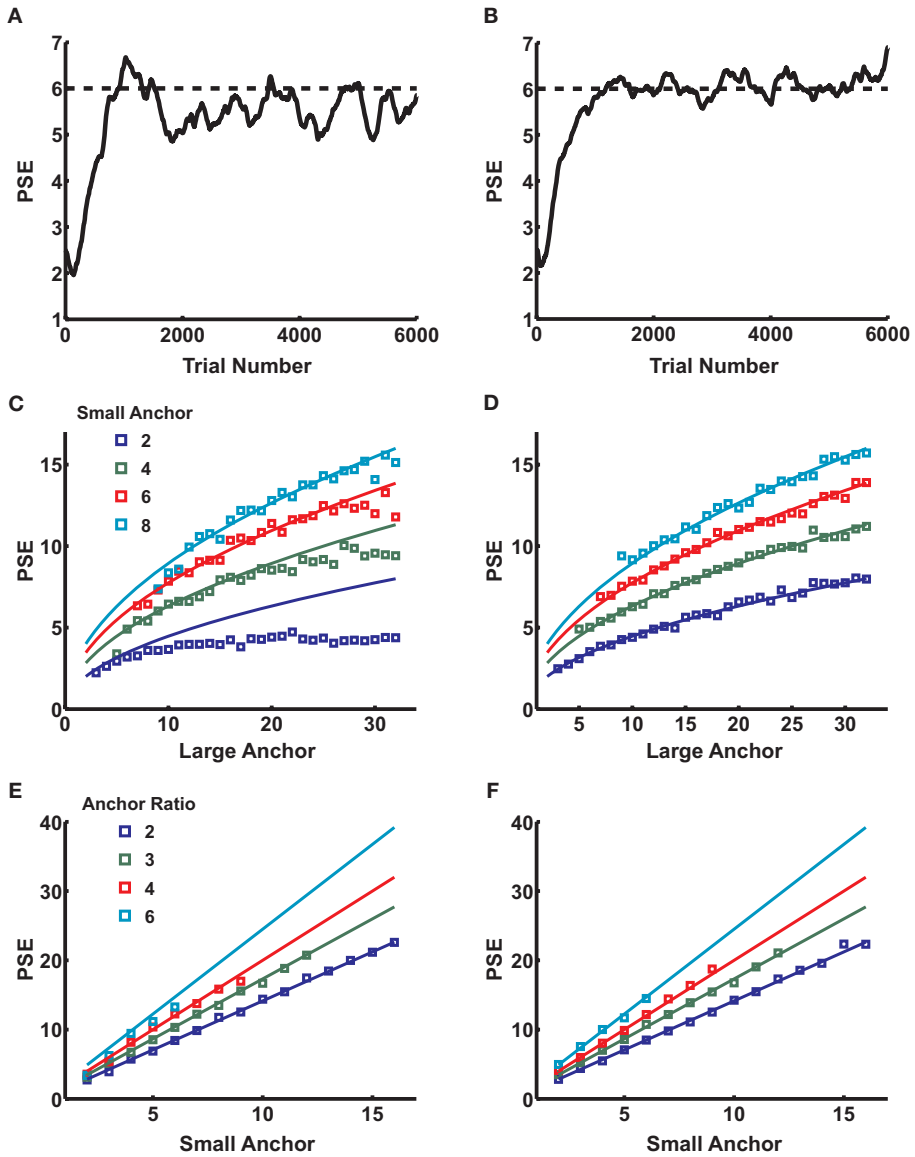


FIGURE 4 | A simple learning algorithm reproduces location of the PSE. **(A)** A sample learning curve for the first 6000 trials of a bisection block with anchors 3 and 12. Dotted line marks the theoretical value of 6. Value learned by the algorithm with linear-hyperbolic inputs is around 5.3. **(B)** Same as **(A)** for the logarithmic model. Note that the mean PSE value is now equal to the geometric mean of the anchors. **(C)** PSE as a function of large anchor value for a series of small anchor values (different colors). Lines represent theoretical predictions based on the geometric mean formula. All series show increasing departures from Weber's Law behavior at high numerosities, particularly for small anchor value 2. **(D)** Same as **(C)** for the logarithmic model. **(E)** PSE as a function of small anchor value for fixed ratios of large to small anchor values (different colors). Simulated data are roughly in line with linear predictions based on the geometric mean formula over the range tested. **(F)** Same as **(E)** for the logarithmic model.

noise, which often results in choices the animal should "know" are wrong, is needed by the reinforcement learning algorithm that learns the task's reward contingencies and the location of the PSE. Because greater sampling from both options leads to better estimates of each option's value, less accurate choice behavior, paradoxically, leads to grater optimality in choosing the location of the PSE that results in maximum reward. Indeed, we conjecture that this need for flexible learning algorithms may explain similar discrepancies between ideal-observer and measured animal behavior in other classification tasks (Shadlen and Newsome, 2001). Finally, our algorithm is noteworthy in that it makes no use of "right" or "wrong" classification behavior, nor requires explicit knowledge of the underlying classification rule. Choosers simply learn the average value of responses in the presence of stimuli, and update the internal model accordingly. As a result, task performance may be viewed through the lens of reward maximization, and our algorithm makes predictions for cases in which responses are differentially or probabilistically rewarded.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/behavioralneuroscience/paper/10.3389/neuro.08/001.2010/

REFERENCES

Andersen, R., and Buneo, C. (2002). Intentional maps in posterior parietal cortex. *Annu. Rev. Neurosci.* 25, 189–220.

Basso, M., and Wurtz, R. (1997). Modulation of neuronal activity by target uncertainty. *Nature* 389, 66–68.

Brannon, E., and Terrace, H. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science* 282, 746.

Brannon, E., Wusthoff, C., Gallistel, C., and Gibbon, J. (2001). Numerical subtraction in the pigeon: evidence for a linear subjective number scale. *Psychol. Sci.* 12, 238–243.

Church, R., and Deluty, M. (1977). Bisection of temporal intervals. *J. Exp. Psychol. Anim. Behav. Process.* 3, 216–228.

Dehaene, S., and Changeux, J. (1993). Development of elementary numerical abilities: a neuronal model. *J. Cogn. Neurosci.* 5, 390–407.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* 84, 279–325.

Gibbon, J. (1981). On the form and location of the psychometric bisection function for time. *J. Math. Psychol.* 24, 58–87.

Gibbon, J., and Church, R. (1981). Time left: linear versus logarithmic subjective time. *J. Exp. Psychol. Anim. Behav. Process.* 7, 87.

Gibbon, J., and Fairhurst, S. (1994). Ratio versus difference comparators in choice. *J. Exp. Anal. Behav.* 62, 409.

Gold, J., and Shadlen, M. (2000). Representation of a perceptual decision in developing oculomotor commands. *Nature* 404, 390–394.

Green, D., and Swets, J. (1989). Signal Detection Theory and Psychophysics. Peninsula Publishing. Available at: http://www.amazon.com/Signal-Detection-Theory-Psychophysics-Marvin/dp/0932146236/ref=sr_1_1?ie=UTF8&s=books&qid=1263939907&sr=8-1.

Grossberg, S., and Repin, D. (2003). A neural model of how the brain represents and compares multi-digit numbers: spatial and categorical processes. *Neural Netw.* 16, 1107–1140.

Jordan, K., and Brannon, E. (2006). Weber's law influences numerical representations in rhesus macaques (*Macaca mulatta*). *Anim. Cogn.* 9, 159–172.

Leon, M., and Shadlen, M. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron* 38, 317–327.

Lo, C., and Wang, X. (2006). Cortico–basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* 9, 956–963.

Machens, C., Romo, R., and Brody, C. (2005). Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Am. Assoc. Adv. Sci.* 307, 1121–1124.

Meck, W., and Church, R. (1983). A mode control model of counting and timing processes. *J. Exp. Psychol. Anim. Behav. Process.* 9, 320–334.

Meck, W., Church, R., and Gibbon, J. (1985). Temporal integration in duration and number discrimination. *J. Exp. Psychol. Anim. Behav. Process.* 11, 591.

Miller, P., Brody, C., Romo, R., and Wang, X. (2003). A recurrent network model of somatosensory parametric working memory in the prefrontal cortex. *Cereb. Cortex* 13, 1208–1218.

Moyer, R., and Landauer, T. (1967). Time required for judgements of numerical inequality. *Nature* 215, 1519.

Nieder, A. (2005). Counting on neurons: the neurobiology of numerical competence. *Nat. Rev. Neurosci.* 6, 177–190.

Nieder, A., Freedman, D., and Miller, E. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science* 297, 1708.

Nieder, A., and Merten, K. (2007). A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *J. Neurosci.* 27, 5986.

Nieder, A., and Miller, E. (2003). Coding of cognitive magnitude compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* 37, 149–157.

Nieder, A., and Miller, E. (2004a). A parieto-frontal network for visual numerical information in the monkey. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7457–7462.

Nieder, A., and Miller, E. (2004b). Analog numerical representations in rhesus monkeys: evidence for parallel processing. *J. Cogn. Neurosci.* 16, 889–901.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., and Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron* 44, 547–555.

Platt, J., and Davis, E. (1983). Bisection of temporal intervals by pigeons. *J. Exp. Psychol. Anim. Behav. Process.* 9, 160–170.

Platt, M., and Glimcher, P. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233.

Roberts, W. (2005). How do pigeons represent numbers? Studies of number scale bisection. *Behav. Processes* 69, 33–43.

Roitman, J., Brannon, E., and Platt, M. (2007). Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS Biol.* 5, e208. doi: 10.1371/journal.pbio.0050208.

Roitman, J., and Shadlen, M. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475.

Shadlen, M., and Newsome, W. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* 86, 1916–1936.

Snyder, L., Batista, A., and Andersen, R. (1997). Coding of intention in the posterior parietal cortex. *Nature* 386, 167–170.

Stevens, S. (1986). Psychophysics: introduction to its perceptual, neural, and social prospects. Transaction Publishers. New Brunswick.

Verguts, T., and Fias, W. (2004). Representation of number in animals and humans: a neural model. *J. Cogn. Neurosci.* 16, 1493–1504.

Verguts, T., Fias, W., and Stevens, M. (2005). A model of exact small-number representation. *Psychon. Bull. Rev.* 12, 66–80.

Zorzi, M., and Butterworth, B. (1999). A computational model of number comparison. In Proceedings of the Twenty First Annual Conference of the Cognitive Science Society, M. Hahn and S. C. Stoness, eds (Mahwah, NJ, Erlbaum), p. 778.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 September 2009; paper pending published: 12 November 2009; accepted: 07 January 2010; published online: 27 January 2010.

*Citation: Pearson J, Roitman JD, Brannon EM, Platt ML and Raghavachari S (2010) A physiologically-inspired model of numerical classification based on graded stimulus coding. Front. Behav. Neurosci. **4**:1. doi: 10.3389/neuro.08.001.2010*

frontiers in
BEHAVIORAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 27 August 2009
doi: 10.3389/neuro.08.021.2009

Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses

R. McKell Carter[1,2†], Jeff J. MacInnes[1†], Scott A. Huettel[1,3]* and R. Alison Adcock[1,4]*

[1] *Center for Cognitive Neuroscience, Duke University, Durham, NC, USA*
[2] *Department of Neurobiology, Duke University Medical Center, Durham, NC, USA*
[3] *Psychology and Neuroscience, Duke University, Durham, NC, USA*
[4] *Department of Psychiatry, Duke University Medical Center, Durham, NC, USA*

Edited by:
Daeyeol Lee, Yale University School of Medicine, USA

Reviewed by:
Hyojung Seo, Yale University School of Medicine, USA
Kenji Kansaku, Research Institute of National Rehabilitation Center for Persons with Disabilities, Japan

***Correspondence:**
Scott A. Huettel, B203 Levine Science Research Center, Duke University, Box 90999, Durham, NC 27708, USA.
e-mail: scott.huettel@duke.edu;
R. Alison Adcock, B203 Levine Science Research Center, Duke University, Box 90999, Durham, NC 27708, USA.
e-mail: alison.adcock@duke.edu

†*R. McKell Carter and Jeff J. MacInnes have contributed equally to the manuscript.*

To represent value for learning and decision making, the brain must encode information about both the motivational relevance and affective valence of anticipated outcomes. The nucleus accumbens (NAcc) and ventral tegmental area (VTA) are thought to play key roles in representing these and other aspects of valuation. Here, we manipulated the valence (i.e., monetary gain or loss) and personal relevance (i.e., self-directed or charity-directed) of anticipated outcomes within a variant of the monetary incentive delay task. We scanned young-adult participants using functional magnetic resonance imaging (fMRI), utilizing imaging parameters targeted for the NAcc and VTA. For both self-directed and charity-directed trials, activation in the NAcc and VTA increased to anticipated gains, as predicted by prior work, but also increased to anticipated losses. Moreover, the magnitude of responses in both regions was positively correlated for gains and losses, across participants, while an independent reward-sensitivity covariate predicted the relative difference between and gain- and loss-related activation on self-directed trials. These results are inconsistent with the interpretation that these regions reflect anticipation of only positive-valence events. Instead, they indicate that anticipatory activation in reward-related regions largely reflects the motivational relevance of an upcoming event.

Keywords: reward, valence, incentive, anticipation, striatum, dopamine, neuroimaging

INTRODUCTION

Neural representations of anticipated reward value are core to models of the mechanisms for learning (Schultz et al., 1997; Sutton and Barto, 1998; O'Doherty et al., 2004; Seymour et al., 2004) and decision making (Montague and Berns, 2002; Bayer and Glimcher, 2005; Balleine et al., 2007; Rangel et al., 2008). These models associate predictive cues with their subsequent outcomes, in order to describe behavior. Accordingly, the subjective experience of the cue-outcome association prior to the occurrence of the outcome reflects "anticipation".

The most common functional neuroimaging paradigms for studying reward anticipation use learned cue-response-outcome contingencies (Delgado et al., 2000; Knutson et al., 2001, 2005). On each trial an initial cue indicates a potential reward (e.g., a monetary gain). Then, following a short delay, a target appears, and if participants respond sufficiently quickly and/or accurately, they receive a reward. Studies using variants of this approach have demonstrated that the ventral striatum (vSTR), particularly its nucleus accumbens (NAcc), exhibits increases in blood oxygenation level-dependent (BOLD) contrast (hereafter, "activation") to anticipated rewards (Knutson et al., 2000; Ernst et al., 2005; Adcock et al., 2006; Knutson and Gibbs, 2007; Dillon et al., 2008). Yet, despite the prevalence of this approach, several important questions about reward anticipation remain incompletely answered: How do these findings generalize to other regions within the dopaminergic system (e.g., the ventral tegmental area, VTA)? Does activation of these regions reflect the motivational salience of cued stimuli (i.e., imperative for action) or the affective properties of the anticipated reward (i.e., valence)? And, is anticipatory activation modulated by decreases in motivational salience if magnitude and valence are held constant? To address these questions we examined brain activation during anticipation of rewards that varied in valence and in personal relevance. Decreased personal relevance should reduce – but not eliminate – motivational salience, while leaving magnitude and valence unchanged.

HOW DOES VTA CONTRIBUTE TO REWARD ANTICIPATION AND LEARNING?

The neural mechanisms that underlie motivation depend on activity of neurons in the NAcc (Wise, 1980, 2004; Kalivas et al., 2005; Berridge, 2007; Salamone et al., 2007), which are themselves modulate by dopaminergic producing neurons in the VTA (Swanson, 1982; Ikemoto, 2007). While much is known about VTA function from single-unit recordings in non-human animals, there have been relatively few neuroimaging studies that report effects in the VTA, largely because of technical constraints. The VTA is a small nucleus within the midbrain, and its boundaries with adjacent nuclei are not readily visible on standard structural magnetic resonance images. Researchers targeting the VTA, therefore, have used a combination of anatomical region-of-interest (ROI) analyses and targeted pulse sequences (e.g., inferior slices, tilted orientation). As one initial example, research by Adcock et al. (2006) evaluated the potential

modulatory role of VTA in shaping memory, demonstrating specifically improved recall for stimuli associated with greater potential rewards. Using a combination of standard regression analyses and functional connectivity measures, they found that voxels within the anatomical location of VTA both increased in activation to larger potential rewards and exhibited functional connectivity with the hippocampus in effective memory formation.

More recently, D'Ardenne et al. (2008) describe VTA responses to the experience of primary and secondary rewards, as a functional neuroimaging analog of the prediction error signals previously reported in single-unit recordings (Ljungberg et al., 1992; Schultz et al., 1997; Bayer and Glimcher, 2005). They found that VTA activation increased to unexpected rewards, both primary (liquid) and secondary (money), consistent with single-unit studies showing that its neurons convey a positive reward prediction error. Of note, D'Ardenne et al. found no significant changes in VTA activation to the omission of an expected liquid reward nor to an unexpected monetary loss, as would be expected if that region also signaled negative reward prediction errors. Where imaging volumes have allowed, some prior studies have reported qualitatively similar results in both NAcc and midbrain (Knutson et al., 2005) and NAcc and VTA (Moll et al., 2006), although a systematic comparison is needed.

WHAT DOES NEURAL ACTIVITY DURING REWARD ANTICIPATION REPRESENT?

Understanding how the brain encodes, represents, and manipulates signals that indicate potential and experienced rewards has been an area of considerable basic (Montague and Berns, 2002; Bayer and Glimcher, 2005; Phillips et al., 2007; Delgado et al., 2008a,b, Knutson and Greer, 2008) and clinical research (Kilts et al., 2001; Grusser et al., 2004; Kienast and Heinz, 2006; Bjork et al., 2008; Knutson et al., 2008a, Schlagenhauf et al., 2008; Scott et al., 2008; Strohle et al., 2008; Pizzagalli et al., 2009). The common thread in this extensive literature is that the neural representation of reward does not reflect any simple unitary construct. In particular, there has been an ongoing debate about whether and how the brain represents two different aspects of reward. The first aspect is the absolute value of the outcome (i.e., important vs. unimportant outcomes), referred to as *energization* (Elliot, 2006), *salience* (Zink et al., 2003), *incentive salience* (Berridge et al., 2009), and *magnitude* (Knutson et al., 2001). A second aspect differentiates positive from negative outcomes; this aspect has been described in terms of *affect* (Knutson and Greer, 2008), *valence*, and *approach/avoidance* (Elliot, 2006). In the current paper, we will refer to these two aspects, which we intended to manipulate separately, as *motivation* and *affective valence*.

In the influential framework advanced by Berridge and colleagues, there are functional and neural dissociations between the valenced and non-valenced aspects of reward (Berridge, 2004; Berridge et al., 2009). Specifically, these authors contend that the response of dopaminergic neurons in the VTA and NAcc reflect a motivational signal associated with information about future rewards (i.e., "wanting" the reward). In contrast, other neurotransmitters (e.g., opioids) affect the valence component of reward [i.e., "liking" the reward (Wise, 1980)]; they make pleasurable stimuli more pleasurable and aversive experiences less aversive (Pecina and Berridge, 2005). These potentially dissociable concepts – motivational significance and affective valence – recur in functional neuroimaging studies of reward anticipation and experience, although some reports discuss activation in these brain regions from the perspective of approach/avoidance behavior (Elliot, 2006), others invoke changes in affect evoked by rewards (Knutson and Greer, 2008), and still others consider responses in these regions as markers of prediction error [both valenced and non-valenced (O'Doherty et al., 2004; Seymour et al., 2007)].

Despite this ongoing debate, motivation and affective valence can be difficult to tease apart experimentally. Rewards in neuroeconomic research are commonly monetary gains implemented in paradigms where they have both motivational significance and affective valence. The resulting activation in NAcc, VTA, or other reward-related regions may thus be attributed to either motivation or valence. Some reports indicate that stimuli of similar motivational significance but different valence (e.g., monetary gains and losses) evoke similar activation in reward-related regions. For example, Cooper and Knutson (2008) show that when an outcome is uncertain, activation in the NAcc increases for both gain and loss anticipation. Other studies have suggested that activation in some components of the reward system does indeed depend on valence, whether because of distinct spatial loci evoked by positive and negative stimuli (Seymour et al., 2007) or because of decreases in activation to negative events (Breiter et al., 2001). Tom et al. (2007) tracked parametric effects of gain and loss magnitudes in a loss-aversion paradigm, and found that activation in regions including the vSTR increased with magnitude for decisions about potential gains and decreased with magnitude for potential losses. Based on these and other conflicts in the literature, how motivation and affective valence information interact within the multiple regions that constitute the reward system remains unknown.

ARE THE NEURAL SUBSTRATES OF OUTCOME ANTICIPATION SIMILAR WHEN PLAYING FOR SELF AND OTHERS?

Finally, there exists considerable evidence that anticipatory activation, at least in the NAcc, generalizes across a wide range of rewards. Most neuroimaging studies of reward have used monetary outcomes, typically repeated opportunities to gain or lose about a dollar (Knutson et al., 2001; Daw, 2007). Yet, similar patterns of NAcc activation can be evoked using fluid rewards (Valentin et al., 2007), food items (Hare et al., 2008, 2009), valuable consumer goods (Knutson et al., 2007, 2008b), social cooperation (Rilling et al., 2002), and even the opportunity to punish others (Singer et al., 2006). Recent studies have related the increases in NAcc activation preceding a decision to the value of rewards earned for others (Moll et al., 2006; Harbaugh et al., 2007). Based on these studies, one natural conclusion is that any anticipated reward, even one with reduced personal relevance (and thus motivational salience), would evoke activation in multiple regions within the reward system (e.g., NAcc and VTA). While plausible, this conjecture has not yet been demonstrated.

OVERVIEW OF THE CURRENT EXPERIMENT

In the current study, we manipulated the valence (i.e., gain vs. loss) and motivational relevance (i.e., oneself vs. charity as beneficiary) of anticipated rewards, using an incentive-compatible response-time

game modeled on common paradigms in the literature (Knutson et al., 2000, 2001). In these paradigms, the trial cue is the earliest possible predictor of the potential gain or loss, and thus initiates anticipation. We focus on reward anticipation, rather than reward outcome, because the motivational and affective explanations for reward-system activation make clear and opposing predictions. If motivational influences alone drive activation during anticipation, and if manipulating the beneficiary of the reward changes the motivational salience (Mobbs et al., 2009), then gain- and loss-related activation should be positively correlated across individuals, with greater responses observed to self- compared to other-directed outcomes. Conversely, if affective valence alone determines anticipatory activation, activation should be greatest when playing for gains and least when playing to avoid losses (relative to neutral outcomes), but with no differences between Self and Charity treatments. Moreover, by assessing participants' reward sensitivity and other-regarding preferences, we obtained independent predictors of individual differences in the neural responses to each reward type.

This paradigm can also test predictions of temporal difference (TD) models of anticipatory association. According to common TD models (Sutton and Barto, 1990), a well-learned reward cue should evoke activation that reflects the value of the expected outcome. This prediction error signal can be described in terms of the value of the associated outcome (i.e., valenced) or the association value (i.e., the strength of the prediction), as discussed further below. In the case where prediction error is valenced, a pattern similar to the valence interpretation of anticipation would be expected: positive for gains, negative for losses. In the case where prediction error mirrors the strength of the association a result similar to the motivational salience signal model would be expected: positive for both gains and losses, with neutral cues producing the least activation. Importantly, both prediction error accounts would dictate identical results in charity and self conditions, unless the predictive system also represented the motivational significance of the cues.

MATERIALS AND METHODS

PARTICIPANTS

Twenty young adults (mean age 24 years; range 19–29 years; 10 females) participated in this study. Two were excluded because of misalignments in acquisition coverage, and one was excluded due to a Beck Depression Inventory (BDI) score indicating depression, leaving 17 participants in the reported data. All participants provided informed consent under a protocol approved by the Duke Medical Center Internal Review Board.

EXPERIMENTAL PROCEDURE

The experimental session comprised initial selection of a charity, task training outside the scanner, an fMRI session using a reward anticipation task, and completion of questionnaires to assess reward attitudes.

Following informed consent, subjects read descriptions of four non-profit organizations – Easter Seals, Durham Literacy Center, Animal Protection Society, and the American Red Cross – and then selected one as their charitable target. They were then provided full information about the task structure and payment contingencies (see below for task details), and were told that no deception was used in the experiment. All participants reported that they understood the task procedures and that they believed that their earnings for charity would go to the selected target. Before entering the scanner, they completed one practice run of the task using only gain trials. We separated gain trials and loss trials into different runs, to minimize cue conflict. Then, the participants were taken to the scanner for the MRI session. During acquisition of initial structural images, each participant completed a second practice run (using only loss trials). Participants then completed four 7-min task runs during collection of fMRI data. The first run always involved monetary gains, so that subjects built up balances within cumulative banks, and the second run always involved monetary losses. The last two runs consisted of one gain run and one loss run, with their order randomly determined.

Each run consisted of 50 trials (**Figure 1**), evenly split between five conditions according to potential outcome: Self $4, Charity $4, Self $0, Charity $0, and Neutral Control $0. Every trial began with a 500-ms cue whose composition indicated the target (picture), monetary amount at stake [background color: red (Self) or blue (Charity) for $4, yellow for $0 control conditions], and valence (gain: square frame, loss: circular frame). Following a variable delay of between 4 and 4.5 s, a target appeared on the screen. The subject's task was to respond by pressing a button with the index finger of the right hand, before the target disappeared. Within gain runs, responses that were sufficiently fast added $4 to the subject's or charity's bank (visually indicated by a coin), and responses that were longer than the current threshold had no financial consequences (visually indicated by a '0'). Within loss runs, responses that were sufficiently fast resulted in no financial consequences (visually indicated by a '0'), whereas responses that were longer than the current threshold subtracted $4 from the subject's or charity's bank (visually indicated by a red circle with a diagonal line). The presentation time of the target was determined by an adaptive algorithm; using information about response times on previous similar trials, the algorithm estimated the response time threshold at which the subject would be successful on approximately 65% of trials. We emphasize that independent thresholds were used for each trial type.

At the end of all runs, the participants exited the scanner and completed a series of behavioral questionnaires (see below). Participants were paid a base sum of $15. In addition, cumulative bank totals were calculated for both the participant (*M* $22.35, SD 11.75) and charity (*M* $22.59, SD $6.62), and participants were paid the full amount of their bank in cash (participants were guaranteed a minimum of $40 for participation). Following completion of data collection from all subjects, the researchers paid the cumulative earnings to each charity.

BEHAVIORAL QUESTIONNAIRES

After completing the experiment, participants were asked to fill out a series of psychological questionnaires. These included: the BDI (a screening tool for depression) (Beck et al., 1961); Behavioral Inhibition System/Behavioral Activation System (BIS/BAS, an index of approach and avoidance tendencies) (Carver and White, 1994); Interpersonal Reactivity Index (IRI, an assessment of other-regarding *behavior*) (Davis, 1983); Personal Altruism Level (PAL, a questionnaire using indices of other-regarding personal *efforts*) (Tankersley et al., 2007); Self Report Altruism Scale (SRAS, an index

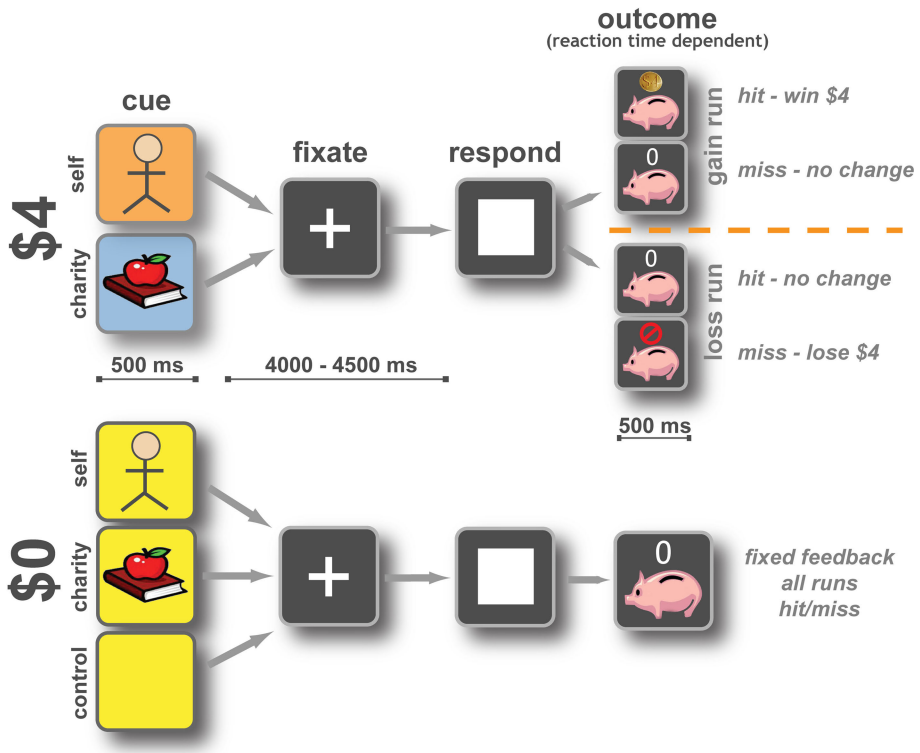


FIGURE 1 | Participants performed a monetary incentive reaction time task. An initial cue marked the start of the trial and indicated whether money was at stake and, if so, who would receive it. Each trial offered either $4 or $0, for the participant (Self), a charity (Charity), or no one. Gain and loss outcomes occurred in separate runs, to minimize cue conflict. After a variable wait (4–4.5 s) a response target appeared indicating that participants were to press a button using their right index finger as quickly as possible. The trial was scored as a hit if the participant responded in time or as a miss if they did not. Changes to the bank as a result of that trial were then displayed for 0.5 s. In gain runs on $4 trials, if the subject responded to the target in time they won $4 for themselves or a charity, if they missed the trial there was no change to that bank. During loss runs on $4 trials, if the subject responded to the target in time there was no change to that bank, if they responded too slowly, they lost $4 for either themselves or their charity. Control trials resulted in no change to the bank but participants were asked to respond as quickly as possible. Reaction time thresholds for hits and misses were set using an adaptive algorithm to allow the subject to win approximately 65% of the time. Thresholds were set independently for each trial type.

of other-regarding *preferences*) (Rushton et al., 1981); Temporal Experience of Pleasure Scale (TEPS, an index of reward experience and anticipation) (Gard et al., 2006). By taking the average of *Z*-score-transformed subscales from these measures, we constructed three individual-difference covariates. We defined the covariates based on *a priori* relations between the above scales: a personal *reward-sensitivity* covariate (BAS and TEPS, combined); an *other-regarding preference* covariate (PAL, IRI, and SRAS); and a *behavioral inhibition* covariate (BIS and BDI). A factor analysis presented by Pulos et al. (2004) suggests the personal-distress subscale of the IRI, included in our other-regarding preference covariate, may differ from the other-regarding trait targeted by the rest of the included subscales. Therefore, as a control test, we also evaluated a more limited empathy covariate that eliminated the personal distress subscale from the IRI.

fMRI ACQUISITION AND PREPROCESSING

At the beginning of the scanning session, we collected initial localizer images to identify the participant's head position within the scanner, followed by IR-SPGR high-resolution whole-volume T1-weighted images to aid in normalization and registration (voxel size: 1 mm × 1 mm × 1 mm). We also collected 17 slice IR-SPGR images, coplanar with the BOLD contrast images described below, for use in registration and normalization.

We collected BOLD contrast images acquired using a standard echo-planar sequence on a 3T GE Signa MRI scanner. Each of the four runs comprised 416 volumes (TR: 1 s; TE: 27 ms; Flip angle: 77°; voxel size: 3.75 mm × 3.75 mm × 3.8 mm) of 17 axial slices positioned to provide coverage of the midbrain and striatum (**Figure 2**). A TR of 1 s, and consequently a smaller acquisition volume, was chosen to increase the sampling rate in our ROIs (NAcc and VTA). We note that the GE Signa EPI sequence automatically passes images through a Fermi filter with a transition width of 10 mm and radius of half the matrix size, which resulted in an effective smoothing kernel of approximately 4.8 mm^3. Thus, we did not include additional smoothing as part of our preprocessing protocol. Following reorientation, raw BOLD images were skull stripped using FSL's BET, corrected for intervolume head motion using MCFLIRT (Jenkinson et al., 2002), intensity normalized by a single multiplicative factor, and subjected to a high-pass temporal filter (Gaussian-weighted least-squares straight line fitting, with sigma = 50.0 s). Registration to high-resolution structural and standard-space images were

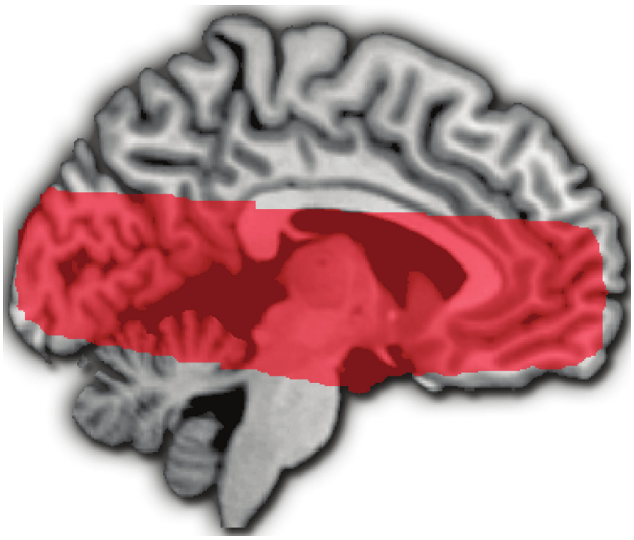


FIGURE 2 | Medial surface sagittal image showing overlap of fMRI volumes acquired in the 17 included participants.

carried out using FLIRT (Jenkinson and Smith, 2001; Jenkinson et al., 2002). All coordinates are reported in MNI space.

fMRI ANALYSIS: GENERAL LINEAR MODEL

All fMRI analyses were carried out using FEAT (FMRI Expert Analysis Tool) Version 5.92, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). Time-series statistical analyses used FILM with local autocorrelation correction (Woolrich et al., 2001).

Our first-level (i.e., within-run) analysis model included five regressors for the anticipation period with two regressors (gain and loss) for the outcome period of each trial type. The anticipation period was modeled as a unit-amplitude response with 1 s duration following the disappearance of the trial indicator cue. The outcome period was modeled as a unit-amplitude response with 1 s duration following the onset of feedback. Trial timing and numbers are noted in the task description above. Self $4 trials were contrasted against Self $0 trials (and Charity $4 against Charity $0) to examine anticipation of gain and loss. The Neutral Control $0 trials were modeled but not analyzed. Second-level (i.e., across-run, but within-subject) analyses used a fixed-effects model, while third-level (i.e., across-subjects) mixed-effects analyses (FLAME 1) included the main effects of each regressor from the lower level analysis, along with three covariates: reward sensitivity, empathy (other regarding preference), and inhibition. Whole-brain analyses used a voxel significance threshold of $z > 2.3$ and a cluster-significance threshold of $p < 0.05$, fully corrected for all voxels in our imaging volume (Worsley, 2001). Because clustering algorithms do not easily differentiate large areas of activation, **Tables 1–4** report the top ten peak voxels present using the elevated threshold indicated in each table.

fMRI ANALYSIS: REGIONS OF INTEREST

Our primary analyses used two anatomically defined ROIs: NAcc and VTA. Hand drawn anatomical ROIs were identified based on the average of all participant's normalized high-resolution anatomical images. The NAcc ROIs were drawn in each hemisphere according to (Breiter et al., 1997). The VTA ROI was drawn by isolating the region medial and anterior to the substantia nigra, following work of Adcock et al. (2006). Only ROI voxels that fell within the group coverage area were included in the analysis.

Table 1 | Self, Gain $4 > Gain $0.

	Peak (*x, y, z*)	*Z*-max
Right intraparietal sulcus	28, –80, 18	5.1
Dorsal striatum	–12, 14, 6	5.05
Ventral striatum	14, 22, –4	5.31
Visual cortex	6, –76, –4	4.95
Left intraparietal sulcus	–26, –84, 18	5.02
Dorsal-medial thalamus	–6, –18, 12	5
Visual cortex	14, –84, 4	5
Dorsal-medial thalamus	–6, –6, 8	4.91
Visual cortex	4, –64, 4	4.94
Visual cortex	16, –84, –4	4.91

Ten maximum cluster peaks, Z-cluster threshold 4.9. Coordinates are millimeters in MNI space.

Table 2 | Self, Lose $4 > Lose $0.

	Peak (*x, y, z*)	*Z*-max
Ventral striatum	10, 16, –2	5.29
Striatum	–10, 10, 6	5.26
Left insula/operculum	–36, 14, 8	5.26
Right insula/operculum	38, 14, 8	5.19
Dorsal-medial thalamus	14, –4, 16	5.14
Visual cortex	2, –76, –6	4.91
Left dorsal-medial thalamus	–10, –2, 12	4.88
Left insula/operculum	–44, 12, –2	4.85
Right dorsal-medial thalamus	16, 12, –4	4.86
Visual cortex	6, –84, –4	4.83

Ten maximum cluster peaks, Z-cluster threshold 4.826. Coordinates are millimeters in MNI space.

Table 3 | Charity, Gain $4 > Gain $0.

	Peak (*x, y, z*)	*Z*-max
Ventral striatum	12, 16, –2	4.43
Right insula/operculum	32, 20, 8	4.46
Thalamus	–10, –14, 8	4.32
Thalamus	12, –12, 6	4.37
Midbrain	10, –12, –6	4.24
Midbrain	–6, –22, –6	4.29
Midbrain	8, –18, –2	4.31
Ventral striatum	–16, 14, –2	4.32
Dorsal striatum	–8, 14, 0	4.13
Ventral striatum	–18, 20, –2	4.28

Ten maximum cluster peaks, Z-cluster threshold 4.1. Coordinates are millimeters in MNI space.

Table 4 | Charity, Lose $4 > Lose $0.

	Peak (*x, y, z*)	*Z*-max
Ventral striatum	22, 6, 2	4.35
Right insula/operculum	32, 16, 8	4.2
Ventral striatum	–18, 14, 2	4.11
Midbrain	–6, –18, –2	4.12
Thalamus	8, –10, 4	4.04
Visual cortex	0, –76, 4	4.04
Ventral striatum	–22, 0, 2	4.05
Thalamus	8, –22, 0	4.13
Ventral striatum	–18, 16, –4	4.08
Ventral striatum	22, 20, –6	4.2

Ten maximum cluster peaks, Z-cluster threshold 4.035. Coordinates are millimeters in MNI space.

RESULTS

BEHAVIORAL DATA

The proportion of successful responses (i.e., those faster than the adaptive response-time threshold) was similar across all four self and charity reward conditions (Self $4: *M* 64%, SD 7%; Charity $4: *M* 64%, SD 4%; Self $0: *M* 63%, SD 6%; Charity $0: *M* 62%, SD 5%), indicating that our adaptive algorithm successfully matched reward rates. Reaction times to $4 gain trials (*M* 208 ms, SD 23 ms) were not significantly different from reaction times to $4 loss trials (*M* 207 ms, SD 24 ms). Reaction times on $4 trials were faster than $0 trials, and $4 trials played for Self were faster than $4 trials played for Charity [2 (beneficiary: Self vs. Charity) × 2 (magnitude: $4 vs. $0) repeated-measures ANOVA; main effect of magnitude: $F(1, 16) = 25.7$, $p < 0.01$, $4; beneficiary × magnitude interaction: $F(1,16) = 7.5$, $p < 0.05$); paired comparison (*M* –22 ms, SEM 4 ms, $p < 0.001$) of $4 (*M* 207 ms, SD 23 ms) vs. $0 (*M* 210 ms, SD 24 ms), $p < 0.001$; paired comparison (*M* –5ms, SEM 2 ms, $p < 0.05$) of Self $4 (*M* 205 ms, SD 24 ms) vs. Charity $4 (*M* 210 ms, SD 23 ms)]. There were no significant differences in reaction times on $0 trials (Self $0: *M* 230 ms, SD 28 ms; Charity $0: *M* 228 ms, SD 23 ms).

fMRI RESULTS: WHOLE-VOLUME ANALYSES

Anticipating gain and loss for self

All analyses reported in this manuscript use regressors associated with reward anticipation (i.e., time-locked to the disappearance of the initial reward cue). We first contrasted parameter estimates between trials that offered the chance to make $4 and trials where no money was at stake (Self-Gain $4 > Self-Gain $0). Activation associated with anticipated monetary gains was widely distributed throughout the imaged volume (**Table 1**), with peaks in the dorsal striatum and vSTR, bilateral operculum/insula (**Figure 3A**, top), midbrain (**Figure 3A**, bottom), mediodorsal thalamus, medial prefrontal, medial orbitofrontal, anterior pole, and visual cortex. These results replicate those found in previous studies of gain anticipation (Knutson et al., 2001; Knutson and Greer, 2008).

Next, we conducted a similar analysis for anticipated monetary losses, by contrasting trials that offered the chance to avoid losing $4 and trials where no money was at stake (Self-Loss $4 vs. Self-Loss $0). Activations in this loss-anticipation contrast (**Table 2**) were distributed similarly to the gain condition. Peaks of activation were also similar to those noted under the gain condition, including in the dorsal striatum and vSTR, bilateral operculum/insula (**Figure 3B**, top), midbrain (shown in **Figure 3B**, bottom), mediodorsal thalamus, and orbitofrontal and visual cortex.

The direct contrast between gain and loss anticipation (Self-Gain $4 > Self-Loss $4) identified only one cluster along the inferior parietal sulcus ($Z = 3.2$; max: 32, –82, 20), and no differential activation overlapping our ROIs or in other regions implicated in reward anticipation by prior literature. No significant clusters of activation were identified in the reverse contrast (Self-Loss $4 > Self-Gain $4). Moreover, *no clusters* exhibited significantly *decreased* activation during either self-directed gain or loss trials compared to control trials (i.e., Self-Gain $0 > Self-Gain $4, or Self-Loss $0 > Self-Loss $4).

Anticipating gain and loss for charity

We repeated all of the analyses from the previous section for trials that offered the chance to gain or lose money for the selected charity. Anticipating potential gains and losses for a charity evoked activation in regions within the dorsal striatum and vSTR, midbrain, thalamus, prefrontal cortex, bilateral insula, and visual cortex. Note that there was very good match between the peak loci of activation for self-directed and charity-directed rewards (**Tables 3 and 4**). Direct contrasts of trials involving potential gains and potential losses (Charity-Gain $4 > Charity-Loss $4, or Charity-Loss $4 > Charity-Gain $4) revealed no clusters of activation that survived whole-volume correction.

Playing for self vs. playing for charity

We next identified regions that exhibited significant differences in activation depending on whether participants were anticipating playing for themselves or for their charity. The direct contrast of self-directed gains greater than charity-directed gains (Self-Gain $4 > Charity-Gain $4) identified activations similar to those found for self-gains (i.e., Self-Gain $4 > Self-Gain $0); i.e., within reward-related regions like the NAcc and VTA. Activation in these regions was greatest to self-directed rewards, intermediate to charity-directed rewards, and least on trials where no reward could be obtained. Additional regions whose activation increased to self-directed gains (**Table 5**) included the prefrontal cortex, temporal–parietal–occipital junction (TPO), and posterior insula/inferior parietal lobule (IPL). Likewise, the direct contrast of self-directed losses greater than charity-directed losses (Self-Loss $4 > Charity-Loss $4, **Table 6**) evoked activation in reward-related regions, along with additional clusters in the TPO and IPL.

The only region exhibiting greater charity-directed activation than self-directed activation was the posterior cingulate cortex (PCC). This activation survived whole-volume correction for the loss trials (Charity-Loss $4 > Self-Loss $4, **Table 7**), but not for the gain trials (Charity-Gain $4 > Loss-Gain $4; $z = 2.8$ for coordinates: 2, –56, –18).

fMRI RESULTS: ROI ANALYSIS

Anticipatory activations in the VTA and NAcc are similar for self and charity

We defined ROIs in the VTA and NAcc, collapsed across hemispheres (see Section "Materials and Methods" for details). For each

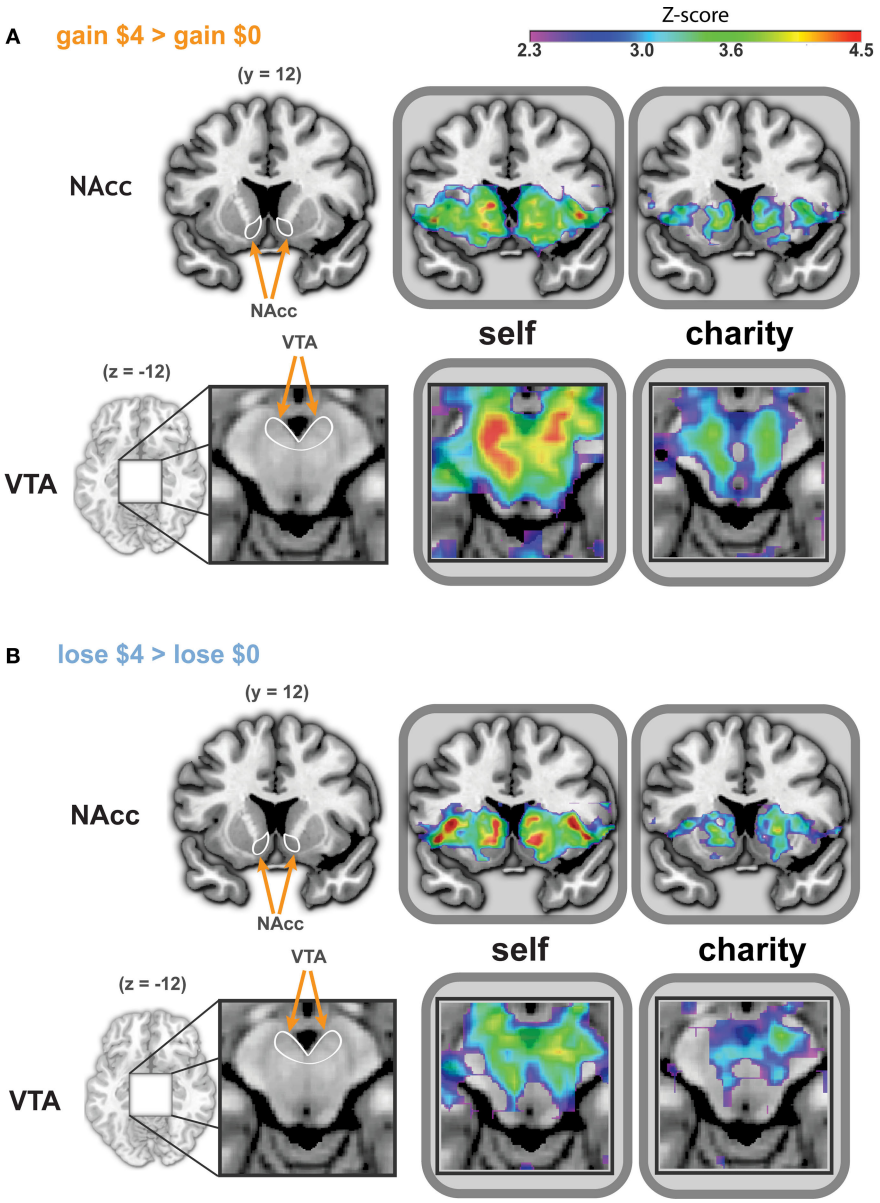


FIGURE 3 | Whole-brain analysis reveals similar patterns of activation during anticipation of gains and losses, whether participants played for self or a charity. Activated regions were larger and more significant in the Self conditions. Activation peaks were present in the NAcc and VTA in all four treatments (i.e., anticipating gain, anticipating loss, playing for self, playing for a charity). ROIs for bilateral NAcc (**A** and **B**, *top*) are shown on a coronal image ($y = -12$). The ROI for the VTA (**A** and **B**, *bottom*) is shown on a magnified axial image ($z = -12$). ROIs are indicated in white on an anatomical image to the left of the statistical maps. The left side of each image corresponds to the participant's left. All statistical map colors reflect the *Z*-score color scale in the upper right corner. Other significant peaks in each condition are listed in **Tables 1–4**.

subject, we calculated parameter estimates for each ROI and reward type within a two-factor (beneficiary: Self vs. Charity; valence: gain vs. loss) repeated measures ANOVA. Note that for each trial type, we subtracted the mean activation associated with the matched $0-reward trial (e.g., Self-Gain $4 minus Self-gain $0), to control for non-task-related processing (e.g., cue perception).

We found that both VTA and NAcc showed greater activation to self-directed rewards compared to charity-directed rewards [VTA: $F(1,13) = 7.41, p < 0.05$; NAcc: $F(1,13) = 12.31, p < 0.05$] though on average activations were positive in both the VTA [$F(1,13) = 70.14$, $p < 0.05$] and NAcc [$F(1,13) = 79.97, p < 0.05$]. Neither the VTA nor the NAcc ROIs showed significant main effect of valence though the VTA did exhibit a trend [Gain vs. Loss, $F(1,13) = 3.96$ $p = 0.07$]. However, the VTA did show a significant effect of valence that scaled with our Reward Sensitivity covariate [Gain vs. Loss × Reward Sensitivity, $F(1,13) = 5.74$, $p < 0.05$]. Although the NAcc did not show any main effects or direct interactions of valence, it did show a three-way interaction incorporating an effect of valence [Self vs. Charity × Gain vs. Loss × Reward Sensitivity, $F(1,13) = 5.81$, $p = 0.031$; **Figure 4**]. We also note that we found no significant

Table 5 | Self-Gain $4 > Charity-Gain $4.

	Peak (*x, y, z*)	*Z*-max
Left prefrontal cortex	–22, 56, –4	3.5
	–33, 46, 14	3.1
	–8, 36, 18	3.3
Right temporal–parietal–occipital junction	46, –60, 2	4.09
Left temporal–parietal–occipital junction	–50, –68, –6	3.8
Right inferior parietal lobule	46, –26, 12	3.6
Left inferior parietal lobule	–35, –28, 16	3.7

The peaks listed are only significantly active when playing for self.

Table 6 | Self-Lose $4 > Charity-Lose $4.

	Peak (*x, y, z*)	*Z*-max
Right temporal–parietal–occipital junction	44, –68, –8	3.8
Left temporal–parietal–occipital junction	–48, –68, –4	4.1
Right inferior parietal lobule	49, –47, 16	3.3
Left inferior parietal lobule	–62, –24, 24	4.0

The peaks listed are only significantly active when playing for self.

Table 7 | Charity-Lose $4 > Self-Lose $4.

	K (voxels)	Peak (*x, y, z*)	*Z*-max
Posterior cingulate cortex	272	–4, –50, 20	3.45

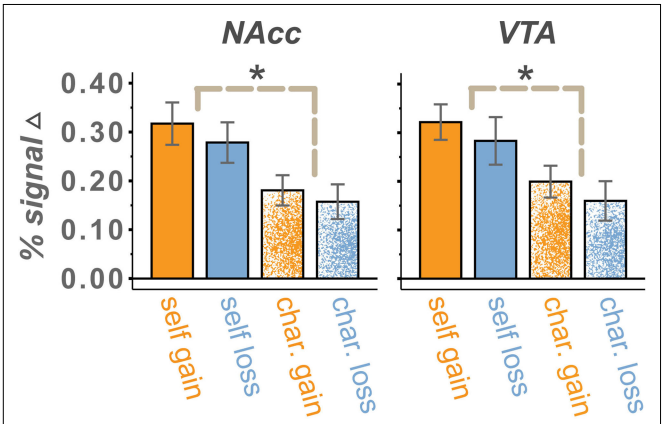


FIGURE 4 | Percent signal change in the NAcc and VTA for $4 vs. $0 trials. Mean activations, relative to $0 conditions, were positive for all trial types. Activations were larger in the Self than Charity treatment condition, reflecting reliable differences on both gain and loss runs. A trend for a main effect of valence was present in the VTA but not the NAcc. Valence effects that are modulated by the reward sensitivity of the participant were present in both regions. We found no significant differences between $0 conditions. Error bars are ±standard error of the mean.

differences in these regions between mean signal changes in the $0 conditions, indicating that these effects are contingent upon the presence of anticipated reward.

To assess the localization of each ROI and test for potential spatial inhomogeneity, we also restricted our analyses to the single voxel with the highest *Z*-score (i.e., most significant) to self-directed gains within the NAcc (MNI coordinates: 12, 6, –6) and VTA (MNI coordinates: 4, –16, –10) ROIs. Results of these ANOVAs are consistent with the results of the whole ROI ANOVAs in both the VTA and NAcc with two exceptions. In the NAcc, the three-way interaction present in the complete NAcc ROI (Gain vs. Loss × Reward Sensitivity) was non-significant for the peak voxel alone [$F(1,13) = 2.38$, $p = 0.15$]. Second, in the VTA, the peak reward-anticipation-sensitive voxel showed a significant main effect of valence [Gain vs. Loss: $F(1,13) = 8.30$, $p < 0.05$], an effect only significant at the trend level in the analysis of the complete VTA ROI. Note that this increase in significance may simply reflect a selection bias, given that this voxel was selected for its robust responses in the self-gain condition. As further confirmation of a motivational salience signal, significant increases in activation to self-directed losses, to charity-directed gains, and to charity-directed losses were also present in the whole-brain analysis from this voxel. In addition, there were no voxels in the VTA or NAcc that showed negative activity on loss trials (with respect to the $0 condition) across participants.

Recent work by Matsumoto and Hikosaka (2009) indicates there are two varieties of dopaminergic neurons in the VTA, one population that responds to positive conditions and one population that responds to both positive and negative conditions. With this in mind we also interrogated voxels in the VTA (MNI coordinates: –10, –16, –12) and NAcc (MNI coordinates: –12, 10, –6) that showed the peak activation increase (i.e., greatest *Z*-score) during anticipation on loss trials. The NAcc loss peak results were consistent with those of the complete ROI and gain peak in that they showed a positive average response in all conditions, a main effect of beneficiary, a trend toward an effect of participant reward sensitivity, and significant three-way interaction of Self vs. Charity × Gain vs. Loss × Reward Sensitivity. Consistent with Matsumoto and Hikosaka, the peak loss voxel in the VTA differed from the peak gain voxel in that it exhibited no significant main effect of valence [$F(1,13) = 0.713$, $p = 0.41$]. We caution that these analyses do not directly test spatial inhomogeneity effects and that such results may be attributable to selection bias because although our initial definition of ROIs was independent, definitions of the peak voxels was based on non-independent tests.

Individual gain and loss anticipation traits in the NAcc and VTA

In the current study, a main effect of affective valence would manifest in increased activation for anticipation of gains and decreased activation for anticipation of losses (or vice versa). Conversely, a main effect of motivation would lead to increased activation for anticipation of both gains and losses, compared to trials without the possibility of reward. As described above, our whole-volume analyses provided no suggestions of opposite responses for gains and losses within reward-related regions; to the contrary, we found that gains and losses each evoked significant increases in activation within the NAcc and VTA, among other regions. We repeated these analyses for our anatomically defined ROIs and found a similar result: increased activation for both gain and loss trials, with greater activation for self-directed compared to charity-directed trials. Thus, we found no evidence for group-level main effects of valence in our target regions.

We next investigated whether there were any across-subjects relationships between the magnitude of the responses to gain and loss trials. If there were a *negative* correlation across individuals between activations to gain and loss trials, even though the mean activation for both types of trials was positive, then that would be strong evidence that both motivation and affective valence modulate activation in reward-related regions. Alternatively, a *positive* correlation between activations to gain and loss trials would provide evidence in favor of a motivational explanation, alone. Our results support the motivation explanation. In the NAcc, activations during gain anticipation scaled positively with loss anticipation (**Figure 5**), with a significant correlation in self-directed-trials ($r = 0.64$) and a non-significant but numerically positive correlation on charity-directed trials. In the VTA (**Figure 6**), activations during gain and loss anticipation were positively correlated for both self-directed ($r = 0.58$) and charity-directed ($r = 0.63$) trials.

We next used a hierarchical regression analysis to evaluate whether the neural bias toward gains, compared to losses, was predicted by our reward sensitivity covariate. We found that there were strong correlations between reward sensitivity and the differential activation between gains and losses (e.g., Self-Gain \$4 minus Self-Loss \$4) in both the NAcc and VTA (**Figures 5A and 6A**). Individuals who had the greatest reward sensitivity exhibited the greatest relative increment in activation gains compared to losses. (We note that this is a fully independent correlation, in that we are using an independent behavioral test, an anatomical ROI, and the residual activation following a contrast of conditions.) This effect was significant for self-directed trials in both NAcc and VTA, but not for charity-directed trials in either ROI. We conducted similar analyses using covariates for other-regarding preferences and behavioral inhibition, and found no significant effects. Based on these results, we conducted a *post hoc* test looking at the relationship between our reward sensitivity covariate and activation to each trial type (as opposed to the difference between trial types described above). We found that, within our sample, the NAcc and VTA responses to self-directed gains were largely similar regardless of reward sensitivity, but that high reward-sensitivity scores correlated with a relative decrease in activation on the other trial

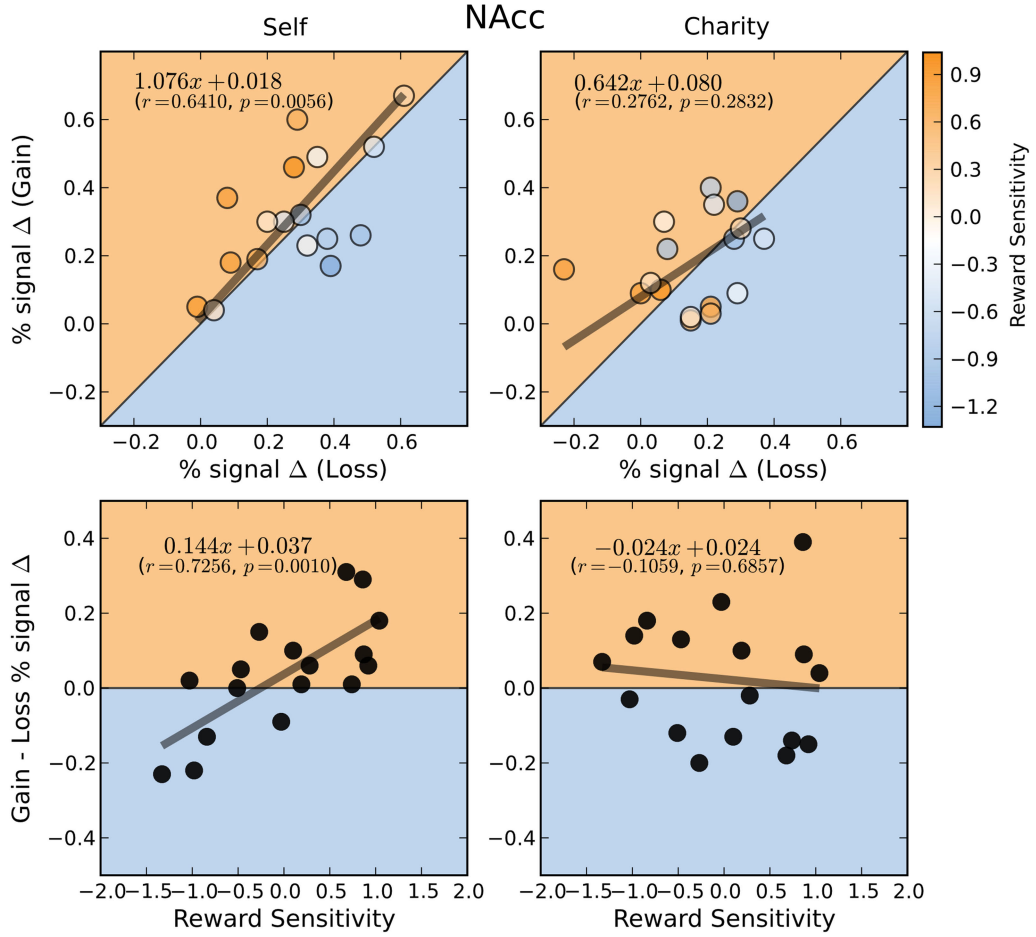


FIGURE 5 | BOLD responses in the NAcc during gain and loss anticipation are positively correlated when participants play for themselves. *Top*: Average percent signal change differences (paid-control) for anticipation of gain and loss trials for the Self (left) and Charity (right) treatments. Each point is colored according to the participant's relative reward sensitivity index (z transformed). *Bottom*: Individual Gain vs. Loss signal change differences are plotted against the participant's reward sensitivity index (z transformed). Each plot includes the orthogonal distance regression best fit line, as well as the correlation coefficient (r) and the p-value of that correlation (p). Only the regression in the Self condition was significant.

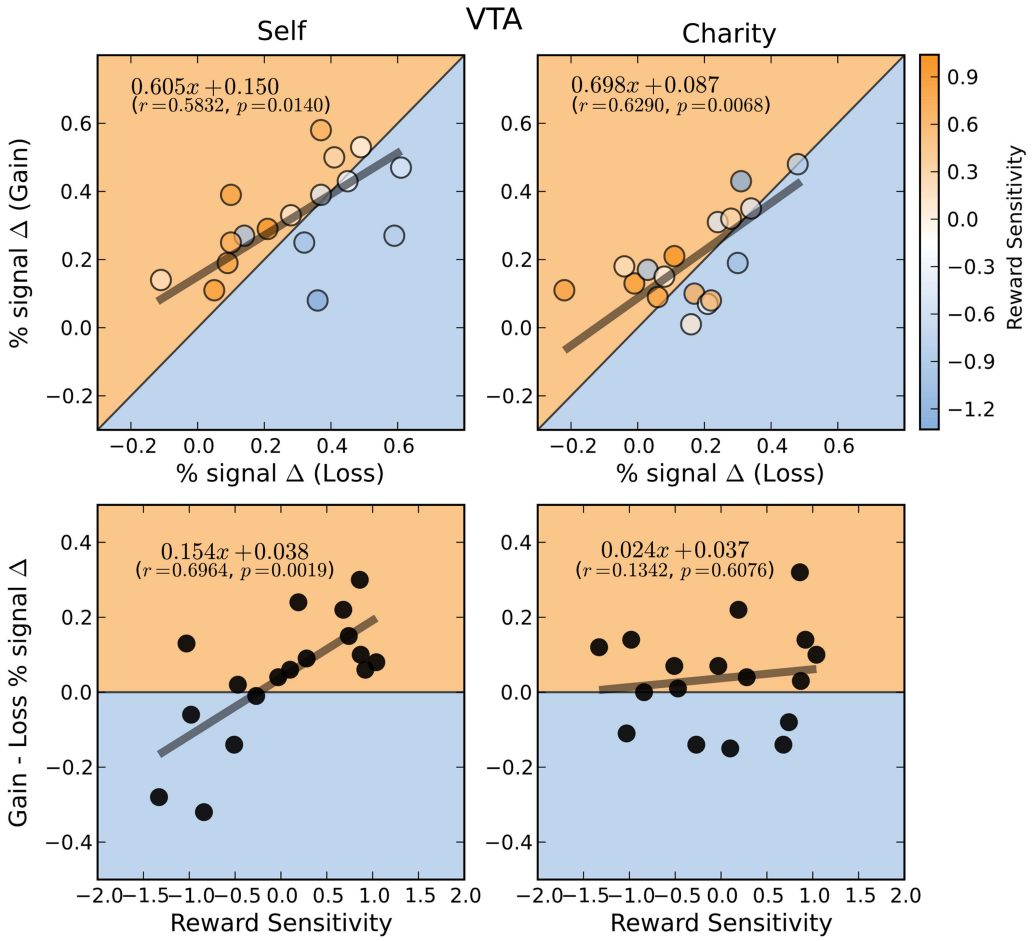


FIGURE 6 | BOLD responses in the VTA during the anticipation of gain and loss are positively correlated whether participants play for themselves or for a charity. *Top*: Average percent signal change differences (paid-control) for anticipation of gain and loss trials for the Self (*left*) and Charity (*right*) treatments. Each point is colored according to the participant's relative reward sensitivity index (z transformed). Gain and loss responses differences are positively correlated in both the self and charity conditions. *Bottom*: Individual Gain vs. Loss signal change differences are plotted against the participant's reward sensitivity index (z transformed). Each plot includes the orthogonal distance regression best fit line, as well as the correlation coefficient (*r*) and the *p*-value of that correlation (*p*). Only the regression in the Self condition was significant.

types (**Figure 7**; see also colored circles on the upper right panels of **Figures 5 and 6**).

DISCUSSION

We examined brain activation during the anticipation of monetary rewards that varied in their valence (i.e., gain vs. loss) and beneficiary (i.e., self-directed vs. charity-directed). We found that activation in putatively reward-related regions, specifically the NAcc and VTA, increased during both gain- and loss-anticipation, with greater responses to self-directed than charity-directed trials. Moreover, there was a strong positive correlation between these responses across individuals, such that those individuals with the greatest anticipatory response to potential gains also had the greatest response to potential losses. Together, these results indicate that anticipatory activation reflects the motivational properties of the potential reward, not its valence. However, we found evidence, using an independent behavioral covariate, that individual differences in reward sensitivity modulated the relative response to gains and losses, with more reward-sensitive individuals exhibiting relatively more activation to gains compared to losses. Below, we consider the implications of each of these results.

REWARD ANTICIPATION: MOTIVATION VS. VALENCE

In group analyses, we found no evidence that anticipatory activations in either VTA or NAcc reflect a univariate value signal that scales according to both the valence and magnitude of the potential reward (i.e., gain > neutral > loss). Both potential gains and potential losses evoked increased activation compared to control stimuli in the NAcc and VTA, as shown within a whole-volume analysis, an anatomical ROI analysis, and in an analysis restricted to the most-active voxel in each region. And, as even stronger evidence that anticipatory activations reflect motivational salience, we found that activations associated with gains and with losses were positively correlated across participants. These results lie in contrast to some previous studies that have shown increased NAcc activation to anticipated gains, compared to anticipated losses(Knutson et al., 2001), or have

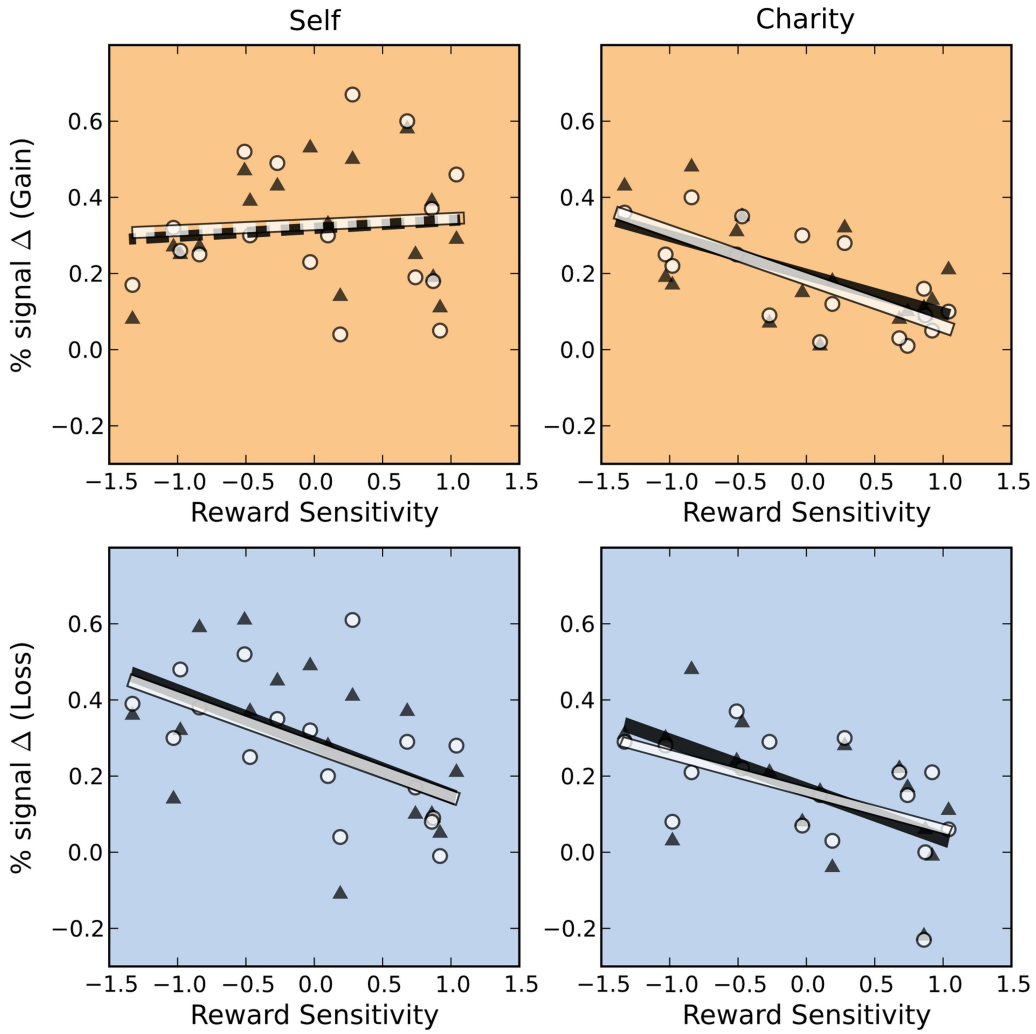


FIGURE 7 | BOLD responses in the Self-Gain condition are unrelated to reward sensitivity. BOLD responses ($4 > $0) in both the NAcc (white circle) and VTA (black triangles) were anticorrelated with reward sensitivity across subjects except in the Self-Gain condition (NAcc self-gain: $r = 0.06$, $p = 0.81$; VTA self-gain: $r = 0.11$, $p = 0.68$; NAcc charity-gain: $r = -0.72$, $p = 0.001$; VTA charity-gain: $r = -0.58$, $p = 0.01$; NAcc self-loss: $r = -0.57$, $p = 0.02$; VTA self-loss: $r = -0.50$, $p = 0.04$; NAcc charity-loss: $r = -0.51$, $p = 0.04$; VTA charity-loss: $r = -0.58$, $p = 0.02$). Solid lines indicate significant regressions. Dashed lines indicate regressions that were not significant.

failed to find increased activation to anticipated losses compared to a neutral control condition (Knutson et al., 2003, 2008a). For example, a study of loss aversion by Tom et al. (2007) showed that activation in the vSTR to decisions about mixed gambles (i.e., that involved a potential gain and a potential loss) increased with increased size of gain, but decreased with increasing size of loss. This result was interpreted as reflecting the response of a single reward mechanism that codes for both gains and losses along a single axis of reward value. We note that gains and losses were always paired in the design of Tom et al. (2007), such that the magnitude of the loss attenuated the overall value (i.e., magnitude) of the gamble. Within our study, in contrast, the potential losses were presented in isolation and thus reflected an independent and negative potential outcome, allowing a differentiation between magnitude and valence.

Prior research has suggested that under certain conditions NAcc activation reflects task factors other than value of a potential reward. Activation in the NAcc has been reported to correlate with both the salience of the stimulus presented (Zink et al., 2003) as well as the unpredictability of the potential outcome (Berns et al., 2001). It could be argued that salience or risk are inherently rewarding. However, there is also evidence that NAcc responses positively correlate with aversive stimuli (Delgado et al., 2004; Jensen et al., 2007; Salamone et al., 2007; Levita et al., 2009), as well.. In the current study, although activations to gain and loss anticipation both exceeded those present during all control conditions ($0), we found evidence that reward valence modulates the amplitude of this activation in the NAcc and VTA. In the VTA we found valence modulations related to reward sensitivity. In the NAcc modulation of valence was dependent on both the beneficiary of the reward and the reward sensitivity of the participant. One reasonable possibility is that the VTA and NAcc are primarily sensitive to aspects of motivational salience but that those responses

are modulated by affective valence (Cooper and Knutson, 2008), especially in those participants who are most reward sensitive. The relative strength of affective valence modulation would then likely be dependent upon task context. This mixed signal may also reflect spatial inhomogeneity within the VTA and NAcc, as discussed below.

A striking result came from the imperfect matching between the neurometric responses to potential gains and to potential losses. While the gain:loss ratio across the entire subject sample was approximately 1:1, some subjects showed a relatively increased response to gains, while others showed a relatively increased response to losses. This residual variation turned out to be systematically related to participants' reward-sensitivity scores. This behavior–brain correlation could reflect a contribution of some subcomponent of these reward-related regions, or the influence of another region that itself was sensitive to affective valence. An important direction for future research will be identifying the pattern of functional connectivity across regions that predict both trial-to-trial effects of cue value and across-subjects factors that bias those value signals.

THE ROLE OF THE VTA IN REWARD ANTICIPATION

Most prior neuroimaging studies of reward processing have focused on the vSTR, specifically the NAcc, which has been reliably reported to exhibit increased activation during anticipation. Much less evidence exists for the modulatory effects of anticipation in the VTA, the primary dopaminergic input to the NAcc (Swanson, 1982; Ikemoto, 2007). Prior research on VTA function, mostly using single-unit recording in non-human primates, has implicated that region in the processing of rewards, generally, and in transient responses to changing reward expectations (Ljungberg et al., 1992). Based on data showing that VTA neurons respond to both unexpected primary rewards and cues that predict future rewards, it has been theorized that these neurons code a reward prediction error, critical for TD learning (Schultz et al., 1997). It would be difficult to account for our results using prediction error signals that treat gains and losses as a single continuum. Because we separated our gain and loss cues into separate blocks, and used two types of rewards, a single continuum prediction error model would predict that we should observe the greatest anticipatory responses to Self-Gain cues, smallest (or most negative) responses to Self-Loss cues, responses in the same directions, but possibly attenuated, to both types of Charity cues, and minimal responses to the non-rewarded control cues. In contrast, we found very similar activation, both in spatial pattern and amplitude, for anticipated gains and anticipated losses, with both gains and losses greater than control cues or charitable cues. Alternatively, the opportunity to avoid losses may be might be seen as rewarding. In fact, there is evidence that relief from pain (Seymour et al., 2005) and even avoiding potential negative outcomes (Kim et al., 2006) can be viewed as rewards. However, this kind of "pure valence" explanation is inconsistent with the observation that activations on loss trials were still greater than neutral ($0) trials.

We emphasize that these results are not necessarily incompatible with the numerous prior demonstrations that prediction errors modulate the responses of VTA neurons, for three reasons. First, as is proposed by Seymour et al. (2007), there may be multiple and potentially valence-dependent prediction error signals. That is, separate neuronal prediction-error signals may increase in anticipation of gains and of losses, each contributing to observed BOLD activation. Second, monetary losses may not have similar psychological and neural effects as omitted primary rewards or aversive stimuli. In particular, the loss of money reflects an opportunity cost that affects the total value of a future reward, rather than an immediate negative consequence (e.g., a painful shock). Accordingly, humans frequently reframe decision problems to minimize decision difficulty or to maximize perceived value (Tversky and Kahneman, 1974); in our paradigm, like many others, the loss cue may have been reframed as an opportunity to avoid negative consequences. Third, activation measured using fMRI does not necessarily map onto the firing rate of individual neurons. Substantial methodological work suggests that the amplitude of BOLD activation matches best to local field potential and multi-unit activity within a region (Goense and Logothetis, 2008), and less well to single-unit activity. The relatively coarse timescale of fMRI data collection, combined with the filtering effects of the BOLD hemodynamic response, precludes determination of the relative timing of the contributing neuronal activity. In addition, evidence from Ungless et al. (2004) indicates the VTA may not be homogenous in its responsiveness to gain and loss. They find evidence of two distinct populations of neurons in the VTA, one responsive to positive stimuli and the other to aversive stimuli. Inhomogeneity in the VTA within dopaminergic neurons is supported in recent work by Matsumoto and Hikosaka (2009) who show not only distinct populations of neurons responsive to positively valenced stimuli but also provide evidence of a dorsal/ventral spatial distinction. Preliminary findings in the current study suggest that an fMRI study designed to look for spatial separation of gain-specific neuronal populations in the VTA may be able to isolate them from those responsive to both gains and losses. Given these caveats, our results should be interpreted as showing that some aspect of information processing in VTA is driven by motivational properties of anticipated rewards or by a prediction error that increases with the magnitude of anticipated punishment. We also note that individuals who are more reward sensitive display effects of valence not present in those relatively less sensitive to reward.

MODULATION OF ANTICIPATORY REWARD SIGNALS BY SELF- VS. OTHER-DIRECTED CONTEXT

The NAcc not only responds to meaningful self-directed outcomes, it also responds to a variety of other-directed outcomes: social cooperation (Rilling et al., 2002), altruistic punishment (Singer et al., 2006), and rewards for a favored charity (Moll et al., 2006; Harbaugh et al., 2007), among others. In these latter cases, the reward may be emulated as if it were being personally received and is therefore represented within the same system, albeit with reduced magnitude. We note that prior research showing activation in the reward system to charitable rewards used tasks involving active decisions or passive receipt of those rewards. Here, we show that mere anticipation of potential reward is sufficient to evoke activation within the NAcc; moreover, like the work of Moll et al., we extend our conclusions to include VTA, as well.

Notably, all of our main-effect analyses indicated that self-directed and charity-directed rewards evoked very similar patterns of activation: for both types of rewards, activation in the NAcc and VTA increased for both anticipated gains and anticipated losses. What differentiated these two reward types was our participants' relative reward sensitivity, such that individuals with higher reward-sensitive individuals showed lower responses for all charitable rewards. Somewhat surprisingly, we found no similar across-participant effect of our other-regarding-preference covariate. We note that prior studies have shown that the relative subjective value of different charitable rewards, defined by the participant's willingness to engage in a transaction as opposed to individual differences in overall other-regarding-preferences, modulates activation of the vSTR (Moll et al., 2006; Harbaugh et al., 2007). In contrast, self-reported trait measures of other-regarding preferences have been reported to relate to structural (Yamasue et al., 2008) and functional (Tankersley et al., 2007) differences in other brain regions associated with social cognition. The independence of other-regarding preferences and likelihood of engaging in a charitable transaction is worthy of further investigation.

We have presented evidence that motivational salience modulates activation in the VTA and NAcc. Activations during the anticipation phase of all trial types were positive with respect to a $0 trial. However, the magnitude of this positive activation was modulated by three factors. First, the beneficiary: activations were smaller in magnitude when the outcome of the trial was not directed toward the participant, suggesting that a single system processes social and personal rewards according to their motivational salience. Second, the valence: in the VTA, the anticipation of gains evokes greater activation than the anticipation of losses, even though both conditions are greater than trials where no reward or punishment could be obtained. Third, the reward sensitivity of the individual: for participants who are more reward sensitive, the magnitude of activations to anticipation in the VTA and NAcc is largest on gain trials played for themselves. We conclude that both the VTA and NAcc provide anticipatory signals that largely reflect the motivational significance of potential rewards.

ACKNOWLEDGEMENTS

Julia Parker Goyer and Elizabeth B. Johnson provided assistance in task programming and data collection. We thank John Clithero, O'Dhaniel Mullette-Gillman, David Smith, and Vinod Venkatraman for manuscript comments. This project was supported by NIMH 70685 (Scott A. Huettel), NINDS 41328 (Scott A. Huettel), and NARSAD (R. Alison Adcock). R. McKell Carter is supported by NIH Fellowship (NIH 51156). Scott A. Huettel is supported by an Incubator Award from the Duke Institute for Brain Sciences. R. Alison Adcock is supported by the Alfred P. Sloan Foundation.

REFERENCES

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., and Gabrieli, J. D. E. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron* 50, 507–517.

Balleine, B. W., Delgado, M. R., and Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* 27, 8161–8165.

Bayer, H. M., and Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Beck, A. T., Erbaugh, J., Ward, C. H., Mock, J., and Mendelsohn, M. (1961). An inventory for measuring depression. *Arch. Gen. Psychiatry* 4, 561–571.

Berns, G. S., McClure, S. M., Pagnoni, G., and Montague, P. R. (2001). Predictability modulates human brain response to reward. *J. Neurosci.* 21, 2793–2798.

Berridge, K. C. (2004). Motivation concepts in behavioral neuroscience. *Physiol. Behav.* 81, 179–209.

Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology* 191, 391–431.

Berridge, K. C., Robinson, T. E., and Aldridge, J. W. (2009). Dissecting components of reward: 'liking', 'wanting', and learning. *Curr. Opin. Pharmacol.* 9, 65–73.

Bjork, J. M., Knutson, B., and Hommer, D. W. (2008). Incentive-elicited striatal activation in adolescent children of alcoholics. *Addiction* 103, 1308–1319.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., and Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron* 30, 619–639.

Breiter, H. C., Gollub, R. L., Weisskoff, R. M., Kennedy, D. N., Makris, N., Berke, J. D., Goodman, J. M., Kantor, H. L., Gastfriend, D. R., Riorden, J. P., Mathew, R. T., Rosen, B. R., and Hyman, S. E. (1997). Acute effects of cocaine on human brain activity and emotion. *Neuron* 19, 591–611.

Carver, C. S., and White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS Scales. *J. Pers. Soc. Psychol.* 67, 319–333.

Cooper, J. C., and Knutson, B. (2008). Valence and salience contribute to nucleus accumbens activation. *Neuroimage* 39, 538–547.

D'Ardenne, K., McClure, S. M., Nystrom, L. E., and Cohen, J. D. (2008). BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* 319, 1264–1267.

Davis, M. H. (1983). Measuring individual-differences in empathy – evidence for a multidimensional approach. *J. Pers. Soc. Psychol.* 44, 113–126.

Daw, N. D. (2007). Dopamine: at the intersection of reward and action. *Nat. Neurosci.* 10, 1505–1507.

Delgado, M. R., Gillis, M. M., and Phelps, E. A. (2008a). Regulating the expectation of reward via cognitive strategies. *Nat. Neurosci.* 11, 880–881.

Delgado, M. R., Li, J., Schiller, D., and Phelps, E. A. (2008b). The role of the striatum in aversive learning and aversive prediction errors. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 363, 3787–3800.

Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D. C., and Fiez, J. A. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *J. Neurophysiol.* 84, 3072–3077.

Delgado, M. R., Stenger, V. A., and Fiez, J. A. (2004). Motivation-dependent responses in the human caudate nucleus. *Cereb. Cortex* 14, 1022–1030.

Dillon, D. G., Holmes, A. J., Jahn, A. L., Bogdan, R., Wald, L. L., and Pizzagalli, D. A. (2008). Dissociation of neural regions associated with anticipatory versus consummatory phases of incentive processing. *Psychophysiology* 45, 36–49.

Elliot, A. J. (2006). The hierarchical model of approach-avoidance motivation. *Motiv. Emot.* 30, 111–116.

Ernst, M., Nelson, E. E., Jazbec, S., McClure, E. B., Monk, C. S., Leibenluft, E., Blair, J., and Pine, D. S. (2005). Amygdala and nucleus accumbens in responses to receipt and omission of gains in adults and adolescents. *Neuroimage* 25, 1279–1291.

Gard, D. E., Gard, M. G., Kring, A. M., and John, O. P. (2006). Anticipatory and consummatory components of the experience of pleasure: a scale development study. *J. Res. Pers.* 40, 1086–1102.

Goense, J. B. M., and Logothetis, N. K. (2008). Neurophysiology of the BOLD fMRI signal in awake monkeys. *Curr. Biol.* 18, 631–640.

Grusser, S. M., Wrase, J., Klein, S., Hermann, D., Smolka, M. N., Ruf, M., Weber-Fahr, W., Flor, H., Mann, K., Braus, D. F., and Heinz, A. (2004). Cue-induced activation of the striatum and medial prefrontal cortex is associated with subsequent relapse in abstinent alcoholics. *Psychopharmacology* 175, 296–302.

Harbaugh, W. T., Mayr, U., and Burghart, D. R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316, 1622–1625.

Hare, T. A., Camerer, C. F., and Rangel, A. (2009). Self-control in decision-

making involves modulation of the vmPFC valuation system. *Science* 324, 646–648.

Hare, T. A., O'Doherty, J., Camerer, C. F., Schultz, W., and Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* 28, 5623–5630.

Ikemoto, S. (2007). Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. *Brain Res. Rev.* 56, 27–78.

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841.

Jenkinson, M., and Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156.

Jensen, J., Smith, A. J., Willeit, M., Crawley, A. P., Mikulis, D. J., Vitcu, I., and Kapur, S. (2007). Separate brain regions code for salience vs. valence during reward prediction in humans. *Hum. Brain Mapp.* 28, 294–302.

Kalivas, P. W., Volkow, N., and Seamans, J. (2005). Unmanageable motivation in addiction: a pathology in prefrontal-accumbens glutamate transmission. *Neuron* 45, 647–650.

Kienast, T., and Heinz, A. (2006). Dopamine and the diseased brain. *CNS Neurol. Disord. Drug Targets* 5, 109–131.

Kilts, C. D., Schweitzer, J. B., Quinn, C. K., Gross, R. E., Faber, T. L., Muhammad, F., Ely, T. D., Hoffman, J. M., and Drexler, K. P. G. (2001). Neural activity related to drug craving in cocaine addiction. *Arch. Gen. Psychiatry* 58, 334–341.

Kim, H., Shimojo, S., and O'Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *Plos Biol* 4, e233. doi: 10.1371/journal.pbio.0040233

Knutson, B., Adams, C. M., Fong, G. W., and Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21, RC159.

Knutson, B., Bhanji, J. P., Cooney, R. E., Atlas, L. Y., and Gotlib, I. H. (2008a). Neural responses to monetary incentives in major depression. *Biol. Psychiatry* 63, 686–692.

Knutson, B., Wimmer, G. E., Rick, S., Hollon, N. G., Prelec, D., and Loewenstein, G. (2008b). Neural antecedents of the endowment effect. *Neuron* 58, 814–822.

Knutson, B., Fong, G. W., Bennett, S. M., Adams, C. M., and Hommer, D. (2003). A region of mesial prefrontal cortex tracks monetarily rewarding outcomes: characterization with rapid event-related fMRI. *Neuroimage* 18, 263–272.

Knutson, B., and Gibbs, S. E. B. (2007). Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology* 191, 813–822.

Knutson, B., and Greer, S. M. (2008). Anticipatory affect: neural correlates and consequences for choice. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 363, 3771–3786.

Knutson, B., Rick, S., Wirnmer, G. E., Prelec, D., and Loewenstein, G. (2007). Neural predictors of purchases. *Neuron* 53, 147–156.

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., and Glover, G. (2005). Distributed neural representation of expected value. *J. Neurosci.* 25, 4806–4812.

Knutson, B., Westdorp, A., Kaiser, E., and Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *Neuroimage* 12, 20–27.

Levita, L., Hare, T. A., Voss, H. U., Glover, G., Ballon, D. J., and Casey, B. J. (2009). The bivalent side of the nucleus accumbens. *Neuroimage* 44, 1178–1187.

Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163.

Matsumoto, M., and Hikosaka, O. (2009). Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* 459, 837–841.

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A. J., Schweizer, S., Frith, C. D., and Dalgleish, T. (2009). A key role for similarity in vicarious reward. *Science* 324, 900.

Moll, J., Krueger, F., Zahn, R., Pardini, M., de Oliveira-Souzat, R., and Grafman, J. (2006). Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl. Acad. Sci. USA* 103, 15623–15628.

Montague, P. R., and Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron* 36, 265–284.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454.

Pecina, S., and Berridge, K. C. (2005). Hedonic hot spot in nucleus accumbens shell: where do mu-opioids cause increased hedonic impact of sweetness? *J. Neurosci.* 25, 11777–11786.

Phillips, P. E., Walton, M. E., and Jhou, T. C. (2007). Calculating utility: preclinical evidence for cost-benefit analysis by mesolimbic dopamine. *Psychopharmacology (Berl.)* 191, 483–495.

Pizzagalli, D. A., Holmes, A. J., Dillon, D. G., Goetz, E. L., Birk, J. L., Bogdan, R., Dougherty, D. D., Iosifescu, D. V., Rauch, S. L., and Fava, M. (2009). Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder. *Am. J. Psychiatry* 166, 702–710.

Pulos, S., Elison, J., and Lennon, R. (2004). The hierarchical structure of the Interpersonal Reactivity Index. *Soc. Behav. Pers.* 32, 355–359.

Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556.

Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., and Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.

Rushton, J. P., Chrisjohn, R. D., and Fekken, G. C. (1981). The altruistic personality and the Self-Report Altruism Scale. *Pers. Individ. Dif.* 2, 293–302.

Salamone, J. D., Correa, M., Farrar, A., and Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology* 191, 461–482.

Schlagenhauf, F., Juckel, G., Koslowski, M., Kahnt, T., Knutson, B., Dembler, T., Kienast, T., Gallinat, J., Wrase, J., and Heinz, A. (2008). Reward system activation in schizophrenic patients switched from typical neuroleptics to olanzapine. *Psychopharmacology* 196, 673–684.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Scott, D. J., Stohler, C. S., Egnatuk, C. M., Wang, H., Koeppe, R. A., and Zubieta, J. K. (2008). Placebo and nocebo effects are defined by opposite opioid and dopaminergic responses. *Arch. Gen. Psychiatry* 65, 220–231.

Seymour, B., Daw, N., Dayan, P., Singer, T., and Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *J. Neurosci.* 27, 4826–4831.

Seymour, B., O'Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., Friston, K. J., and Frackowiak, R. S. (2004). Temporal difference models describe higher-order learning in humans. *Nature* 429, 664–667.

Seymour, B., O'Doherty, J. P., Koltzenburg, M., Wiech, K., Frackowiak, R., Friston, K., and Dolan, R. (2005). Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. *Nat. Neurosci.* 8, 1234–1240.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., and Frith, C. D. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439, 466–469.

Strohle, A., Stoy, M., Wrase, J., Schwarzer, S., Schlagenhauf, F., Huss, M., Hein, J., Nedderhut, A., Neumann, B., Gregor, A., Juckel, G., Knutson, B., Lehmkuhl, U., Bauer, M., and Heinz, A. (2008). Reward anticipation and outcomes in adult males with attention-deficit/hyperactivity disorder. *Neuroimage* 39, 966–972.

Sutton, R. S., and Barto, A. G. (1990). Foundations of adaptive networks. In Learning and Computational Neuroscience, M. Gabriel and J. Moore, eds (Cambridge, MA, MIT Press), pp. 497–537.

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

Swanson, L. W. (1982). The projections of the ventral tegmental area and adjacent regions – a combined fluorescent retrograde tracer and immunofluorescence study in the rat. *Brain Res. Bull.* 9, 321–353.

Tankersley, D., Stowe, C. J., and Huettel, S. A. (2007). Altruism is associated with an increased neural response to agency. *Nat. Neurosci.* 10, 150–151.

Tom, S. M., Fox, C. R., Trepel, C., and Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science* 315, 515–518.

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty – heuristics and biases. *Science* 185, 1124–1131.

Ungless, M. A., Magill, P. J., and Bolam, J. P. (2004). Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* 303, 2040–2042.

Valentin, V. V., Dickinson, A., and O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed

learning in the human brain. *J. Neurosci.* 27, 4019–4026.

Wise, R. (1980). The dopamine synapse and the notion of pleasure centers in the brain. *Trends Neurosci.* 3, 91–95.

Wise, R. A. (2004). Dopamine, learning and motivation. *Nat. Rev. Neurosci.* 5, 483–494.

Woolrich, M. W., Ripley, B. D., Brady, M., and Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage* 14, 1370–1386.

Worsley, K. J. (2001). Statistical analysis of activation images. In Functional MRI: An Introduction to Methods, P. Jezzard, P. M. Matthews, S. M. Smith, eds (Oxford, Oxford University Press), pp. 251–270.

Yamasue, H., Abe, O., Suga, M., Yamada, H., Rogers, M. A., Aoki, S., Kato, N., and Kasai, K. (2008). Sex-linked neuroanatomical basis of human altruistic cooperativeness. *Cereb. Cortex* 18, 2331–2340.

Zink, C. F., Pagnoni, G., Martin, M. E., Dhamala, M., and Berns, G. S. (2003). Human striatal response to salient nonrewarding stimuli. *J. Neurosci.* 23, 8092–8097.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 May 2009; paper pending published: 02 June 2009; accepted: 11 August 2009; published online: 27 August 2009.

*Citation: Carter RM, MacInnes JJ, Huettel SA and Adcock RA (2009) Activation in the VTA and nucleus accumbens increases in anticipation of both gains and losses. Front. Behav. Neurosci. **3**:21. doi: 10.3389/neuro.08.021.2009*

frontiers in
BEHAVIORAL NEUROSCIENCE

REVIEW ARTICLE
published: 07 September 2009
doi: 10.3389/neuro.08.013.2009

Dopamine, behavioral economics, and effort

John D. Salamone*, Merce Correa†, Andrew M. Farrar, Eric J. Nunes and Marta Pardo†

Department of Psychology, University of Connecticut, Storrs, CT, USA

Edited by:
Paul Phillips, University of Washington, USA

Reviewed by:
Daeyeol Lee, Yale University, USA
Kent Berridge, University of Michigan, USA

****Correspondence:***
John D. Salamone, Board of Trustees Distinguished Professor, Division of Behavioral Neuroscience, Department of Psychology, University of Connecticut, Storrs, CT 06269-1020, USA.
e-mail: john.salamone@uconn.edu

†Present address: Area de Psicobiol., Dept. Psic., Universitat de Jaume I, Castelló, 12071, Spain.

There are numerous problems with the hypothesis that brain dopamine (DA) systems, particularly in the nucleus accumbens, directly mediate the rewarding or primary motivational characteristics of natural stimuli such as food. Research and theory related to the functions of mesolimbic DA are undergoing a substantial conceptual restructuring, with the traditional emphasis on hedonia and primary reward yielding to other concepts and lines of inquiry. The present review is focused upon the involvement of nucleus accumbens DA in behavioral activation and effort-related processes. Viewed from the framework of behavioral economics, the effects of accumbens DA depletions and antagonism on food-reinforced behavior are highly dependent upon the work requirements of the instrumental task, and DA depleted rats are more sensitive to increases in response costs (i.e., ratio requirements). Moreover, interference with accumbens DA transmission exerts a powerful influence over effort-related choice behavior. Rats with accumbens DA depletions or antagonism reallocate their instrumental behavior away from food-reinforced tasks that have high response requirements, and instead these rats select a less-effortful type of food-seeking behavior. Nucleus accumbens DA and adenosine interact in the regulation of effort-related functions, and other brain structures (anterior cingulate cortex, amygdala, ventral pallidum) also are involved. Studies of the brain systems regulating effort-based processes may have implications for understanding drug abuse, as well as energy-related disorders such as psychomotor slowing, fatigue or anergia in depression and other neurological disorders.

Keywords: nucleus accumbens, motivation, reward, reinforcement, activation, anergia, psychomotor slowing, depression

LIMITATIONS OF THE REWARD HYPOTHESIS OF DOPAMINERGIC FUNCTION

The last several years have seen substantial theoretical developments related to the hypothesized behavioral functions of nucleus accumbens dopamine (DA). It has become evident to many investigators that there are conceptual limitations and empirical problems with the traditional DA hypothesis of "reward" (Baldo and Kelley, 2007; Barbano and Cador, 2007; Salamone et al., 1997, 2005, 2007; Salamone, in press). Even the use of the term "reward" itself often is problematic (Cannon and Bseikri, 2004; Salamone et al., 2005; Salamone, 2006; Sanchis-Segura and Spanagel, 2006; Yin et al., 2008). Researchers rarely define what they mean by "reward" when they are using it to describe a psychological process; some use it as though it were a synonym for "reinforcement", or in reference to "appetite" or "primary motivation", while still others employ it as a code word to mean "pleasure". In some papers, the word "reward" seems to be used as a rather monolithic, all-encompassing term that refers to any and all aspects of appetitive learning, motivation and emotion, whether conditioned or unconditioned. Used in this way, the term reward is a rather blunt instrument. These problems are not merely semantic, as it is difficult to test a hypothesis which maintains that a neurotransmitter mediates such an ill-defined set of functions. It has been suggested that it is advantageous to maintain the distinction between the terms reward and reinforcement; with this usage, reinforcement refers more directly to instrumental learning mechanisms (Wise, 2004; Sanchis-Segura and Spanagel, 2006), while reward connotes the primary motivational and emotional effects of reinforcing stimuli (Everitt and Robbins, 2005; Salamone et al., 2005, 2007).

Against the backdrop of these conceptual and terminological issues, there is a tremendous weight of empirical evidence that has built up against the various iterations of the DA hypothesis of "reward". It is somewhat ironic that the processes most directly linked to the use of the term reward (i.e., primary motivation, subjective pleasure) are the ones that have proven to be most problematic in terms of demonstrating the involvement of mesolimbic DA (Salamone et al., 2007). For example, low doses of DA antagonists and depletions of nucleus accumbens DA have been shown to produce effects that do not closely resemble extinction (Salamone, 1986; Salamone et al., 1995, 1997; Rick et al., 2006), pre-feeding (Salamone et al., 1991; Aberman and Salamone, 1999), or appetite suppressant drugs (Cousins et al., 1994; Salamone et al., 2002; Sink et al., 2008). Although it is well known that whole forebrain DA depletions can produce aphagia (i.e., lack of eating), it is DA depletions in the lateral or ventrolateral caudate/putamen, rather than the nucleus accumbens, which have most conclusively been linked to this effect (Ungerstedt, 1971; Dunnett and Iversen, 1982; Salamone et al., 1993a). It has been shown repeatedly that nucleus accumbens DA depletions or antagonism do not substantially impair appetite for food, or produce a general disruption of primary food motivation (Ungerstedt, 1971; Koob et al., 1978; Bakshi and Kelley, 1991; Salamone et al., 1993a). In DA deficient mice, restoration of DA production in caudate putamen, but not nucleus accumbens, was able to rescue feeding behavior (Szczypka

et al., 2001). In summarizing their findings that injections of DA D_1 or D_2 family antagonists into either the core or the shell subregions of nucleus accumbens impaired locomotion and rearing, but did not suppress food intake, Baldo et al. (2002) stated that DA receptor blockade "did not abolish the primary motivation to eat" (p. 176).

Furthermore, the idea that nucleus accumbens DA mediates the pleasure associated with positive reinforcers has been strongly challenged (Berridge, 2007; Salamone et al., 2007; Berridge and Kringlebach, 2008). Interference with accumbens DA transmission does not impair appetitive taste reactivity for sucrose (Berridge, 2007; Berridge and Kringlebach, 2008). Several studies in humans have reported that DA antagonists did not blunt the subjective euphoria produced by drugs of abuse (Gawin, 1986; Brauer and De Wit, 1997; Haney et al., 2001; Nann-Vernotica et al., 2001; Wachtel et al., 2002). Moreover, the potential role of DA systems in instrumental behavior or learning is not limited to situations involving appetitive motivation. There is considerable evidence that striatal mechanisms in general, and mesolimbic DA in particular, also participate in aspects of aversive learning and aversive motivation (Salamone, 1994; Munro and Kokkinidis, 1997; Blazquez et al., 2002; Pezze and Feldon, 2004; Delgado et al., 2008; Faure et al., 2008; Martinez et al., 2008). Although imaging studies often are used to support the idea that nucleus accumbens mediates pleasure (e.g., Sarchiapone et al., 2006; Wacker et al., 2009), this appears to be oversimplified; indeed, research employing various imaging methods has demonstrated that the human nucleus accumbens also responds to stress, aversion and hyperarousal/irritability (Liberzon et al., 1999; Jensen et al., 2003; Pavic, 2003; Phan et al., 2004; Pruessner et al., 2004; Levita et al., 2009). Physiological and neurochemical studies in animals clearly indicate that DA neuron activity is not simply tied to the delivery of primary reinforcers or pleasurable stimuli. Rather, VTA neuron activity and DA release can be activated by a number of different appetitive and aversive conditions (McCullough and Salamone, 1992; McCullough et al., 1993; Guarraci and Kapp, 1999; Roitman et al., 2004; Young, 2004; Anstrom and Woodward, 2005; Broom and Yamamoto, 2005; Marinelli et al., 2005; Schultz, 2007a,b; Brischoux et al., 2009), with changes seen across varying time scales, including tonic, slow phasic and fast phasic signals (Salamone, 1996; Salamone et al., 2007; Schultz, 2007a,b; Salamone, in press; see also Lapish et al., 2007 for a discussion of various time scales associated with the postsynaptic effects of DA release and DA receptor stimulation).

Of course, one would not want to throw the baby out with the bathwater. It is apparent that mesolimbic DA participates in several complex functions related to aspects of instrumental behavior, learning and incentive motivation, and pavlovian/instrumental interactions (Wise, 2004; Everitt and Robbins, 2005; Kelley et al., 2005; Salamone et al., 2005, 2007; Berridge, 2007; Robbins and Everitt, 2007; Redgrave et al., 2008; Yin et al., 2008). The more difficult aspect of research and theory in this area is to ask – which specific aspects? Exploration of these diverse areas of dopaminergic function has become a rich and fruitful area of inquiry. Indeed, this literature is so extensive that a thorough review of the behavioral functions of nucleus accumbens DA is beyond the scope of the present article (see Salamone et al., 2007). For the purposes of this special issue, the present review will focus upon the role of nucleus accumbens DA in effort-related processes, with a special emphasis on effort-based choice behavior that depends upon cost/benefit analyses.

BEHAVIORAL ACTIVATION, EXERTION OF EFFORT, AND NUCLEUS ACCUMBENS DA

Even as the popularity of the DA hypothesis of reward was growing during the 1980s, it was becoming apparent that there were alternative conceptual frameworks available for organizing what was known about the behavioral functions of DA systems, particularly mesolimbic DA. Mogenson et al. (1980) suggested that nucleus accumbens acted as a functional interface between the limbic system and the motor system, facilitating the ability of information related to emotion and motivation to impinge upon the neural systems involved in the instigation of action. It had been emphasized for several decades that behavioral activation, i.e., the vigor, persistence and effort seen in the pursuit of motivational stimuli, and the heightened activity induced by conditioned stimuli that predict reinforcers, was a fundamental aspect of motivation (e.g., Cofer and Appley, 1964). Several investigators suggested that DA systems were involved in behavioral activation. DA antagonists or accumbens DA depletions were shown to suppress the activities such as excessive drinking, wheel running, and locomotion that are induced by scheduled presentation of food (Robbins and Koob, 1980; Wallace et al., 1983; Salamone, 1986, 1988). It also was reported that the effects of DA antagonists on reinforced behavior interacted powerfully with the kinetic requirements of the instrumental response. For example, doses of DA antagonists that suppressed reinforced lever pressing had minimal effects on reinforced nose poking behavior (Ettenberg et al., 1981; Mekarski, 1988). Although 0.1 mg/kg haloperidol severely reduced responding on a fixed ratio (FR) 20 schedule of lever pressing, a dose four times that size had no effect on the reinforced response of simply being in proximity to the food dish on a fixed interval 30 s schedule (Salamone, 1986). As this research was being reported, investigators began to employ economic concepts, such as exertion of effort and cost-benefit analyses, to describe the behavioral functions of accumbens DA. Neill and Justice (1981) hypothesized that injection of amphetamine into nucleus accumbens could be increasing the "willingness" of rats to exert effort to obtain a given level of reinforcement. In a contemporary review of the behavioral functions of DA systems (Salamone, 1987), it was noted that DA in nucleus accumbens could be involved in the "exertion of effort", and it was suggested that future experiments could "offer animals choices between various reinforcers that are associated with operants of varying difficulty" (p. 602) so that researchers could determine if the allocation of behavioral resources could be biased toward or away from more or less effortful responses by administration of dopaminergic drugs.

This recognition of dopaminergic involvement in the exertion of effort, and effort-based choices related to cost benefit analyses, fit nicely with an emerging emphasis in the behavioral literature on work, response costs or constraints, and economic models of operant behavior. Several behavioral investigators have emphasized how response costs or constraints affect operant response output (Staddon, 1979; Kaufman, 1980; Kaufman et al., 1980; Foltin, 1991). Collier and colleagues studied how work requirements, such as the number of lever presses necessary for obtaining food, could

serve as determinants of response output and affect consumption parameters (Collier and Jennings, 1969; Johnson and Collier, 1987). Economic models of operant behavior have emphasized how a number of factors, including not only reinforcement value, but also conditions related to the characteristics of the instrumental response, can determine behavioral output (Lea, 1978; Allison, 1981, 1993; Bickel et al., 2000). Hursh et al. (1988) suggested that, in terms of behavioral economics, the price of food reinforcement as a commodity is a cost/benefit ratio expressed as the effort expended per unit of food value consumed. Optimal foraging theory was proposed to account for the observation that the amount of effort or time expended to obtain motivational stimuli was an important determinant of foraging choice (Krebs, 1977), an idea that is still very influential in the ethology research today (e.g., Hengeveld et al., 2009).

Over the last two decades, several lines of evidence have converged to strengthen the original observation that the effects of interference with DA transmission interact powerfully with the work requirements of an instrumental task. One of the ways of controlling work requirements in an operant schedule is to vary the ratio requirement (i.e., the number of times the animal must press the lever to receive a unit of reinforcement). The effects of the DA antagonist haloperidol on food-reinforced behavior were shown to be dependent upon the particular ratio schedule that was used [i.e., FR1 vs. progressive ratio; Caul and Brindle (2001)]. Accumbens DA depletions also produce effects that interact powerfully with the ratio requirement of the schedule employed. Ishiwari et al. (2004) found that accumbens DA depletions that substantially impaired FR5 lever pressing had no significant effect on FR1 performance. Aberman and Salamone (1999) systematically studied a wide range of ratio schedules (FR1, 4, 16, and 64) to assess the effects of accumbens DA depletions. While FR1 performance was not affected by DA depletion, and FR4 responding was only transiently and mildly suppressed, the schedules with large ratio requirements (i.e., FR16 and FR64) were severely impaired (**Figure 1A**). In fact, DA depleted animals responding on the FR64 schedule showed significantly fewer responses than those responding on the FR16 schedule (Aberman and Salamone, 1999). This pattern indicates that accumbens DA depletions exacerbate an effect known as *ratio strain*. In untreated animals, the overall relation between ratio size and response output is inverted-U shaped. Up to a point, as ratio requirements get larger, animals adjust to this challenge by increasing response output. However, if the ratio requirement is high enough (i.e., if the cost is too high), the animal reaches the point at which additional responses being required actually tend to suppress responding. For normal rats, responding at levels of FR64, FR100 or higher, even if there is only one 45 mg food pellet being delivered, does not seem to be problematic. A completely different function is shown by rats with accumbens DA depletions, which are much more sensitive to the size of the ratio requirement. In behavioral economic terms, this pattern can be described as reflecting an increase in the elasticity of demand for food reinforcement (Salamone et al., 1997; Aberman and Salamone, 1999; see **Figure 1B**). The term elasticity is widely used in economics, but price elasticity of demand refers to the sensitivity of consumption to changes in price (Vuchinich and Heather, 2003). Thus, if the ratio requirement is analogous to the price of the commodity (in this case, reinforcement pellets), it appears that rats with accumbens DA depletions are more sensitive than control animals to the price of the food reinforcers. Of course, rats do not use currency to purchase operant pellets; rather, it has been argued that an operant procedure is more of a barter system, in which the rat trades its work (or reductions in leisure) for a commodity (Rachlin, 2003). Thus, another way of describing this effect of impaired DA transmission is to say that rats with accumbens DA depletions are more sensitive than control animals to work-related response costs, or that they are less likely to trade high levels of work for food. In another study (Salamone et al., 2001), the increased effects of accumbens DA depletions with increasing ratio requirements were observed when rats were tested across a broader range of ratio schedules as high as FR300, even when the overall relation

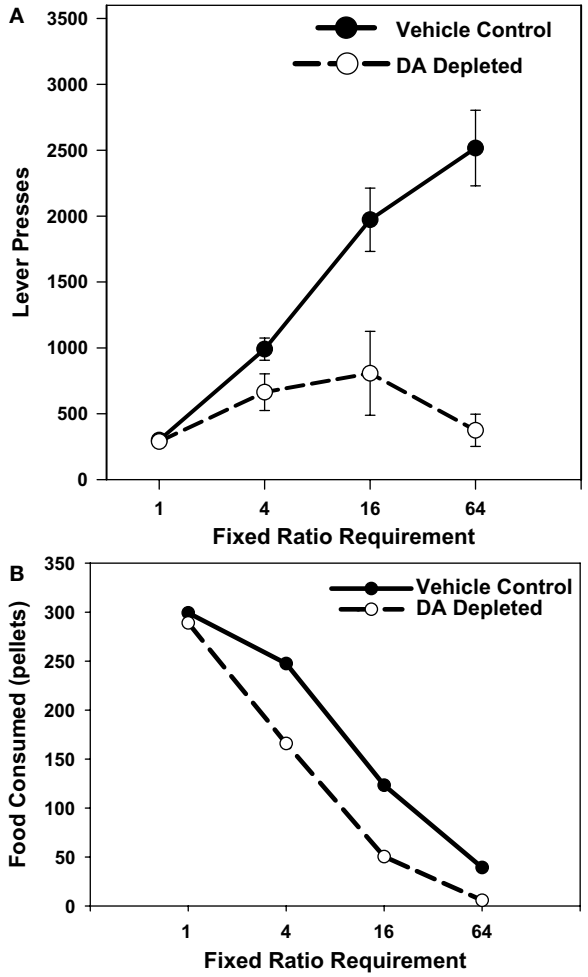


FIGURE 1 |This figure shows the effect of ratio requirement on the number of lever presses emitted and operant pellets consumed in rats with accumbens DA depletions compared to rats in the vehicle control group. Figure **(A)** is re-drawn based upon Aberman and Salamone (1999); these data are depicted in terms of number of responses, as in the original article. For **(B)**, the data are represented as number of operant pellets consumed. Each data point shown is the mean value from each group at each ratio level. Although comparable levels of consumption in DA depleted and control groups were seen with the FR1 schedule, DA-depleted rats showed markedly reduced consumption relative to the control group at higher ratio levels.

between lever pressing and food delivered per lever press was kept constant (i.e., FR50, one pellet every 50 responses; FR100, two pellets every 100 responses; FR200, four pellets every 200 responses; FR300, six pellets every 300 responses; Salamone et al., 2001). Thus, both the magnitude and the organization of the ratio requirement appear to be critical determinants of the sensitivity of an operant schedule to the effects of accumbens DA depletions.

In order to be sure that these results reflected the influence of ratio size, as opposed to other variables such as time, additional studies examined the effects of accumbens DA depletions on tandem schedules, in which a ratio requirement was attached to an interval requirement. In a conventional variable interval (VI) schedule, a time interval must elapse before the first response is reinforced, and the particular time interval varies around an average value. A tandem VI/FR schedule has an additional ratio requirement attached to the interval. For example, with a tandem VI 30 s/FR5 schedule, the animal is reinforced for the fifth response after the interval elapses, rather than the first. In this way, one can vary the ratio requirement of a schedule while keeping the programmed time intervals the same. Research employing tandem VI/FR schedules with varying combinations (e.g., VI 30 s/FR5, VI 60 s/FR10, VI 120 s/FR10) has yielded a consistent pattern; accumbens DA depletions do not impair overall response output in rats responding on the conventional VI schedules (i.e., those requiring only one response after the interval), but do substantially reduce responding on the corresponding VI schedule with the higher ratio requirement attached (Correa et al., 2002; Mingote et al., 2005). These results are consistent with research showing that accumbens DA antagonism did not impair performance on a progressive interval task (Wakabayashi et al., 2004), and suggest that interval requirements *per se* do not pose a severe constraint to rats with compromised DA transmission in nucleus accumbens. This serves to underscore the critical importance of ratio requirements as providing a work-related challenge to rats with accumbens DA depletions or antagonism.

In summarizing these results, Salamone and Correa (2002) stated that nucleus accumbens DA depletions appear to have two major effects: (1) they reduce the response-enhancing effects that moderate-size ratio requirements have on operant responding (i.e., the ascending limb of the function relating ratio requirement to response output), and (2) they enhance the response-suppressing effects that very large ratios have on operant responding (i.e., the descending limb of the function, enhancing ratio strain). Furthermore, finer grained analyses of detailed patterns of responding reveal more insights into the behavioral manifestations of accumbens DA depletions. Accumbens DA depletions produce a slight reduction in the local rate of responding, as indicated by the distribution of inter-response times (Salamone et al., 1993b, 1999; Mingote et al., 2005). In addition, they enhance pauses in responding (Salamone et al., 1993b; Mingote et al., 2005). The latter may indicate a fragmentation in the pattern of responding (Mingote et al., 2005), a reduction in the ability to sustain uninterrupted response output, or a lack of engagement in the task (Nicola, 2007). Recently, computational approaches have been used to analyze these effects of accumbens DA depletions on response rate (e.g., Niv et al., 2007; Phillips et al., 2007). This relation between response output and DA function has been interpreted to mean that DA release in nucleus accumbens could provide a window of opportunistic drive during which the threshold cost expenditure to obtain the reward is decreased (Phillips et al., 2007).

In discussing the effects of dopaminergic manipulations on ratio performance, it is useful to mention the term "reinforcement efficacy", which is sometimes used to describe the effects of drug manipulations on progressive ratio performance. With progressive ratio schedules, the ratio requirement increases as successive ratios are completed, and the "break point" is said to occur at the point at which the animal essentially ceases to respond. One can operationally define reinforcement efficacy in terms of the break point in a progressive ratio schedule (and also by measuring ratio strain in rats responding across different FR schedules). The determination of reinforcement efficacy can be a very useful tool for characterizing some of the fundamental reinforcing actions of drugs that are self-administered, and for comparing self-administration behavior across different substances or classes of substances (e.g., Richardson and Roberts, 1996; Marinelli et al., 1998; Woolverton and Ranaldi, 2002). Used in this manner, reinforcement efficacy is essentially being employed as an empirical descriptor of a particular behavioral outcome. Nevertheless, given the terminological problems mentioned above, it is worth emphasizing that the term "reinforcement efficacy" should not be used simply as a replacement for "reward", nor should progressive ratio breakpoints be viewed as necessarily providing some direct and unambiguous measure related to the subjective pleasure produced by the stimulus (Salamone, 2006). Changes in progressive ratio break points can reflect more than just changes in the appetitive motivational properties of a reinforcing stimulus (Richardson and Roberts, 1996; Hamill et al., 1999). For example, changing the kinetic requirements of the instrumental response (e.g., increasing the height of the lever) was shown to decrease progressive ratio break points (Skjoldager et al., 1993; Schmelzeis and Mittleman, 1996). Although some researchers have maintained that the break point provides a direct measure of the appetitive motivational characteristics of a stimulus, it is, as explicitly stated in a classic review by Stewart (1974), more directly a measure of how much work the organism will do in order to obtain that stimulus. Progressive ratio break points and measures of ratio strain are essentially outcomes that result from effort-related decision making processes. The animal is making a cost/benefit choice about whether or not to continue to respond, based partly on factors related to the reinforce itself, but also upon the work-related response costs and time constraints imposed by the ratio schedule. For these reasons, interpretations of the actions of drugs or lesions on progressive ratio break points should be done with caution, as should be the case for any individual task. A drug that alters the break point could do so for many different reasons; it may be affecting functions related to the processing of reward value, or alternatively it could be affecting exertion of effort, or decision making processes.

RESPONSE ALLOCATION, EFFORT-RELATED CHOICE BEHAVIOR, AND NUCLEUS ACCUMBENS DA

The ability to exert effort, sustain work, overcome obstacles, and attain access to motivationally relevant stimuli is necessary for survival. But it is only part of the story. In a complex environment, which affords many opportunities for obtaining significant stimuli, and multiple paths for accessing them, organisms must make choices. The variables that need to be evaluated to make these decisions are complex

and multidimensional, but among the most important are those involving cost/benefit assessments based upon effort and reinforcement value (Salamone and Correa, 2002; Salamone et al., 2003, 2005, 2007; van den Bos et al., 2006; Walton et al., 2006). Considerable evidence indicates that nucleus accumbens DA, along with other transmitters and structures, participates in the overall circuitry that regulates effort-based choice behavior (Salamone et al., 2003, 2005, 2007; Floresco et al., 2008a; Hauber and Sommer, 2009).

One of the procedures that has been used to assess the contribution of accumbens DA to response allocation and effort-related choice behavior is a task that offers rats the option of either lever pressing to obtain a relatively preferred food (e.g., Bioserve pellets; usually obtained on a FR5 schedule), or approaching and consuming a less preferred food (lab chow) that is concurrently available in the chamber. Well trained rats under baseline conditions typically get most of their food by lever pressing, and consume only small quantities of chow (Salamone et al., 1991). Low-to-moderate doses of DA antagonists, which block either D_1 or D_2 family receptor subtypes, produce a substantial alteration of response allocation in rats performing on this task. The DA antagonists *cis*-flupenthixol, haloperidol, raclopride, eticlopride, SCH 23390, SKF83566, and ecopipam all decreased lever pressing for food but substantially increased intake of the concurrently available chow (Salamone et al., 1991, 1996, 2002; Cousins et al., 1994; Sink et al., 2008; Worden et al., 2009). The use of this task for assessing effort-related choice behavior has been validated in many ways. For example, the low dose of haloperidol that produced the shift from lever pressing to chow intake (0.1 mg/kg) did not affect total food intake or alter preference between these two specific foods in free-feeding choice tests (Salamone et al., 1991). Although DA antagonists have been shown to reduce FR5 lever pressing and increase chow intake, appetite suppressants from different classes, including amphetamine (Cousins et al., 1994), fenfluramine (Salamone et al., 2002), and cannabinoid CB1 antagonists (Sink et al., 2008), failed to increase chow intake at doses that suppressed lever pressing. Similarly, pre-feeding to reduce food motivation was shown to suppress both lever pressing and chow intake (Salamone et al., 1991). Furthermore, attachment of higher ratio requirements (up to FR20) caused animals that were not drug treated to shift from lever pressing to chow intake (Salamone et al., 1997), indicating that this task is sensitive to work load. Together with other results, these findings demonstrate that interference with DA transmission does not simply reduce appetite, but does act to alter response allocation between alternative sources of food that can be obtained through different instrumental responses.

The shift from lever pressing to chow intake in rats performing on this task is associated with DA depletions in nucleus accumbens, but not the neostriatum. Although it has been suggested that caudate/putamen DA may have some types of motivational functions related to feeding (Palmiter, 2007), DA depletions in anteroventromedial neostriatum, which is dorsal to nucleus accumbens, had no behavioral effect, while ventrolateral neostriatal DA depletions produced severe motor impairments that merely decreased both lever pressing and feeding (Cousins et al., 1993). In contrast, decreases in lever pressing and increases in chow intake occur as a result of accumbens DA depletions, as well as intra-accumbens injections of D_1 or D_2 antagonists (Salamone et al., 1991; Cousins et al., 1993; Cousins and Salamone, 1994; Sokolowski and Salamone, 1998; Koch et al., 2000; Nowend et al., 2001). The shift from lever pressing to chow intake on this task has been shown to occur in rats if D_1 or D_2 family antagonist are injected into the medial core, lateral core, or dorsal shell subregions of the accumbens (Salamone et al., 1991; Nowend et al., 2001). Thus, although lever pressing is decreased by accumbens DA antagonism or depletions, the rats show a compensatory reallocation of behavior and select a new path to an alternative food source. Consistent with these effects observed in rats that have impaired DA transmission, DA transporter knockdown mice, which have enhanced DA transmission, show increased selection of lever pressing relative to chow intake when tested with this task (Cagniard et al., 2006).

Salamone et al. (1994) also developed a T-maze procedure in order to assess the effects of DA antagonists and accumbens DA depletions on effort-related decision making. With this procedure, the two choice arms of the maze can have different reinforcement densities (e.g., four vs. two food pellets, or four vs. zero food pellets), and under some conditions a 44-cm barrier can be placed in the arm with the higher density of food reinforcement to present an effort-related challenge to the rat. When no barrier is placed in the arm with the high reinforcement density, rats mostly choose that arm, and neither haloperidol nor accumbens DA depletion alters their response choice (Salamone, 1994). When the arm with the barrier contained four pellets, but the other arm contained no pellets, rats with accumbens DA depletions were very slow, but still managed to choose the high density arm, climb the barrier, and consume the pellets (Cousins et al., 1996). Yet accumbens DA depletions dramatically altered choice behavior when the high density arm (four pellets) had the barrier in position, and the arm without the barrier contained an alternative food source (two pellets). In this case, DA depletions or antagonism decreased choice for the high density arm, and increased choice for the low density arm (Salamone, 1994; Cousins et al., 1996; Denk et al., 2005; Mott et al., 2009). Like the operant concurrent choice task, the T-maze task for measuring effort-based choice behavior also has undergone considerable behavioral validation and evaluation (Salamone, 1994; Cousins et al., 1996; van den Bos et al., 2006; Correa et al., 2009). For example, in a recent T-maze choice study with mice, it was confirmed that haloperidol reduced choice of the arm with the barrier, and it also was demonstrated that haloperidol had no effect on choice when both arms had a barrier in place (Correa et al., 2009). Thus, dopaminergic manipulations did not alter the preference for the high density of food reward over the low density, and did not affect discrimination or memory processes related to arm preference. Over the last several years, variants of this task have been used by several laboratories to characterize the effects of brain lesions or drug manipulations (Salamone, 1994; Walton et al., 2003; Denk et al., 2005; Schweimer and Hauber, 2005; van den Bos et al., 2006; Bardgett et al., 2009; Hauber and Sommer, 2009; Mott et al., 2009). The results of the T-maze studies in rodents, together with the findings from the operant concurrent choice studies reviewed above, indicate that low doses of DA antagonists and accumbens DA depletions cause animals to reallocate their instrumental response selection based upon the response requirements of the task, and select lower effort alternatives for obtaining rewards (see reviews by Salamone et al., 2003, 2005, 2007; Floresco et al., 2008a).

Recent papers have used effort discounting procedures to study the effects of dopaminergic manipulations. Floresco et al. (2008b) investigated the effects of dopaminergic and glutamatergic drugs on both effort and delay discounting. The DA antagonist haloperidol altered effort discounting even when the effects of time delay were controlled for (Floresco et al., 2008b). A T-maze effort discounting task was recently developed (Bardgett et al., 2009), in which the amount of food in the high density arm of the maze was diminished each trial on which the rats selected that arm (i.e., an "adjusting-amount" discounting variant of the T-maze procedures, which allows for the determination an indifference point for each rat). Administration of both the D_1 family antagonist SCH23390 and the D_2 family antagonist haloperidol altered effort discounting, making it more likely that rats would choose the arm with the smaller reward. Increasing DA transmission by administration of amphetamine blocked the effects of SCH23390 and haloperidol, and also biased rats toward choosing the high reward/high cost arm, which is consistent with operant choice studies using DA transporter knockdown mice (Cagniard et al., 2006). Together with other results, the findings reported by Bardgett et al. (2009) and Floresco et al. (2008b) support the suggestion that, across a variety of conditions, DA transmission exerts a bidirectional influence over effort-related decision making.

INTERACTIONS BETWEEN DA AND ADENOSINE

As reviewed above, considerable research has demonstrated that DA antagonists and accumbens DA depletions affect behavioral activation, instrumental response output, response allocation, and effort-related processes (Salamone et al., 1991, 2007; Salamone and Correa, 2002; Phillips et al., 2007; Robbins and Everitt, 2007; Floresco et al., 2008a). Clearly, DA does not participate in effort-related processes in isolation, and for that reason it is important to review how other brain areas and neurotransmitters interact with dopaminergic mechanisms. Within the last few years, considerable emphasis has been placed upon interactions between DA and adenosine. Caffeine and other methylxanthines, which act as minor stimulants, are non-selective adenosine antagonists (Ferré et al., 2008). Recently, there has been a rapid growth of research on adenosine receptor neurochemistry and pharmacology, particularly concerning the A_{2A} subtype of adenosine receptor. DA-rich striatal areas, including both the caudate/putamen (neostriatum) and the nucleus accumbens, have a very high degree of adenosine A_{2A} receptor expression (Schiffmann et al., 1991; DeMet and Chicz-DeMet, 2002; Ferré et al., 2004). There is considerable evidence of a functional interaction between striatal DA D_2 and adenosine A_{2A} receptors (Fink et al., 1992; Ferré, 1997; Hillion et al., 2002; Fuxe et al., 2003). This interaction frequently has been studied in regard to neostriatal motor functions that are related to parkinsonism (Ferré et al., 1997, 2001; Hauber and Munkel, 1997; Svenningsson et al., 1999; Hauber et al., 2001; Wardas et al., 2001; Morelli and Pinna, 2002; Correa et al., 2004; Jenner, 2005; Pinna et al., 2005; Ishiwari et al., 2007; Salamone et al., 2008a,b). Several reports also have characterized aspects of adenosine A_{2A} receptor function related to cognitive processes (Takahashi et al., 2008), anxiety (Correa and Font, 2008), and motivation (Salamone et al., 2007; Mingote et al., 2008).

Adenosine A_{2A} receptors also are involved in aspects of behavioral activation and effort-related processes (Farrar et al., 2007; Font et al., 2008; Mingote et al., 2008; Mott et al., 2009; Worden et al., 2009). Injections of the adenosine A_{2A} agonist CGS 21680 directly into the accumbens can produce effects that resemble those of accumbens DA depletions or antagonism. Intra-accumbens injections of CGS 21680 were shown to reduce locomotor activity (Barraco et al., 1993). Local infusions of CGS 21680 into the accumbens reduced responding on a VI 60 s schedule with a FR10 requirement attached, but did not impair performance on a conventional VI 60 s schedule (Mingote et al., 2008); this pattern is similar to that previously shown with accumbens DA depletions (Mingote et al., 2005). In rats responding on the operant FR5/chow feeding concurrent choice procedure, injections of CGS 21680 into the accumbens decreased lever pressing and increased chow intake (Font et al., 2008), a pattern of effects similar to that produced by accumbens DA depletions and antagonism. Consistent with the observation that an adenosine A_{2A} agonist could produce actions similar to those resulting from DA depletion or blockade, it also has been reported the locomotor suppression induced by the DA antagonist haloperidol was reduced by injections of the adenosine A_{2A} antagonist MSX-3 into nucleus accumbens core, but not into the shell or the ventrolateral neostriatum (Ishiwari et al., 2007). Furthermore, it has been demonstrated that adenosine A_{2A} receptor antagonists can reverse the effects of DA D_2 antagonists on both the operant concurrent choice task (Farrar et al., 2007; Salamone et al., 2009; Worden et al., 2009) and the T-maze choice procedure (Correa et al., 2009; Mott et al., 2009). Recently, studies with intracranial injections revealed that systemic or intra-accumbens injections of the adenosine A_{2A} antagonist MSX-3 were able to block the effects of intra-accumbens injections of the D_2 antagonist eticlopride in rats responding on the operant concurrent choice task (Farrar, 2009, unpublished doctoral dissertation, University of Connecticut).

These studies afford an interesting opportunity to assess the overall interaction between DA and adenosine receptor subtypes. Adenosine A_{2A} receptor antagonists MSX-3 and KW 6002 reliably attenuate the effects of D_2 antagonists such as haloperidol and eticlopride in rats responding on the operant concurrent choice procedure (Farrar et al., 2007; Salamone et al., 2009; Worden et al., 2009). In contrast, MSX-3 was relatively ineffective at reducing the effects of the D_1 antagonist ecopipam (SCH 39166; Worden et al., 2009) on this task. Although the non-selective adenosine antagonist caffeine was able to partially reverse the effects of haloperidol on the concurrent choice task, DPCPX, which is highly selective for the adenosine A_1 receptor subtype, was ineffective (Salamone et al., 2009). Similar results were obtained with rats and mice responding on the T-maze barrier choice task. Although MSX-3 was able to reverse the effect of haloperidol on selection of the arm with the barrier, the A_1 antagonists DPCPX and CPT were not (Correa et al., 2009; Mott et al., 2009).

The results described above demonstrate that there is a relatively selective interaction between DA D_2 and adenosine A_{2A} receptor subtypes (**Table 1**). Based upon anatomical studies, it appears that this is likely to be due to the pattern of cellular localization of adenosine A_1 and A_{2A} receptors in striatal areas, including the nucleus accumbens (Ferré, 1997). Adenosine A_{2A} receptors are typically co-localized on striatal and accumbens enkephalin-positive

Table 1 | Interactions between dopamine and adenosine receptor antagonists.

	Adenosine receptor antagonist			
	Non-selective	**A_1**	**A_{2A}**	
CONCURRENT FR5/FREE CHOW				
D_2 receptor antagonist	Reversal	No reversal	Reversal	Farrar et al. (2007), Salamone et al. (2009), Worden et al. (2009)
D_1 receptor antagonist	–	No reversal	Partial reversal[1]	Worden et al. (2009)[1], Nunes et al. (2009)[2]
T-MAZE WITH BARRIER				
D_2 receptor antagonist	Reversal	No reversal	Reversal	Mott et al. (2009), Pardo (2009)[3]

[1]*There was a mild increase in lever pressing in ecopipam-treated rats that received the A_{2A} antagonist MSX-3, but no reversal of the chow intake effect of the D_1 antagonist.*
[2]*Data from Nunes et al. (2009).*
[3]*Data from Pardo (2009), unpublished masters thesis, University of Jaume I.*

medium spiny neurons with DA D_2 family receptors, and these receptors converge onto the same signal transduction pathways and show the capacity for forming heteromeric complexes (Fink et al., 1992; Ferré, 1997; Svenningsson et al., 1999; Hillion et al., 2002; Fuxe et al., 2003). Thus, adenosine A_{2A} receptor antagonists appear to be so effective in reversing the effort-related actions of D_2 antagonists because of direct interactions between DA D_2 and adenosine A_{2A} receptors located on the same neurons (**Figure 2**). On the other hand, DA D_1 receptors are more likely to be co-localized with adenosine A_1 receptors (Ferré, 1997), which could help to explain why it is more difficult for adenosine A_1 receptor antagonists to reverse the effects of D_2 receptor blockade. Interestingly, despite the fact that D_1 and A_1 receptors tend to be co-localized on the same neurons, the A_1 antagonists DPCPX and CPT were unable to reverse the effects of the D_1 antagonist ecopipam in rats responding on the concurrent choice operant procedure (Nunes et al., 2009). This suggests that A_{2A} antagonists exert an overall greater effect than A_1 antagonists on effort-related functions of nucleus accumbens.

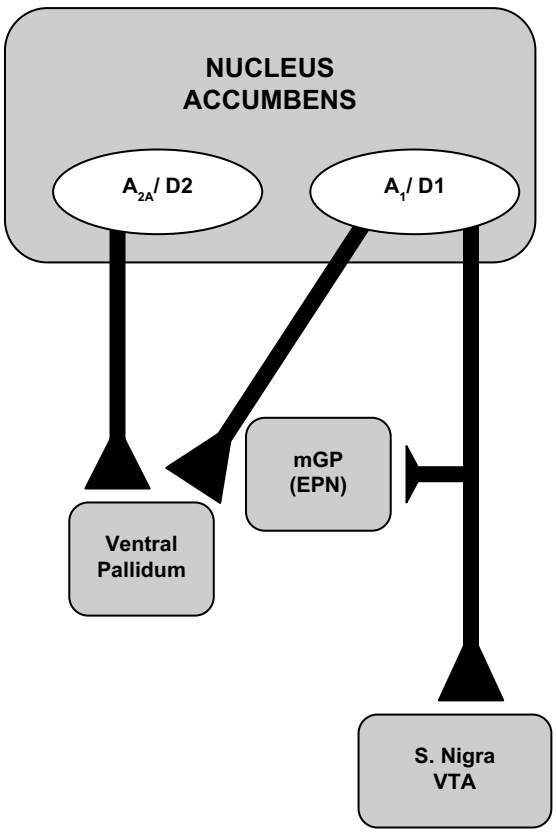


FIGURE 2 | Anatomical diagram depicting the pattern of DA and adenosine receptor localization in nucleus accumbens. See text for details (see also Ferré, 1997; Hillion et al., 2002; Fuxe et al., 2003). mGP, medial globus pallidus; epn, entopeduncular nucleus; s. nigra, substantia nigra; VTA, ventral tegmental area.

BEHAVIORAL THEORY AND ANALYSES: FURTHER EVALUATION OF EFFORT-RELATED PROCESSES

Research on the brain mechanisms involved in effort-related processes may lead to new ways of thinking about behavioral analysis and theory in behavioral economics. One of the contributions that behavioral neuroscience can make to behavioral theory is to use manipulations (e.g., drugs, lesions) that dissociate complex behavioral processes into component parts (Salamone et al., 2007). In this regard, it is useful to consider that a given parameter that is generated from curve-fitting analyses, when viewed in terms of its biological characteristics, has many factors that contribute to it. A good example of this is the ED_{50}, which is used in pharmacology to provide a measure of the potency of a drug based upon dose-response analysis. Empirically, the ED_{50} is the dose that produces an effect that is 50% of the maximal effect. Although the ED_{50} is expressed as one number, that simplicity is deceptive because many biochemical factors contribute to it, including the affinity of a drug for a receptor, duration of action, drug metabolism, and penetration into the target tissue. A useful example of this principle from the behavioral neuroscience literature is the progressive ratio break point; as discussed above, this measure also has many factors that can contribute to it. Another case in which this point is important to consider is threshold measures used in intracranial self-stimulation studies. Such measures often are viewed as providing "rate-free" indices of reinforcement value, nevertheless, they are influenced by lever pressing ratio requirements as well as the electrical current level (Fouriezos et al., 1990).

Some research related to behavioral economics, reinforcer value, and the functions of DA systems has used response-reinforcement matching methods (e.g., Heyman and Monaghan, 1987; Aparicio, 2007). Matching equations have been employed to describe the results of studies with both conventional and concurrent VI schedules, and one of the parameters (R_e) can be used to represent reinforcement value (e.g., Herrnstein, 1974; see equation below for single-lever conventional VI schedules, in which B represents response rate, R represents reinforcement density, k is the constant for maximal responding, and R_e represents the reinforcement level that generates 50% of maximum responding).

$$B = k\,R/(R + R_e)$$

However, used in this way, R_e does not selectively represent only the reinforcement value of food *per se*; actually, it reflects the relative value of lever pressing for and consuming the food reinforcer compared to the reinforcing value of all other stimuli and responses available (Salamone et al., 1997). Several factors can contribute to this composite measure, which is one of the reasons why other matching equations have been developed that account for deviations from matching by allowing for estimates of reinforcer sensitivity, as well as response preference or bias (Baum, 1974; Williams, 1988; Aparicio, 2001). Clearly, a drug or lesion manipulation could yield apparent effects on "reinforcement value" that actually reflect changes in response bias (Salamone, 1987; Salamone et al., 1997).

For these reasons, it may be useful to think more deeply about how terms such as value are used in neuroeconomics research. The aggregate reinforcement value of an instrumental activity (e.g., lever pressing for and consuming food) should perhaps be viewed as a composite measure that includes both the reinforcing value of the reinforcer itself, plus any net value or costs associated with the instrumental response that is required to obtain the reinforcer. Viewed in this way, the effects of dopaminergic manipulations on effort-related choice behavior could be described in terms of actions upon the response costs associated with the particular reinforcer, rather than the reinforcing value of the food stimulus itself. Although the effects of haloperidol on bias may be minimal when two levers that are relatively similar are used (e.g., Aparicio, 2007), they may be much larger when very different responses are compared (e.g., lever pressing vs. sniffing; lever pressing vs. unrestricted access to food; barrier climbing vs. locomotion). Future research will determine if measures of bias based upon the matching equations, or some other type of mathematical formulation, would be the best way to capture these drug effects quantitatively.

SUMMARY AND CONCLUSIONS

In summary, DA and adenosine in the nucleus accumbens interact to regulate effort-related functions. Additional research has shown that a number of components of the cortico-striato-pallidal loop system also are involved (Walton et al., 2006; Floresco and Ghods-Sharifi, 2007; Farrar et al., 2008; Mingote et al., 2008; Hauber and Sommer, 2009). Disconnection studies have revealed that serial connections between basolateral amygdala, anterior cingulate cortex, nucleus accumbens, and ventral pallidum are involved in the exertion of effort and effort-related choice behavior (Floresco and Ghods-Sharifi, 2007; Farrar et al., 2008; Mingote et al., 2008; Hauber and Sommer, 2009). Within the last few years, there has been considerable progress in characterizing the functional anatomy underlying this important aspect of motivation and decision making. Several transmitters across multiple brain regions are involved in effort-related functions, and researchers are only beginning to piece together the complex puzzle of all the potential brain systems that are involved. Presently, the specific way in which each structure contributes to the overall function of the system is unclear. It is uncertain which brain areas are involved in the exertion of effort, or the perception of effort, vs. the actual decision making process itself. For example, it is possible that nucleus accumbens is involved in the actual decision making processes, but it also is possible that it is mainly involved in regulating energy output, or setting effort-related constraints or feedback that in turn influences decisions made at other levels in the system. If the latter is true, then it is possible that the decision making effects of drug or lesion manipulations of nucleus accumbens are an outcome reflecting the constraints that are set after compromised DA function in accumbens, rather than a direct effect upon decision making processes *per se*. Future research will be necessary to tease apart these distinct aspects of effort-related function.

In addition to providing insights into aspects of animal behavior and natural motivation, research on effort-related processes also has clinical implications. Within the last few years, there has been a greater emphasis upon effort-related functions involved in drug self-administration (e.g., Vezina et al., 2002). Drug seeking behavior in humans involves many psychological processes, including effort. Addicts will go to great lengths to obtain their preferred drug, overcoming numerous obstacles and constraints, both behavioral and economic. Furthermore, addiction is characterized not only by a re-organization of the preference structure of the person, but also by a dramatic change in the allocation of behavioral resources toward the addictive substance; there is a heightened emphasis upon drug seeking and drug taking, typically at the expense of other motivational activities. As well as being related to aspects of drug taking and addiction, research on behavioral activation and effort has implications for understanding the neural basis of psychiatric symptoms such as psychomotor slowing, anergia, fatigue and apathy, which are seen in depression as well as other psychiatric or neurological conditions (Salamone et al., 2006, 2007). These motivational symptoms, which can have devastating behavioral manifestations (Stahl, 2002; Demyttenaere et al., 2005), represent impairments in aspects of behavioral activation and effort that can lead to problems in the workplace, as well as limitations in terms of life function, interaction with the environment, and responsiveness to treatment. There is considerable overlap between the neural circuitry involved in effort-related functions in animals and the brain systems that have been implicated in psychomotor slowing and anergia in depression (Salamone et al., 2006). Thus, research on effort-related behavioral processes, and their neural regulation, could have substantial impact on clinical research related to addiction, depression, and other disorders.

ACKNOWLEDGMENTS

Much of the work cited in this review was supported by a grant to John D. Salamone from the US NIH/NIMH (MH078023), and to Merce Correa from C. Sanitat G.V. AP04108.

REFERENCES

Aberman, J. E., and Salamone, J. D. (1999). Nucleus accumbens dopamine depletions make rats more sensitive to high ratio requirements but do not impair primary food reinforcement. *Neuroscience* 92, 545–552.

Allison, J. (1981). Economics and operant conditioning. In Predictability, Correlation and Contiguity, P. Harzem and M. D. Zeiler, eds (New York, John Wiley and Sons), pp 321–353.

Allison, J. (1993). Response deprivation, reinforcement, and economics. *J. Exp. Anal. Behav.* 60, 129–140.

Anstrom, K. K., and Woodward, D. J. (2005). Restraint increases dopaminergic burst firing in awake rats. *Neuropsychopharmacology* 30, 1832–1840.

Aparicio, C. F. (2001). Overmatching in rats: the barrier choice paradigm. *J. Exp. Anal. Behav.* 75:93–106.

Aparicio, C. F. (2007). Haloperidol, dynamics of choice, and the parameters of the matching law. *Behav. Processes* 75, 206–212.

Bakshi, V. P., and Kelley, A. E. (1991). Dopaminergic regulation of feeding behavior: I. Differential effects of haloperidol microinjection in three striatal subregions. *Psychobiology* 19, 223–232.

Baldo, B. A., and Kelley, A. E. (2007). Discrete neurochemical coding of distinguishable motivational processes: insights from nucleus accumbens control of feeding. *Psychopharmacology (Berl.)* 191, 439–459.

Baldo, B. A., Sadeghian, K., Basso, A. M., and Kelley, A. E. (2002). Effects of selective dopamine D1 or D2 receptor blockade within nucleus accumbens subregions on ingestive behavior and associated motor activity. *Behav. Brain Res.* 137, 165–177.

Barbano, M. F., and Cador, M. (2007). Opioids for hedonic experience and dopamine to get ready for it. *Psychopharmacology* 191, 497–506.

Bardgett, M. E., Depenbrock, M., Downs, N., Points, M., and Green, L. (2009). Dopamine modulates effort-based decision making in rats. *Behav. Neurosci.* 123, 242–251.

Barraco, R. A., Martens, K. A., Parizon, M., and Normile, H. J. (1993). Adenosine A2a receptors in the nucleus accumbens mediate locomotor depression. *Brain Res. Bull.* 31, 397–404.

Baum, W. M. (1974). On two types of deviation from the matching law: bias and undermatching. *J. Exp. Anal. Behav.* 22, 231–242.

Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology* 191, 391–431.

Berridge, K. C., and Kringlebach, M. L. (2008). Affective neuroscience of pleasure: reward in humans and animals. *Psychopharmacology* 199, 457–480.

Bickel, W. K., Marsch, L. A., and Carroll, M. E. (2000). Deconstructing relative reinforcing efficacy and situating the measures of pharmacological reinforcement with behavioral economics: a theoretical proposal. *Psychopharmacology* 153, 44–56.

Blazquez, P. M., Fujii, N., Kojima, J., and Graybiel, A. M. (2002). A network representation of response probability in the striatum. *Neuron* 33, 973–982.

Brauer, L. H., and De Wit, H. (1997). High dose pimozide does not block amphetamine-induced euphoria in normal volunteers. *Pharmacol. Biochem. Behav.* 56, 265–272.

Brischoux, F., Chakraborty, S., Brierley, D. I., and Ungless, M. A. (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proc. Natl. Acad. Sci. USA* 106, 4894–4899.

Broom, S. L., and Yamamoto, B. K. (2005). Effects of subchronic methamphetamine exposure on basal dopamine and stress-induced dopamine release in the nucleus accumbens shell of rats. *Psychopharmacology* 181, 467–476.

Cagniard, B., Balsam, P. D., Brunner, D., and Zhuang, X. (2006). Mice with chronically elevated dopamine exhibit enhanced motivation, but not learning, for a food reward. *Neuropsychopharmacology* 31, 1362–1370.

Cannon, C. M., and Bseikri, M. R. (2004). Is dopamine required for natural reward? *Physiol. Behav.* 81, 741–748.

Caul, W. F., and Brindle, N. A. (2001). Schedule-dependent effects of haloperidol and amphetamine: multiple-schedule task shows within-subject effects. *Pharmacol. Biochem. Behav.* 68, 53–63.

Cofer, C. N., and Appley, M. H. (1964). Motivation: Theory and Research. New York, John Wiley and Sons.

Collier, G. H., and Jennings, W. (1969). Work as a determinant of instrumental performance. *J. Comp. Physiol. Psychol.* 68, 659–662.

Correa, M., Carlson, B. B., Wisniecki, A., and Salamone, J. D. (2002). Nucleus accumbens dopamine and work requirements on interval schedules. *Behav. Brain Res.* 137, 179–187.

Correa, M., and Font, L. (2008). Is there a major role for adenosine A2A receptors in anxiety? *Front. Biosci.* 13, 4058–4070.

Correa, M., Pardo, M., Hockemeyer, J., Müller, C. E., and Salamone, J. D. (2009). Dopamine D2 Antagonism Increases Selection of Less Effortful Food-seeking Behavior in a T-maze Procedure in Mice: Reversal of the Motivational Impairment with Adenosine Antagonists. Neuroscience Meeting Planner. Washington, DC, Society for Neuroscience (online).

Correa, M., Wisniecki, A., Betz, A., Dobson, D. R., O'Neill, M. F., O'Neill, M. J., and Salamone, J. D. (2004). The adenosine A2A antagonist KF17837 reverses the locomotor suppression and tremulous jaw movements induced by haloperidol in rats: possible relevance to parkinsonism. *Behav. Brain Res.* 148, 47–54.

Cousins, M. S., Atherton, A., Turner, L., and Salamone, J. D. (1996). Nucleus accumbens dopamine depletions alter relative response allocation in a T-maze cost/benefit task. *Behav. Brain Res.* 74, 189–197.

Cousins, M. S., and Salamone, J. D. (1994). Nucleus accumbens dopamine depletions in rats affect relative response allocation in a novel cost/benefit procedure. *Pharmacol. Biochem. Behav.* 49, 85–91.

Cousins, M. S., Sokolowski, J. D., and Salamone, J. D. (1993). Different effects of nucleus accumbens and ventrolateral striatal dopamine depletions on instrumental response selection in the rat. *Pharmacol. Biochem. Behav.* 46, 943–951.

Cousins, M. S., Wei, W., and Salamone, J. D. (1994). Pharmacological characterization of performance on a concurrent lever pressing/feeding choice procedure: effects of dopamine antagonist, cholinomimetic, sedative and stimulant drugs. *Psychopharmacology* 116, 529–537.

Delgado, M. R., Li, J., Schiller, D., and Phelps, E. A. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 363, 3787–3800.

DeMet, E. M., and Chicz-DeMet, A. (2002). Localization of adenosine A2A-receptors in rat brain with [3H]ZM-241385. *Naunyn Schmiedebergs Arch. Pharmacol.* 366, 478–481.

Demyttenaere, K., De Fruyt, J., and Stahl, S. M. (2005). The many faces of fatigue in major depressive disorder. *Int. J. Neuropsychopharmacol.* 8, 93–105.

Denk, F., Walton, M. E., Jennings, K. A., Sharp, T., Rushworth, M. F., and Bannerman, D. M. (2005). Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology* 179, 587–596.

Dunnett, S. B., and Iversen, S. D. (1982). Regulatory impairments following selective 6-OHDA lesions of the neostriatum. *Behav. Brain Res.* 4, 195–202.

Ettenberg, A., Koob, G. F., and Bloom, F. E. (1981). Response artifact in the measurement of neuroleptic-induced anhedonia. *Science* 213, 357–359.

Everitt, B. J., and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489.

Farrar, A. M., Font, L., Pereira, M., Mingote, S. M., Bunce, J. G., Chrobak, J. J., and Salamone, J. D. (2008). Forebrain circuitry involved in effort-related choice: injections of the $GABA_A$ agonist muscimol into ventral pallidum alters response allocation in food-seeking behavior. *Neuroscience* 152, 321–330.

Farrar, A. M., Pereira, M., Velasco, F., Hockemeyer, J., Muller, C. E., and Salamone, J. D. (2007). Adenosine A(2A) receptor antagonism reverses the effects of dopamine receptor antagonism on instrumental output and effort-related choice in the rat: implications for studies of psychomotor slowing. *Psychopharmacology* 191, 579–586.

Faure, A., Reynolds, S. M., Richard, J. M., and Berridge, K. C. (2008). Mesolimbic dopamine in desire and dread: enabling motivation to be generated by localized glutamate disruptions in nucleus accumbens. *J. Neurosci.* 28, 7184–7192.

Ferré, S. (1997). Adenosine–dopamine interactions in the ventral striatum. Implications for the treatment of schizophrenia. *Psychopharmacology*, 133, 107–120.

Ferré, S., Ciruela, F., Borycz, J., Solinas, M., Quarta, D., Antoniou, K., Quiroz, C., Justinova, Z., Lluis, C., Franco, R., and Goldberg, S. R. (2008). Adenosine A1-A2A receptor heteromers: new targets for caffeine in the brain. *Front. Biosci.* 13, 2391–2399.

Ferré, S., Ciruela, F., Canals, M., Marcellino, D., Burgueno, J., Casado, V., Hillion, J., Torvinen, M., Fanelli, F., Benedetti, P. P., Goldberg, S. R., Bouvier, M., Fuxe, K., Agnati, L. F., Lluis, C., Franco, R., and Woods. A. (2004). Adenosine A2A-dopamine D2 receptor–receptor heteromers. Targets for neuro-psychiatric disorders. *Parkinsonism Relat. Disord.* 10, 265–271.

Ferré, S., Fredholm, B. B., Morelli, M., Popoli, P., and Fuxe, K. (1997). Adenosine–dopamine receptor–receptor interactions as an integrative mechanism in the basal ganglia. *Trends Neurosci.* 20, 482–487.

Ferré, S., Popoli, P., Giménez-Llort, L., Rimondini, R., Müller, C. E.,

Strömberg, I., Ögren, S. O., and Fuxe, K. (2001). Adenosine/dopamine interaction: implications for the treatment of Parkinson's disease. *Parkinsonism Relat. Disord.* 7, 235–241.

Fink, J. S., Weaver, D. R., Rivkees, S. A., Peterfreund, R. A., Pollack, A. E., Adler, E. M., and Reppert, S. M. (1992). Molecular cloning of the rat A_{2A} adenosine receptor: selective co-expression with D_2 dopamine receptors in rat striatum. *Brain Res. Mol. Brain Res.* 14, 186–195.

Floresco, S. B., and Ghods-Sharifi, S. (2007). Amygdala-prefrontal cortical circuitry regulates effort-based decision making. *Cereb. Cortex* 17, 251–260.

Floresco, S. B., St Onge, J. R., Ghods-Sharifi, S., and Winstanley, C. A. (2008a). Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. *Cogn. Affect. Behav. Neurosci.* 8, 375–389.

Floresco, S. B., Tse, M. T., and Ghods-Sharifi, S. (2008b). Dopaminergic and glutamatergic regulation of effort- and delay-based decision making. *Neuropsychopharmacology* 33, 1966–1979.

Foltin, R. W. (1991). An economic analysis of "demand" for food in baboons. *J. Exp. Anal. Behav.* 56, 445–454.

Font, L., Mingote, S., Farrar, A. M., Pereira, M., Worden, L., Stopper, C., Port, R. G., and Salamone, J. D. (2008). Intra-accumbens injections of the adenosine A(2A) agonist CGS 21680 affect effort-related choice behavior in rats. *Psychopharmacology* 199, 515–526.

Fouriezos, G., Bielajew, C., and Pagotto, W. (1990). Task difficulty increases thresholds of rewarding brain stimulation. *Behav. Brain Res.* 37, 1–7.

Fuxe, K., Agnati, L. F., Jacobsen, K., Hillion, J., Canals, M., Torvinen, M., Tinner-Staines, B., Staines, W., Rosin, D., Terasmaa, A., Popoli, P., Leo, G., Vergoni, V., Lluis, C., Ciruela, F., Franco, R., and Ferré, S. (2003). Receptor heteromerization in adenosine A2A receptor signaling: relevance for striatal function and Parkinson's disease. *Neurology* 61, S19–S23.

Gawin, F. H. (1986). Neuroleptic reduction of cocaine-induced paranoia but not euphoria? *Psychopharmacology* 90, 142–143.

Guarraci, F. A., and Kapp, B. S. (1999). An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential pavlovian fear conditioning in the awake rabbit. *Behav. Brain Res.* 99, 169–179.

Hamill, S., Trevitt, J. T., Nowend, K. L., Carlson, B. B., and Salamone, J. D. (1999). Nucleus accumbens dopamine depletions and time-constrained progressive ratio performance: effects of different ratio requirements. *Pharmacol. Biochem. Behav.* 64, 21–27.

Haney, M., Ward, A. S., Foltin, R. W., and Fischman, M. W. (2001). Effects of ecopipam, a selective dopamine D1 antagonist, on smoked cocaine self-administration by humans. *Psychopharmacology* 155, 330–337.

Hauber, W., and Munkel, M. (1997). Motor depressant effects mediated by dopamine D_2 and adenosine A_{2A} receptors in the nucleus accumbens and the caudate-putamen. *Eur. J. Pharmacol.* 323, 127–131.

Hauber, W., Neuscheler, P., Nagel, J., and Muller, C. E. (2001). Catalepsy induced by a blockade of dopamine D_1 or D_2 receptors was reversed by a concomitant blockade of adenosine A_{2A} receptors in the caudate putamen of rats. *Eur. J. Neurosci.* 14, 1287–1293.

Hauber, W., and Sommer, S. (2009). Prefrontostriatal circuitry regulates effort-related decision making. *Cereb. Cortex* [Epub ahead of print]. doi: 10.1093/cercor/bhn241.

Hengeveld, G. M., van Langevelde, F., Groen, T. A., and de Knegt, H. J. (2009). Optimal foraging for multiple resources in several food species. *Am. Nat.* 17, 102–110.

Herrnstein, R. J. (1974). Formal properties of the matching law. *J. Exp. Anal. Behav.* 21, 159–164.

Heyman, G. M., and Monaghan, M. M. (1987). The effect of changes in the response requirement and deprivation on the parameters of the matching law equation: new data and review. *J. Exp. Psych. An. Behav. Proc.* 13, 384–394.

Hillion, J., Canals, M., Torvinen, M., Casado, V., Scott, R., Terasmaa, A., Hansson, A., Watson, S., Olah, M. E., Mallol, J., Canela, E. I., Zoli, M., Agnati, L. F., Ibañez, C. F., Lluis, C., Franco, R., Ferré, S., and Fuxe, K. (2002). Coaggregation, cointernalization, and codesensitization of adenosine A_{2A} receptors and dopamine D_2 receptors. *J. Biol. Chem.* 277, 18091–18097.

Hursh, S. R., Raslear, T. G., Shurtleff, D., Bauman, R., and Simmons, L. (1988). A cost-benefit analysis of demand for food. *J. Exp. Anal. Behav.* 50, 419–440.

Ishiwari, K., Madson, L. J., Farrar, A. M., Mingote, S. M., Valenta, J. P., DiGianvittorio, M. D., Frank, L. E., Correa, M., Hockemeyer, J., Muller, C., and Salamone, J. D. (2007). Injections of the selective adenosine A_{2A} antagonist MSX-3 into the nucleus accumbens core attenuate the locomotor suppression induced by haloperidol in rats. *Behav. Brain Res.* 178, 190–199.

Ishiwari, K., Weber, S. M., Mingote, S., Correa, M., and Salamone, J. D. (2004). Accumbens dopamine and the regulation of effort in food-seeking behavior: modulation of work output by different ratio or force requirements. *Behav. Brain Res.* 151, 83–91.

Jenner, P. (2005). Istradefylline, a novel adenosine A_{2A} receptor antagonist, for the treatment of Parkinson's disease. *Expert Opin. Investig. Drugs* 14, 729–738.

Jensen, J., McIntosh, A. R., Crawley, A. P., Mikulis, D. J., Remington, G., and Kapur, S. (2003). Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron* 40, 1251–1257.

Johnson, D. F., and Collier, G. H. (1987). Caloric regulation and patterns of food choice in a patchy environment: the value and cost of alternative foods. *Physiol. Behav.* 39, 351–359.

Kaufman, L. W. (1980). Foraging costs and meal patterns in ferrets. *Physiol. Behav.* 25, 139–141.

Kaufman, L. W., Collier, G., Hill, W. L., and Collins, K. (1980). Meal cost and meal patterns in an uncaged domestic cat. Physiol. Behav. 25, 135–137.

Kelley, A. E., Baldo, B. A., Pratt, W. E., and Will, M. J. (2005). Corticostriatal-hypothalamic circuitry and food motivation: integration of energy, action and reward. *Physiol. Behav.* 86, 773–795.

Koch, M., Schmid, A., and Schnitzler, H. U. (2000). Role of nucleus accumbens dopamine D1 and D2 receptors in instrumental and Pavlovian paradigms of conditioned reward. *Psychopharmacology* 152, 67–73.

Koob, G. F., Riley, S. J., Smith, S. C., and Robbins, T. W. (1978). Effects of 6-hydroxydopamine lesions of the nucleus accumbens septi and olfactory tubercle on feeding, locomotor activity, and amphetamine anorexia in the rat. *J. Comp. Physiol. Psychol.* 92, 917–927.

Krebs, J. R. (1977). Optimal foraging: theory and experiment. *Nature* 268, 583–584.

Lapish, C. C., Kroener, S., Durstewitz, D., Lavin, A., Seamans, J. K. (2007). The ability of the mesocortical dopamine system to operate in distinct temporal modes. *Psychopharmacology* 191, 609–625.

Lea, S. E. G. (1978). The psychology and economics of demand. *Psychol. Bull.* 85, 441–466.

Levita, L., Hare, T. A., Voss, H. U., Glover, G., Ballon, D. J., and Casey, B. J. (2009). The bivalent side of the nucleus accumbens. *Neuroimage* 44, 1178–1187.

Liberzon, I., Taylor, S. F., Amdur, R., Jung, T. D., Chamberlain, K. R., Minoshima, S., Koeppe, R. A., and Fig, L. M. (1999). Brain activation in PTSD in response to trauma-related stimuli. *Biol. Psychiatry* 45, 817–826.

Marinelli, M., Barrot, M., Simon, H., Oberlander, C., Dekeyne, A., Le Moal, M., and Piazza, P. V. (1998). Pharmacological stimuli decreasing nucleus accumbens dopamine can act as positive reinforcers but have a low addictive potential. *Eur. J. Neurosci.* 10, 3269–3275.

Marinelli, S., Pascucci, T., Bernardi, G., Puglisi-Allegra, S., and Mercuri, N. B. (2005). Activation of TRPV1 in the VTA excites dopaminergic neurons and increases chemical- and noxious-induced dopamine release in the nucleus accumbens. *Neuropsychopharmacology* 30, 864–875.

Martinez, R. C. R., Oliveira, A. R., Macedo, C. E., Molina, V. A., and Brandao, M. L. (2008). Involvement of dopaminergic mechanisms in the nucleus accumbens core and shell subregions in the expression of fear conditioning. *Neurosci. Lett.* 446, 112–116.

McCullough, L. D., and Salamone, J. D. (1992) Anxiogenic drugs beta-CCE and FG 7142 increase extracellular dopamine levels in nucleus accumbens. *Psychopharmacology* 109, 379–382.

McCullough, L. D., Sokolowski, J. D., and Salamone, J. D. (1993). A neurochemical and behavioral investigation of the involvement of nucleus accumbens dopamine in instrumental avoidance. *Neuroscience* 52, 919–925.

Mekarski, J. E. (1988). Main effects of current and pimozide on prepared and learned self-stimulation behaviors are on performance not reward. *Pharmacol. Biochem. Behav.* 31, 845–853.

Mingote, S., Font, L., Farrar, A. M., Vontell, R., Worden, L. T., Stopper, C. M., Port, R. G., Sink, K. S., Bunce, J. G., Chrobak, J. J., and Salamone, J. D. (2008). Nucleus accumbens adenosine A2A receptors regulate exertion of effort by acting on the ventral striatopallidal pathway. *J. Neurosci.* 28, 9037–9046.

Mingote, S., Weber, S. M., Ishiwari, K., Correa, M., and Salamone, J. D. (2005). Ratio and time requirements on operant schedules: effort-related effects of nucleus accumbens dopamine depletions. *Eur. J. Neurosci* 21, 1749–1757.

Mogenson, G. J., Jones, D. L., and Yim, C. Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. *Prog. Neurobiol.* 14, 69–97.

Morelli, M., and Pinna, A. (2002). Interaction between dopamine and adenosine A_{2A} receptors as a basis for the treatment of Parkinson's disease. *Neurol. Sci.* 22, 71–72.

Mott, A. M, Nunes, E. J., Collins, L. E., Port, R. G., Sink, K. S., Hockemeyer, J., Müller, C. E., and Salamone, J. D. (2009). The adenosine A_{2A} antagonist MSX-3 reverses the effects of the dopamine antagonist haloperidol on effort-related decision making in a T-maze cost/benefit procedure. *Psychopharmacology* 204, 103–112.

Munro, L. J., and Kokkinidis, L. (1997). Infusion of quinpirole and muscimol into the ventral tegmental area inhibits fear-potentiated startle: implications for the role of dopamine in fear expression. *Brain Res.* 746, 231–238.

Nann-Vernotica, E., Donny, E. C., Bigelow, G. E., and Walsh, S. L. (2001). Repeated administration of the D1/5 antagonist ecopipam fails to attenuate the subjective effects of cocaine. *Psychopharmacology* 155, 338–347.

Neill, D. B., and Justice, J. B. (1981). An hypothesis for a behavioral function of dopaminergic transmission in nucleus accumbens. In The Neurobiology of Nucleus Accumbens, R. B. Chronister and J. F. Defrance, eds (Brunswick, Huer Institute).

Nicola, S. M. (2007). The nucleus accumbens as part of a basal ganglia action selection circuit. *Psychopharmacology* 191, 521–550.

Niv, Y., Daw, N. D., Joel, D., and Dayan, P. (2007). Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology* 191, 507–520.

Nowend, K. L., Arizzi, M., Carlson, B. B., and Salamone, J. D. (2001). D_1 or D_2 antagonism in nucleus accumbens core or dorsomedial shell suppresses lever pressing for food but leads to compensatory increases in chow consumption. *Pharmacol. Biochem. Behav.* 69, 373–382.

Nunes, E. J., Farrar, A. M., Font, L., Patel, V., Schlar, D. E., Collins, L. E., Sager, T. N., Correa, M., and Salamone, J. D. (2009). Differential Actions of Adenosine A1 and A2A Antagonists on the Effort-related Effects of Dopamine Antagonists. Neuroscience Meeting Planner. Washington, DC, Society for Neuroscience (online).

Palmiter, R. D. (2007). Is dopamine a physiologically relevant mediator of feeding behavior? *Trends Neurosci.* 30, 375–381.

Pavic, L. (2003). Alterations in brain activation in posttraumatic stress disorder patients with severe hyperarousal symptoms and impulsive aggressiveness. *Eur. Arch. Psychiatry Clin. Neurosci.* 253, 80–83.

Pezze, M. A., and Feldon, J. (2004). Mesolimbic dopaminergic pathways in fear conditioning. *Prog. Neurobiol.* 74, 301–320.

Phan, K. L., Taylor, S. F., Welsh, R. C., Ho, S. H., Britton, J. C., and Liberzon, I. (2004). Neural correlates of individual ratings of emotional salience: a trial-related fMRI study. *Neuroimage* 21, 768–780.

Phillips, P. E., Walton, M. E., and Jhou, T. C. (2007). Calculating utility: preclinical evidence for cost-benefit analysis by mesolimbic dopamine. *Psychopharmacology* 191, 483–495.

Pinna, A., Wardas, J., Simola, N., and Morelli, M. (2005). New therapies for the treatment of Parkinson's disease: adenosine A_{2A} receptor antagonists. *Life Sci.* 77, 3259–3267.

Pruessner, J. C., Champagne, F., Meaney, M. J., and Dagher, A. (2004). Dopamine release in response to a psychological stress in humans and its relationship to early life maternal care: a positron emission tomography study using [11C]raclopride. *J. Neurosci.* 24, 2825–2831.

Rachlin, H. (2003). Economic concepts in the behavioral study of addiction. In Choice, Behavioral Economics and Addiction, R. E. Vuchinich and N. Heather, eds (Oxford, Elsevier), pp. 129–149.

Redgrave, P., Gurney, K., and Reynolds, J. (2008). What is reinforced by phasic dopamine signals? *Brain Res. Rev.* 58, 322–339.

Richardson, N. R., and Roberts, D. C. (1996). Progressive ratio schedules in drug self-administration studies in rats: a method to evaluate reinforcing efficacy. *J. Neurosci. Methods* 66, 1–11.

Rick, J. H., Horvitz, J. C., and Balsam, P. D. (2006). Dopamine receptor blockade and extinction differentially affect behavioral variability. *Behav. Neurosci.* 120, 488–492.

Robbins, T. W., and Everitt, B. J. (2007). A role for mesencephalic dopamine in activation: commentary on Berridge (2006). *Psychopharmacology* 191, 433–437.

Robbins, T. W., and Koob, G. F. (1980). Selective disruption of displacement behaviour by lesions of the mesolimbic dopamine system. *Nature* 285, 409–412.

Roitman, M. F., Stuber, G. D., Phillips, P. E., Wightman, R. M., and Carelli, R. M. (2004). Dopamine operates as a subsecond modulator of food seeking. *J. Neurosci.* 24, 1265–1271.

Salamone, J. D. (1986). Different effects of haloperidol and extinction on instrumental behaviours. *Psychopharmacology* 88, 18–23.

Salamone, J. D. (1987). The actions of neuroleptic drugs on appetitive instrumental behaviors. In Handbook of Psychopharmacology, L. L. Iversen, S. D. Iversen and S. H. Snyder, eds (New York, Plenum Press), pp. 575–608.

Salamone, J. D. (1988). Dopaminergic involvement in activational aspects of motivation: effects of haloperidol on schedule induced activity, feeding and foraging in rats. *Psychobiology* 16, 196–206.

Salamone, J. D. (1994). The involvement of nucleus accumbens dopamine in appetitive and aversive motivation. *Behav. Brain Res.* 61, 117–133.

Salamone, J. D. (1996). The behavioral neurochemistry of motivation: methodological and conceptual issues in studies of the dynamic activity of nucleus accumbens dopamine. *J. Neurosci. Methods* 64, 137–149.

Salamone, J. D. (in press). Involvement of nucleus accumbens dopamine in behavioral activation and effort-related functions. In Dopamine Handbook, L. L. Iversen, S. D. Iversen, S. B., Dunnett and A. Bjorkland, eds (Oxford, Oxford University Press).

Salamone, J. D. (2006). Will the last person who uses the term 'reward' please turn out the lights? Comments on processes related to reinforcement, learning, motivation and effort. *Addict. Biol.* 11, 43–44.

Salamone, J. D., Aberman, J. E., Sokolowski, J. D., and Cousins, M. S. (1999). Nucleus accumbens dopamine and rate of responding: neurochemical and behavioral studies. *Psychobiology* 27, 236–247.

Salamone, J. D., Arizzi, M., Sandoval, M. D., Cervone, K. M., and Aberman, J. E. (2002). Dopamine antagonists alter response allocation but do not suppress appetite for food in rats: contrast between the effects of SKF 83566, raclopride and fenfluramine on a concurrent choice task. *Psychopharmacology* 160, 371–380.

Salamone, J. D., Betz, A. J., Ishiwari, K., Felsted, J., Madson, L., Mirante, B., Clark, K., Font, L., Korbey, S., Sager, T. N., Hockemeyer, J., and Muller, C. E. (2008a). Tremorolytic effects of adenosine A2A antagonists: implications for parkinsonism. *Front. Biosci.* 13, 3594–3605.

Salamone, J. D., Ishiwari, K., Betz, A. J., Farrar, A. M., Mingote, S. M., Font, L., Hockemeyer, J., Müller, C. E., and Correa, M. (2008b). Dopamine/adenosine interactions related to locomotion and tremor in animal models: possible relevance to parkinsonism. *Parkinsonism Relat. Disord.* 14, S130–S134.

Salamone, J. D., and Correa, M. (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav. Brain Res.* 137, 3–25.

Salamone, J. D., Correa, M., Farrar, A., and Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology* 191, 461–482.

Salamone, J. D., Correa, M., Mingote, S., and Weber, S. M. (2003). Nucleus accumbens dopamine and the regulation of effort in food-seeking behavior: implications for studies of natural motivation, psychiatry, and drug abuse. *J. Pharmacol. Exp. Ther.* 305, 1–8.

Salamone, J. D., Correa, M., Mingote, S. M., and Weber, S. M. (2005). Beyond the reward hypothesis: alternative functions of nucleus accumbens dopamine. *Curr. Opin. Pharmacol.* 5, 34–41.

Salamone, J. D., Correa, M., Mingote, S. M., Weber, S. M., and Farrar, A. M. (2006). Nucleus accumbens dopamine and the forebrain circuitry involved in behavioral activation and effort-related decision making: implications for understanding anergia and psychomotor slowing in depression. *Curr. Psychiatry Rev.* 2, 267–280.

Salamone, J. D., Cousins, M. S., and Bucher, S. (1994). Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure. *Behav. Brain Res.* 65, 221–229.

Salamone, J. D., Cousins, M. S., Maio, C., Champion, M., Turski, T., and Kovach, J. (1996). Different behavioral effects of haloperidol, clozapine and thioridazine in a concurrent lever pressing and feeding procedure. *Psychopharmacology* 125, 105–112.

Salamone, J. D., Cousins, M. S., and Snyder, B. J. (1997). Behavioral functions of nucleus accumbens dopamine: empirical and conceptual problems with the anhedonia hypothesis. *Neurosci. Biobehav. Rev.* 21, 341–359.

Salamone, J. D., Farrar, A. M., Font, L., Patel, V., Schlar, D. E., Nunes, E. J., Collins, L. E., and Sager, T. N. (2009). Differential actions of adenosine A1 and A2A antagonists on the effort-related effects of dopamine D2 antagonism. *Behav. Brain Res.* 201, 216–222.

Salamone, J. D., Kurth, P., McCullough, L. D., and

Sokolowski, J. D. (1995). The effects of nucleus accumbens dopamine depletions on continuously reinforced operant responding: contrasts with the effects of extinction. *Pharmacol. Biochem. Behav.* 50, 437–443.

Salamone, J. D., Mahan, K., and Rogers, S. (1993a). Ventrolateral striatal dopamine depletions impair feeding and food handling in rats. *Pharmacol. Biochem. Behav.* 44, 605–610.

Salamone, J. D., Kurth, P. A., McCullough, L. D., Sokolowski, J. D., and Cousins, M. S. (1993b). The role of brain dopamine in response initiation: effects of haloperidol and regionally specific dopamine depletions on the local rate of instrumental responding. *Brain Res.* 628, 218–226.

Salamone, J. D., Steinpreis, R. E., McCullough, L. D., Smith, P., Grebel, D., and Mahan, K. (1991). Haloperidol and nucleus accumbens dopamine depletion suppress lever pressing for food but increase free food consumption in a novel food choice procedure. *Psychopharmacology* 104, 515–521.

Salamone, J. D., Wisniecki, A., Carlson, B. B., and Correa, M. (2001). Nucleus accumbens dopamine depletions make animals highly sensitive to high fixed ratio requirements but do not impair primary food reinforcement. *Neuroscience* 105, 863–870.

Sanchis-Segura, C., and Spanagel, R. (2006). Behavioural assessment of drug reinforcement and addictive features in rodents: an overview. *Addict. Biol.* 11, 2–38.

Sarchiapone, M., Carli, V., Camardese, G., Cuomo, C., Di Guida, D., Calgagni, M. L., Focacci, C., and De Riso, S. (2006). Dopamine transporter binding in depressed patients with anhedonia. *Psychiatry Res. Neuroimag.* 147, 243–248.

Schiffmann, S. N., Jacobs, O., and Vanderhaeghen, J. J. (1991). Striatal restricted adenosine A_{2A} receptor (RDC8) is expressed by enkephalin but not by substance P neurons: an in situ hybridization histochemistry study. *J. Neurochem.* 57, 1062–1071.

Schmelzeis, M. C., and Mittleman, G. (1996). The hippocampus and reward: effects of hippocampal lesions on progressive-ratio responding. *Behav. Neurosci.* 110, 1049–1066.

Schultz, W. (2007a). Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* 30, 259–288.

Schultz, W. (2007b). Behavioral dopamine signals. *Trends Neurosci.* 30, 203–210.

Schweimer, J., and Hauber, W. (2005). Involvement of the rat anterior cingulate cortex in control of instrumental responses guided by reward expectancy. *Learn. Mem.* 12, 334–342.

Sink, K. S., Vemuri, V. K., Olszewska, T., Makriyannis, A., and Salamone, J. D. (2008). Cannabinoid CB1 antagonists and dopamine antagonists produce different effects on a task involving response allocation and effort-related choice in food-seeking behavior. *Psychopharmacology* 196, 565–574.

Skjoldager, P., Pierre, P. J., and Mittlman, G. (1993). Reinforcer magnitude and progressive ratio responding: effects of increased effort, prefeeding and extinction. *Learn. Motiv.* 24, 303–343.

Sokolowski, J. D., and Salamone, J. D. (1998). The role of accumbens dopamine in lever pressing and response allocation: effects of 6-OHDA injected into core and dorsomedial shell. *Pharmacol. Biochem. Behav.* 59, 557–566.

Staddon, J. E. R. (1979). Operant behavior as adaptation to constraint. *J. Exp. Psychol. Gen.* 108, 48–67.

Stahl, S. M. (2002). The psychopharmacology of energy and fatigue. *J. Clin. Psychiatry* 63, 7–8.

Stewart, W. J. (1974). Progressive reinforcement schedules: a review and evaluation. *Aust. J. Psychol.* 27, 9–22.

Svenningsson, P., Le Moine, C., Fisone, G., and Fredholm, B. B. (1999). Distribution, biochemistry and function of striatal adenosine A_{2A} receptors. *Prog. Neurobiol.* 59, 355–396.

Szczypka, M. S., Kwok, K., Brot, M. D., Marck, B. T., Matsumoto, A. M., Donahue, B. A., and Palmiter, R. D. (2001). Dopamine production in the caudate putamen restores feeding in dopamine-deficient mice. *Neuron* 30, 819–828.

Takahashi, R. N., Pamplona, F. A., and Prediger, R. D. (2008). Adenosine receptor antagonists for cognitive dysfunction: a review of animal studies. *Front. Biosci.* 13, 2614–2632.

Ungerstedt, U. (1971). Adipsia and aphagia after 6-hydroxydopamine induced degeneration of the nigrostriatal dopamine system. *Acta Physiol. Scand. Suppl.* 367, 95–122.

van den Bos, R., van der Harst, J., Jonkman, S., Schilders, M., and Spruijt, B. (2006). Rats assess costs and benefits according to an internal standard. *Behav. Brain Res.* 171, 350–354.

Vezina, P., Lorrain, D. S., Arnold, G. M., Austin, J. D., and Suto, N. (2002). Sensitization of midbrain dopamine neuron reactivity promotes the pursuit of amphetamine. *J. Neurosci.* 22, 4654–4662.

Vuchinich, R. E., and Heather, N. (2003). Introduction: Overview of behavioural economic perspectives on substance use and addiction. In Choice, Behavioral Economics and Addiction, R. E. Vuchinich and N. Heather, eds (Oxford, Elsevier), pp. 1–31.

Wachtel, S. R., Ortengren, A., and de Wit, H. (2002). The effects of acute haloperidol or risperidone on subjective responses to methamphetamine in healthy volunteers. *Drug Alcohol Depend.* 68, 23–33.

Wacker, J., Dillon, D. G., and Pizzagalli, D. A. (2009). The role of the nucleus accumbens and rostral anterior cingulate cortex in anhedonia: integration of resting EEG, fMRI, and volumetric techniques. *Neuroimage* 46, 327–337.

Wakabayashi, K. T., Fields, H. L., and Nicola, S. M. (2004). Dissociation of the role of nucleus accumbens dopamine in responding to reward-predictive cues and waiting for reward. *Behav. Brain Res.* 154, 19–30.

Wallace, M., Singer, G., Finlay, J., and Gibson, S. (1983). The effect of 6-OHDA lesions of the nucleus accumbens septum on schedule-induced drinking, wheelrunning and corticosterone levels in the rat. *Pharmacol. Biochem. Behav.* 18, 129–136.

Walton, M. E., Bannerman, D. M., Alterescu, K., and Rushworth, M. F. S. (2003). Functional specialization within medial frontal cortex of the anterior cingulated for evaluating effort-related decisions. *J. Neurosci.* 23, 6475–6479.

Walton, M. E., Kennerley, S. W., Bannerman, D. M., Phillips, P. E., and Rushworth, M. F. (2006). Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. *Neural Netw.* 19, 1302–1314.

Wardas, J., Konieczny, J., and Lorenc-Koci, E. (2001). SCH 58261, an A_{2A} adenosine receptor antagonist, counteracts parkinsonian-like muscle rigidity in rats. *Synapse* 41, 160–171.

Williams, B. A. (1988). Reinforcement, choice, and response strength. In Stevens' Handbook of Experimental Psychology Vol. 2, R. C. Atkinson, R. J. Herrnstein, G. Lindsey and R. D. Luce, eds (New York, John Wiley and Sons), pp. 167–174.

Wise, R. A. (2004). Dopamine, learning and motivation. *Nat. Rev. Neurosci.* 5, 483–494.

Woolverton, W. L., and Ranaldi, R. (2002). Comparison of the reinforcing efficacy of two dopamine D2-like receptor agonists in rhesus monkeys using a progressive-ratio schedule of reinforcement. *Pharmacol. Biochem. Behav.* 72, 803–809.

Worden, L. T., Shahriari, M., Farrar, A. M., Sink, K. S., Hockemeyer. J., Müller, C., Salamone, JD. (2009). The adenosine A_{2A} antagonist MSX-3 reverses the effort-related effects of dopamine blockade: differential interaction with D1 and D2 family antagonists. *Psychopharmacology* 203, 489–499.

Yin, H. H., Ostlund, S. B., and Balleine, B. W. (2008). Reward-guided learning beyond dopamine in the nucleus accumbens: the integrative functions of cortico-basal ganglia networks. *Eur. J. Neurosci.* 28, 1437–1448.

Young, A. M. (2004). Increased extracellular dopamine in nucleus accumbens in response to unconditioned and conditioned aversive stimuli: studies using 1 min microdialysis in rats. *J. Neurosci. Methods* 138, 57–63.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 June 2009; paper pending published: 17 June 2009; accepted: 21 July 2009; published online: 07 September 2009.

*Citation: Salamone JD, Correa M, Farrar AM, Nunes EJ and Pardo M (2009) Dopamine, behavioral economics, and effort. Front. Behav. Neurosci. **3**:13. doi: 10.3389/neuro.08.013.2009*

frontiers in
BEHAVIORAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 01 October 2009
doi: 10.3389/neuro.08.033.2009

Avoiding negative outcomes: tracking the mechanisms of avoidance learning in humans during fear conditioning

Mauricio R. Delgado[1]*, Rita L. Jou[2], Joseph E. LeDoux[3] and Elizabeth A. Phelps[2,3]

[1] *Department of Psychology, Rutgers University, Newark, NJ, USA*
[2] *Department of Psychology, New York University, New York, NY, USA*
[3] *Center for Neural Science, New York University, New York, NY, USA*

Edited by:
Daeyeol Lee, Yale University School of Medicine, USA

Reviewed by:
Jeansok J. Kim, University of Washington, USA
Hackjin Kim, Korea University, Korea

***Correspondence:**
Mauricio R. Delgado, Department of Psychology, Rutgers University, 340 Smith Hall, 101 Warren Street, Newark, NJ 07102, USA.
e-mail: delgado@psychology.rutgers.edu

Previous research across species has shown that the amygdala is critical for learning about aversive outcomes, while the striatum is involved in reward-related processing. Less is known, however, about the role of the amygdala and the striatum in learning how to exert control over emotions and avoid negative outcomes. One potential mechanism for active avoidance of stressful situations is postulated to involve amygdala–striatal interactions. The goal of this study was to investigate the physiological and neural correlates underlying avoidance learning in humans. Specifically, we used a classical conditioning paradigm where three different conditioned stimuli (CS) were presented. One stimulus predicted the delivery of a shock upon stimulus offset (CS+), while another predicted no negative consequences (CS–). A third conditioned cue also predicted delivery of a shock, but participants were instructed that upon seeing this stimulus, they could avoid the shock if they chose the correct action (AV+). After successful learning, participants could then easily terminate the shock during subsequent stimulus presentations (AV–). Physiological responses (as measured by skin conductance responses) confirmed a main effect of conditioning, particularly showing higher arousal responses during pre (AV+) compared to post (AV–) learning of an avoidance response. Consistent with animal models, amygdala–striatal interactions were observed to underlie the acquisition of an avoidance response. These results support a mechanism of active coping with conditioned fear that allows for the control over emotional responses such as fears that can become maladaptive and influence our decision-making.

Keywords: amygdala, striatum, negative reinforcement, instrumental conditioning, fMRI, reward, punishment

INTRODUCTION

The ability to modify and control our emotional responses is critical for adaptive function and goal-directed behavior. Although learning to fear a potentially dangerous situation is important, it is equally important to be able to modify this fear when new information is available, or use this fear to motivate adaptive action that diminishes the potential threat. Recent research examining the neural systems of regulating fear in humans has highlighted passive extinction techniques (Milad and Quirk, 2002; Knight et al., 2004; Phelps et al., 2004) and the use of cognitive strategies (Kalisch et al., 2005; Ochsner and Gross, 2005; Delgado et al., 2008b). These techniques focus on modifying the fear response in the presence of the fear-eliciting event. Another common response used to regulate fear, however, is to take an action to avoid the potential danger and diminish the fear response. Given how frequently action is used to cope with potential threat and fear outside the laboratory, surprisingly little research conducted in humans has examined the neural system mediating the active coping of fear. Research in non-human animals has suggested that active coping of fear may involve amygdala and striatal interactions (Killcross et al., 1997; Everitt et al., 1999; LeDoux and Gorman, 2001; Cardinal et al., 2002). The goal of the present study is to investigate if an amygdala–striatal circuitry underlies active coping of fear in humans.

The amygdala has been the focus of investigations of aversive learning, particularly Pavlovian fear conditioning paradigms in which fear is expressed passively, such as through autonomic responses (for review see Phelps and LeDoux, 2005). In contrast, the human striatum has been highlighted in investigations of reward-related processing, such as instrumental paradigms, that involve decision-making and action-contingencies (for review see Montague and Berns, 2002; O'Doherty, 2004; Knutson and Cooper, 2005; Delgado, 2007; Rangel et al., 2008). The human striatum, however, is also implicated in aversive learning (e.g., Jensen et al., 2003; Seymour et al., 2004, 2007; Menon et al., 2007; Delgado et al., 2008a). Studies from non-human animals have led to the hypothesis that one role for the striatum in aversive learning may be to aid in the acquisition of avoidance actions that diminish exposure to a fear-eliciting event (LeDoux and Gorman, 2001). For example, an investigation in rodents examining the amygdala subnuclei that mediate fear-motivated action found that the basal nucleus, which projects to the striatum (Mogenson et al., 1980; Robbins et al., 1989), is necessary for the acquisition of a fear-reducing action, but is not necessary for more passive expressions of conditioned fear (Amorapanth et al., 2000). In contrast, the central nucleus, which projects to the brainstem and hypothalamus, is necessary for the passive expression of fear, but not for learning a fear-reducing action. These results suggest that

partially independent neural circuits mediate active and passive means of fear expression and that the amygdala's connectivity with the striatum allows for active coping strategies to develop and diminish fear induced by a conditioned stimulus (Everitt et al., 1991; Amorapanth et al., 2000; Cain and LeDoux, 2008). In support of this hypothesis, both dorsal and ventral striatum in rats have been implicated in various types of avoidance learning (Winocur and Mills, 1969; Allen and Davison, 1973; Neill et al., 1974; McCullough et al., 1993; Li et al., 2004).

In humans, neuroimaging experiments suggest that the striatum is involved in the expectation of an aversive stimulus, whether an opportunity to avoid exists or not (Jensen et al., 2003; Delgado et al., 2008a). However, less is known about the potential striatal–amygdala interactions that may underlie avoidance learning in the human brain. In this experiment, we used a modified aversive conditioning paradigm in conjunction with blood oxygenated level dependent (BOLD) and autonomic measures to explore the acquisition of an avoidance learning response. Participants were instructed that they could: (a) avoid a potential shock by learning a behavioral response (i.e., a button press), and (b) express the behavioral response after successful learning to prevent future shock delivery. As suggested by animal models (e.g., LeDoux and Gorman, 2001), we hypothesized that interactions of the amygdala and striatum would underlie a measure of successful avoidance learning, comparing BOLD responses pre- and post-learning.

MATERIALS AND METHODS

PARTICIPANTS

Thirty-two participants were initially recruited for this study (19F/13M, $M = 19.8$, SD = 2.2). Nine participants were excluded from further analysis due to excessive motion during scanning ($N = 4$, more than 2 mm of movement), failure to learn the task ($N = 3$) or equipment malfunction during session ($N = 2$, shocks not delivered). The final behavioral and neuroimaging analyses were conducted on 23 participants (15F/8M, mean age = 19.9, SD = 2.6). Participants responded to posted advertisement and all participants gave informed consent. The experiment was approved by the University Committee on Activities Involving Human Subjects at New York University.

PROCEDURE

The experiment consisted of an aversive conditioning paradigm with instruction. Participants were presented with three types of colored squares (e.g., blue, yellow, green) that served as conditioned stimuli (CS). Two of the CSs were fully predictable and led to the delivery of a mild shock to the wrist (the unconditioned stimulus, US) with either 100% (CS+ trials) or 0% (CS– trials) probability (certain trials). The third CS also predicted delivery of a mild shock, however participants were afforded a chance to avoid the shock if they learned the appropriate response (avoidable or AV trials). The AV trials were further subdivided into two types of trials according to learning success. During pre-learning trials (AV+ trials) participants attempted to learn how to avoid the negative outcome. Post-learning trials included subsequent presentations of the CS where successful avoidance of the US was maintained by the previously learned response (AV– trials). Thus, the four conditions (CS+, CS–, AV+, AV–) comprised two 2 classes of CS (certain and avoidable) that varied with respect to conditioned response (aversive and safe; **Figure 1**).

Each CS presentation lasted 10 s and was broken down into a CS and a response phase. During the CS phase (4–6 s), participants were presented with the type of CS and instructed to just observe and wait for the response phase. The CS phase was the task period of interest and measures of physiological and BOLD responses reflect activation at this time point, uncontaminated by any motor responses or shock delivery. The response phase was cued by a question-mark in the middle of the colored square (4–6 s). At this time, participants were instructed to make a behavioral response (i.e., a button press). A mild shock was delivered during CS+ and AV+ trials for 200 ms that co-terminated with the end of the response phase. The trial ended with a 14-s inter-trial interval, for a total trial time of 24 s. Each session contained 24 trials, evenly divided into 6 trials for each type of condition (CS+, CS–, AV+, AV–).

For AV trials, participants had a chance to avoid the shock with the correct response during the response phase. Specifically, they were told that one of eight button presses could terminate the shock delivery. Participants were given an MRI compatible button box with four buttons and used their right hand to make one response per trial. They were further instructed that the "correct button" could be the first or second time a button was pressed, thus creating eight possible correct buttons and diminishing excessive motor coordination issues associated with the use of multiple button boxes. Participants were also asked to make a non-contingent button press during the response phase for certain trials (CS+, CS–) to control for motor requirements.

Prior to scanning, participants were instructed what each CS predicted (certain or avoidable outcome). Unbeknownst to participants, however, the correct button press during AV trials was always the

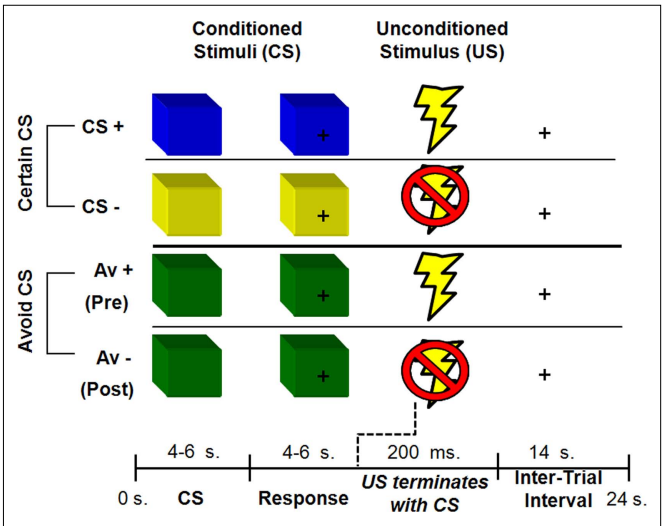


FIGURE 1 | Human avoidance paradigm. Participants were presented with three types of CS. Both CS+ and CS– predicted a certain outcome (an aversive shock or no shock respectively). The AV+ condition predicted a potential shock but afforded the participant an opportunity to avoid a shock with the correct behavioral response. An AV– trial referred to trials post-learning of an avoidance response. Colors were counterbalanced across scanning sessions.

response made in the sixth AV+ trial, irrespective of which button was pressed. The correct button press was then repeated post-learning, during the remaining six AV− trials. This ensured that all participants experienced the same schedule of reinforcement, with each session containing 24 trials, evenly divided into six trials for each type of condition (CS+, CS−, AV+, AV−). Participants who failed to learn the contingency (i.e., failed to repeat the correct button, and thus never experiencing AV− trials), typically reported not paying attention, and were excluded from all further analysis ($N = 3$).

The US constituted mild shocks delivered to the right wrist through a stimulating bar electrode connected to a Grass Medical Instruments stimulator. The stimulator was shielded for magnetic interference and grounded through an RF filter. Participants used a work-up procedure to set the appropriate shock level prior to the experimental session. Specifically, participants experienced a mild shock (10 V, 200 ms, 50 pulses/s) which was gradually increased up to a fixed maximum (60 V). They were instructed to set a level that was deemed uncomfortable, but not painful (mean shock level = 25.69 V, SD = 8.91).

Task events were programmed using E-PRIME software, v1.0 (PST, Pittsburgh, PA, USA). The color of the CSs was counterbalanced across sessions. Stimuli were presented in a black background and projected onto a screen which was visible inside the scanner through a mirror in the head coil. Right handed responses were made using an MRI compatible button box. At the end of the experimental session, participants were debriefed and compensated.

PHYSIOLOGICAL SET-UP, ASSESSMENT AND BEHAVIORAL ANALYSIS

Skin conductance responses (SCRs) were acquired from the participant's middle phalanges of the second and third fingers in the left hand and amplified by BIOPAC Systems skin conductance module. Shielded Ag–AgCl electrodes were grounded through an RF filter panel and served to acquire data. AcqKnowledge software was used to analyze the analog skin conductance waveforms. The level of SCR response was assessed for each trial as the base to peak amplitude difference in skin conductance of the largest deflection in the 0.5–4.5 s latency window after onset of the CS (see LaBar et al., 1995). A minimum response criterion of 0.02 μS was used with lower responses scored as 0. Raw scores were square-root transformed prior to statistical analysis to normalize the distributions (LaBar et al., 1998). Acquired SCRs during the CS phase were then averaged per participant and per type of trial (CS+, CS−, AV+, AV−). A 2×2 repeated measures ANOVA with participants as a random factor was used to test for a main effect of type of CS (certain, avoidable) and conditioned response (aversive, safe). *Post hoc* paired *t*-tests were then conducted to probe differences between the contrast of interest, AV+ and AV− trials.

Additional behavioral data was acquired in the form of reaction time during the response phase, using an MRI compatible button box with four buttons. The primary analysis of the reaction time data was a paired *t*-test comparison of certain (CS+, CS−) and AV trials (AV+, AV−), hypothesized to differ with respect to motivation. Since the schedule of reinforcement was predetermined to reflect learning after six trials, accuracy differences were not expected, and participants who did not learn were excluded as previously described.

FMRI ACQUISITION AND ANALYSIS

A 3T Siemens Allegra head-only scanner and a Siemens standard head coil were used for data acquisition at NYU's Center for Brain Imaging. Anatomical images were acquired using a T1-weighted protocol (256×256 matrix, 176 1-mm sagittal slices) Functional images were acquired using a single-shot gradient echo EPI sequence (TR = 2000 ms, TE = 20 ms, FOV = 192 cm, flip angle = 75°, bandwidth = 4340 Hz/px, echo spacing = 0.29 ms). Thirty-five contiguous oblique-axial slices (3 mm × 3 mm × 3 mm voxels) parallel to the AC–PC line were obtained. Analysis of imaging data was conducted using Brain Voyager software (Brain Innovation, Maastricht, The Netherlands). The data was initially corrected for motion (using a threshold of 2 mm or less), and slice scan time using sinc interpolation was applied. Further, spatial smoothing was performed using a three-dimensional Gaussian filter (4-mm FWHM), along with voxel-wise linear detrending and high-pass filtering of frequencies (three cycles per time course). Structural and functional data of each participant was then transformed to standard Talairach stereotaxic space (Talairach and Tournoux, 1988).

A random-effects analysis was performed on the functional data using a general linear model (GLM) on 23 participants. There were four regressors of interest in the CS phase (CS+, CS−, AV+, AV−). There were also six regressors of no interest that modeled the response phase (separated into four types of trials according to condition) and the shock delivery (CS+_US and AV+_US). The principal contrast served to identify regions of interest (ROIs) involved in processing anticipated aversive outcomes during the CS phase, using a conservative threshold of FDR <0.01 along with a cluster threshold of 10 contiguous voxels. Specifically, trials where an aversive outcome was expected (AV+, CS+) were compared to the most non-aversive control condition (CS−), as some residue conditioned fear could exist in AV− trials. Given the *a priori* hypothesis with respect to amygdala–striatal interactions, an amygdala ROI was functionally defined with this contrast using a more lenient threshold of $p < 0.025$ (uncorrected) along with a cluster threshold of four contiguous voxels (Buchel et al., 1998; LaBar et al., 1998). Mean parameter estimates reflecting effect size of a particular condition were extracted from ROIs in the striatum and amygdala for further analysis. A correlation analysis was then conducted comparing learning changes differences (i.e., mean parameter estimates for AV+ minus mean parameter estimates for AV−) between the functionally defined amygdala and striatum ROIs. Additionally, the time course of activation across the entire functional run for each individual participant was extracted from the amygdala ROI and used in an exploratory connectivity analysis. The time-course data was z-transformed and used as a single predictor in a GLM. The resulting activation map was thresholded at FDR <0.001 and identified regions which hemodynamic patterns correlated with the seed amygdala ROI. Finally, an exploratory analysis was performed comparing AV+ and AV− trials, investigating brain regions associated with avoidance learning changes, and identified ROIs at $p < 0.005$ with four or more contiguous voxels.

RESULTS

BEHAVIORAL AND PHYSIOLOGICAL RESULTS

SCRs were acquired during the CS phase as a measure of physiological arousal. A main effect of conditioned response was observed [$F(1, 21) = 42.34$, $p < 0.0001$; **Figure 2**] suggesting that participants

were more aroused during presentation of CS that predicted aversive (CS+, AV+) compared to safer outcomes (CS–, AV–) consistent with earlier studies of human aversive conditioning (LaBar et al., 1998; Phelps et al., 2004). *Post hoc* paired *t*-tests revealed a differential response between CS+ and CS– trials [$t(22) = 4.71$, $p < 0.0002$] as expected based on previous findings. Importantly, a difference was also observed between AV+ and AV– [$t(22) = 4.73$, $p < 0.0002$], suggesting that learning an avoidable response in this paradigm decreases previously documented conditioned fear responses. No main effect of type of CS [$F(1, 21) = 0.3$, $p = 0.59$] or interaction [$F(2, 21) = 1.3$, $p = 0.27$] was observed using SCRs. Instead, differences across type of CS were seen in the reaction time data in the response phase. Specifically, participants were faster to make a response to avoidable ($M = 593.55$, SD = 161) compared to certain ($M = 671.93$, SD = 146.47) trials [$F(1, 21) = 12.24$, $p < 0.002$], suggesting behavioral reactions were faster when motivation to respond was high (Delgado et al., 2004). These effects were particularly strong when comparing safe trials [AV– and CS–; $t(22) = 4.24$, $p < 0.0005$], but also approaching significance when contrasting the aversive trials [AV+ and CS+; $t(22) = 1.5$, $p = 0.14$]. No effects of conditioning [$F(1, 21) = 1.06$, $p = 0.31$] or interaction [$F(1,21) = 0.31$, $p = 0.58$] were observed in the reaction time data.

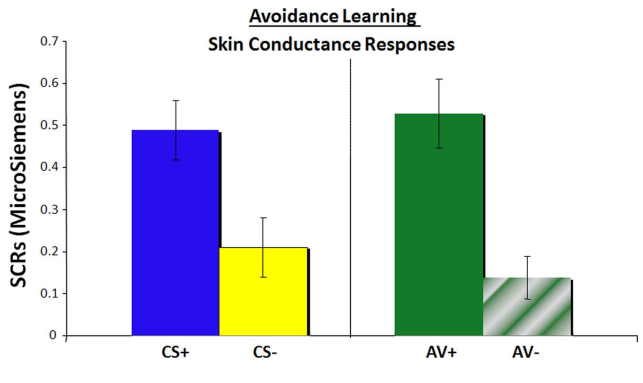


FIGURE 2 | Physiological measure of conditioned fear, as assessed by skin conductance responses, representing arousal responses during the CS phase for all four types of trials.

NEUROIMAGING RESULTS

The main analysis used to identify ROIs involved a contrast of aversive trials, where an aversive outcome was expected (AV+, CS+), and the most non-aversive control condition (CS–). This contrast yielded positive activation patterns in an array of cortical regions (**Table 1**), but central to this investigation, we observed activation in the ventral and dorsal striatum bilaterally. Mean parameter estimates for individual participants were then extracted for the striatum ROIs for further analysis. Within the left ventral striatum ROI identified in this contrast (**Figure 3A**), which included the putamen, a differential response between AV+ and AV– trials was found using *post hoc t*-tests [$t(22) = 3.25$, $p < 0.005$]. This pattern also characterized BOLD responses in the right dorsal striatum ROI [**Figure 3B**; $t(22) = 2.37$, $p < 0.05$]. Interestingly, a differential response between AV+ and CS+ trials was seen in the left ventral striatum ROI [$t(22) = 2.22$, $p < 0.05$], but not within the voxels defined in the right dorsal striatum ROI [$t(22) = 0.94$, $p = 0.36$].

An amygdala ROI was functionally defined with the same contrast of aversive and safe trials, but using a more lenient threshold given the *a priori* hypothesis with respect to amygdala–striatal interactions.

Table 1 | Contrast of aversive (CS+ and AV+) and Safe (CS–) trials at FDR <0.01 and contiguity threshold of 10 voxels.

Region of activation	Brodmann areas	Laterality	Talairach coordinates			#Voxels
			x	*y*	*z*	
Medial frontal gyrus	6	Left	0	–1	52	3965
Superior frontal gyrus	6	Right	28	–5	60	2275
Somatosensory cortex	1, 2, 3	Left	–39	–34	54	754
Superior parietal lobe	7	Left	–33	–51	50	330
Precentral gyrus	4	Left	–35	–23	56	1392
Precuneus	7	Left	–13	–70	48	672
Medial frontal gyrus	6, 8	Right	4	18	44	1433
Dorsolateral PFC	9	Right	43	4	36	445
Dorsolateral PFC	9	Left	–31	52	25	423
Inferior parietal lobe	40	Right	54	–47	30	750
Medial occipital gyrus	19	Left	–23	–86	19	552
Dorsal striatum		Right	15	2	15	2372
Dorsal striatum		Left	–13	–3	17	1008
Ventral striatum		Left	–18	12	2	492
Ventral striatum		Right	17	12	0	428
Inferior frontal gyrus	45	Left	–30	22	7	312
Lingual gyrus	18	Right	12	–62	6	380
Occipital lobe	17, 18		1	–80	–4	10996
Cerebellum		Left	–33	–52	–23	2297
Cerebellum		Right	31	–43	–21	2255

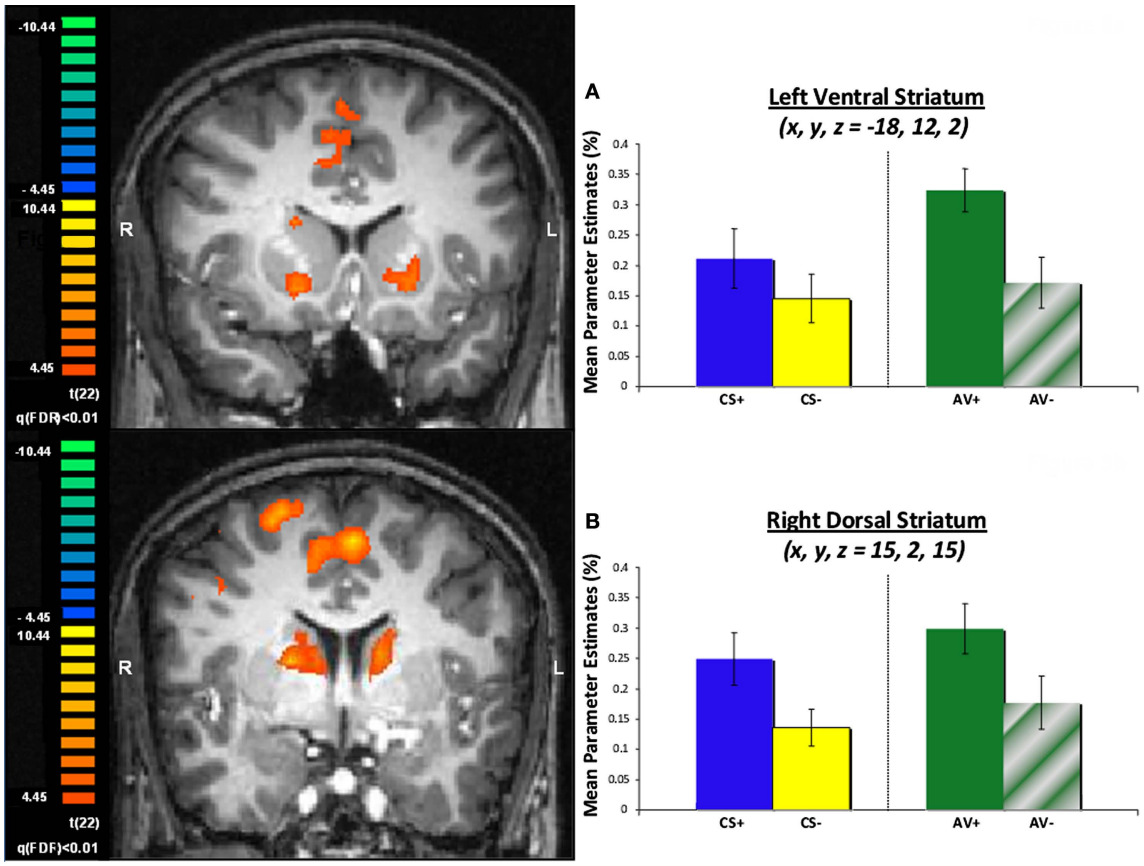


FIGURE 3 | Striatum ROIs defined by a contrast of aversive (CS+ and AV+) and safe (CS–) trials. (A) BOLD signals in the left ventral striatum are depicted as mean parameter estimates and highlight differential response pre (AV+) and post (AV–) learning of an avoidance response. **(B)** BOLD signals in the right dorsal striatum displayed as mean parameter estimates also showing a differential response pre (AV+) and post (AV–) learning of an avoidance response.

Activity within the amygdala ROI in the left hemisphere showed differential responses between CS+ and CS– trials [$t(22) = 2.28$, $p < 0.05$; **Figure 4**], consistent with previous literature and contrast used to define this ROI, while differential response between AV+ and AV– trials approached significance [$t(22) = 1.53$, $p = 0.14$].

A measure of the magnitude of avoidance learning was calculated from mean parameter estimates (i.e., mean beta weights) with the goal of contrasting the *a priori* ROIs (i.e., amygdala and striatum) during task performance. Specifically, we used the difference between mean parameter estimates during AV+ and AV– trials, reflecting the difference between pre- and post-learning of an avoidance response. This measure of the magnitude of avoidance learning for the amygdala ROI was then correlated with the same measure for both left ventral striatum ($r = 0.51$, $p < 0.05$) and right dorsal striatum ($r = 0.54$, $p < 0.05$) ROIs previously described. To better quantify a potential interaction between the amygdala and striatum during avoidance learning, however, an exploratory connectivity analysis was performed where the time course of amygdala activation for each individual subject was used as a single predictor in a GLM. As hypothesized by animal models (LeDoux and Gorman, 2001), it was expected that the seed ROI, the amygdala activation pattern, would correlate with striatum activity during task performance. With the caveat that this analysis included the entire task, and not selected types of trials (e.g., avoidance trials), activation in the striatum bilaterally was observed to correlate with the seed amygdala ROI. Specifically, ROIs in the left ($x, y, z = -7$, 15, 5; 1615 voxels) and right ($x, y, z = 7$, 9, 3; 903 voxels) ventral caudate nucleus were observed in this analysis (**Figure 5A**), with some degree of overlap with the striatum ROI previously defined by the general analysis (**Figure 5B**).

Finally, an exploratory analysis was performed comparing AV+ and AV– trials, which investigated brain regions distinctly associated with learning changes during avoidance trials. Corticostriatal circuits typically involved in reinforcement learning (for review see O'Doherty, 2004; Daw and Doya, 2006; Balleine et al., 2007) were identified in this contrast (**Table 2**). These included ROIs in the dorsal and ventral striatum, along with dorsal (BA 6) and ventromedial (BA 10/32) prefrontal regions. Interestingly, both striatum ROIs showed a pattern of response resembling learning signals, as only the AV+ trials, where learning could occur, elicited a strong BOLD signal as represented by higher mean parameter estimate values.

DISCUSSION

The goal of this study was to explore the neural circuitry underlying active coping of fear in humans using a variant of an aversive conditioning paradigm where conditioned fear could be

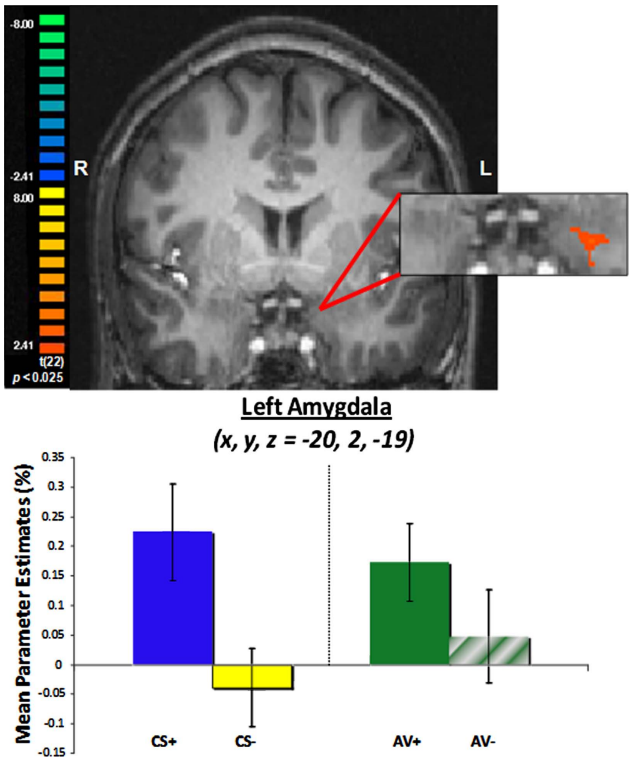


FIGURE 4 | Left amygdala ROI defined by a contrast of aversive (CS+ and AV+) and safe (CS–) trials.

diminished by an instrumental action – an avoidance response. Participants first acquired a behavioral response to terminate delivery of a mild shock, and then continued to use this response to refrain from further aversive outcomes. Physiological arousal, as index by SCRs, supported the behavioral evidence of learning as arousal levels were decreased post-learning of an avoidance response. Additionally, instrumental behavior was faster during avoidance trials, compared to trials where a certain outcome was expected (i.e., non-contingent response), potentially indicating higher motivational levels when an opportunity to exert control over an emotional event is present. A contrast of aversive and safe trials identified *a priori* ROIs in both dorsal and ventral striatum along with amygdala, with BOLD signals within the striatum differing between pre- and post-learning of an avoidance response, a measure that correlated with BOLD signals in the amygdala. This was supported by a connectivity analysis using the amygdala as a seed ROI which found correlations with the striatum. As postulated by non-human animal models (Killcross et al., 1997; Everitt et al., 1999; LeDoux and Gorman, 2001; Cardinal et al., 2002), active coping of fear in humans may involve amygdala and striatal interactions as a means of diminishing conditioned fear and exerting control over emotional responses.

The involvement of the striatum in active avoidance has been previously observed in animal studies (Winocur and Mills, 1969; Allen and Davison, 1973; Neill et al., 1974; McCullough et al., 1993; Li et al., 2004). In the context of this human paradigm, the striatum was particularly involved in learning a behavioral action that allowed for control over conditioned fear, highlighting the role

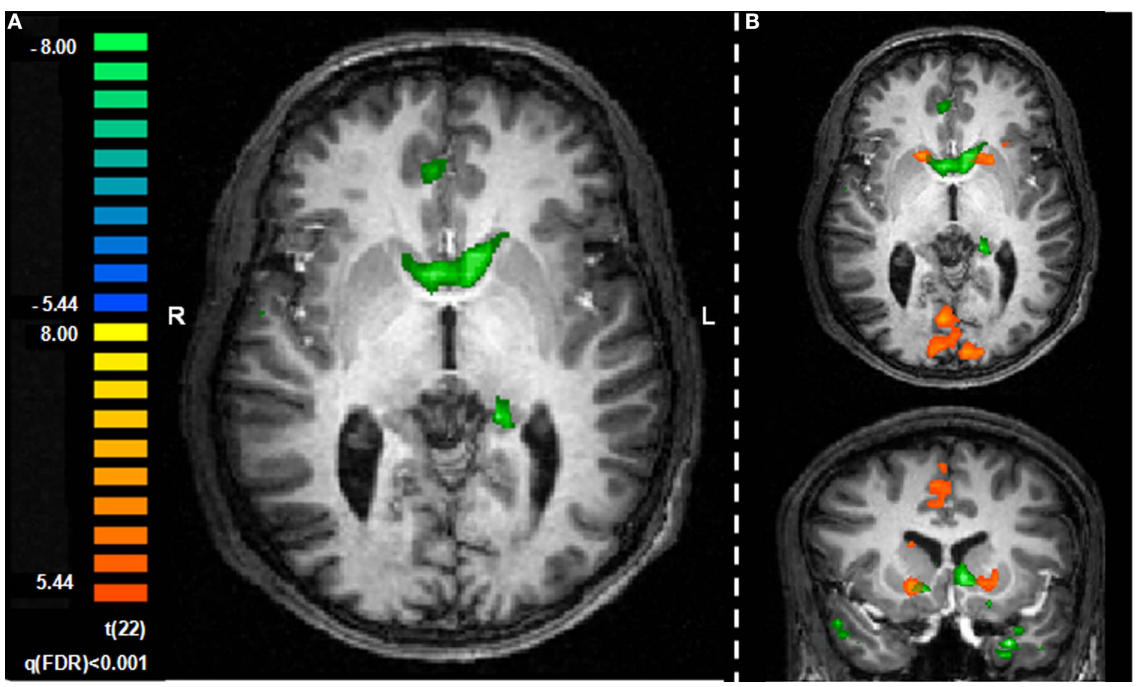


FIGURE 5 | Exploratory connectivity analysis with amygdala as a seed ROI. (A) Time course for the amygdala ROI extracted for each individual participant and entered as a single predictor in a general linear model correlates with activation in the striatum. **(B)** Striatum ROIs defined by the seed analysis (green) overlap with striatum ROIs defined by general contrast of aversive and safe trials (orange). Slices are shown in the axial ($z = 1$) and coronal ($y = 13$) planes.

Table 2 | Contrast of AV+ and AV– trials at $p < 0.005$ and contiguity threshold of four voxels.

Region of activation	Brodmann areas	Laterality	Talairach coordinates			#Voxels
			x	*y*	*z*	
Dorsomedial frontal gyrus	6	Right	33	1	55	377
Dorsal striatum		Right	16	12	19	130
Ventral striatum		Left	–20	11	1	126
Frontal medial gyrus	10	Right	39	46	1	199
Ventromedial PFC	32/10	Right	25	45	–1	248
Ventromedial PFC	32/10	Right	11	37	–6	578

of the striatum in action-contingency during learning (O'Doherty et al., 2004; Tricomi et al., 2004), while extending it to aversive states. Despite its functional heterogeneity and connectivity to regions such as the amygdala, the human striatum is typically discussed in the context of reward processing (for review see Rangel et al., 2008), although evidence continues to suggest the striatum's involvement during affective learning irrespective of type of reinforcer (positive or negative). This paradigm provides additional support for the involvement of the human striatum in processing negative reinforcement.

The amygdala is a structure linked to aversive processes, particularly the acquisition of fears (for review see Phelps and LeDoux, 2005). Studies in animals also link the amygdala with certain forms of avoidance learning (e.g., Killcross et al., 1997; Machado et al., 2009), or escaping from fear (Amorapanth et al., 2000), leading to the hypothesis that amygdala–striatum interaction could underlie one's ability to actively cope with conditioned fear (LeDoux and Gorman, 2001). Given this a priori hypothesis, we used a lenient threshold previously used by other human fear conditioning studies (e.g., LaBar et al., 1998) to investigate the role of the amygdala in human avoidance learning. Although the results have to be carefully considered given the lenient threshold, we observed a correlation between the time course of amygdala activation during task performance and the striatum, supportive of a potential interaction between the two structures during avoidance learning.

Previous research investigating the regulation of fear in humans has examined passive extinction techniques (Milad and Quirk, 2002; Knight et al., 2004; Phelps et al., 2004) and the use of cognitive strategies (for review see Ochsner and Gross, 2005). In the current paper, we examine the role of active coping strategies, particularly taking an action to avoid a potential threat, to adaptively control fear. A common finding across the three types of techniques is that conditioned fear is diminished, as evidenced by a physiological correlate of fear (SCRs) and decreases in BOLD response in the amygdala. Interestingly, the left amygdala ROI identified in a previous cognitive regulation study of conditioned fear (Delgado et al., 2008b; $x, y, z = -20, 0, -20$) was quite similar to the left amygdala ROI identified in the current study using an avoidance paradigm ($x, y, z = -20, 2, -19$). One potential difference across the techniques, however, is the role of the striatum in the control of fears. Striatum activation has been reported in previous papers examining the control of fear through passive extinction and cognitive strategies techniques (Phelps et al., 2004; Delgado et al., 2008b), although the particular contrast was a general effect of conditioning. The current findings suggest that the motivation to avoid a negative outcome with an instrumental response engages the striatum even more than just simple conditioning as evidenced by increased activation in the left striatum during AV+ trials, when subjects were learning the avoidance response, compared to CS+ trials when they were simply anticipating an aversive outcome, further highlighting the involvement of the striatum in learning via negative reinforcement.

The paradigm used for this experiment was adapted from previous animal (Amorapanth et al., 2000) and human (Phelps et al., 2004; Delgado et al., 2008b)studies of aversive conditioning. It is a simple task that has distinct advantages for studying a complex process such as avoidance learning. The inclusion of separate CS and response phase, for instance, allows probing of neural responses to the initial representation of the CS without the potential motor and motor preparation confounds associated with this instrumental procedure. This paradigm can also be adapted to study avoidance learning with secondary reinforcers (e.g., money; see Kim et al., 2006), comparisons between reinforcers of different valence (positive and negative reinforcement), and varying levels of probability or complexity of avoidance response (e.g., manipulation of effort required to successfully avoid negative outcome). This avoidance paradigm also has its limitations, however, such as the minimal amount of trials experienced by a participant per condition (6), which required a fixed reinforcement schedule. The task length was designed to limit the amount of shocks administered and, due to piloting, provide an ideal time window where participants felt that they could indeed be successful. It is also possible that some type of habituation can occur in this design as AV+ and AV– trials are separated in time within a scanning session. An argument against habituation being an explanation for the observed results, however, is that individuals who failed to learn the task did not show the differential responses in the striatum when comparing AV+ and AV– trials, which was characteristic of successful task learning (see Supplementary Materials).

In summary, we used a variant of a fear conditioning study where participants had a chance to avoid a negative outcome with an instrumental behavior. Consistent with animal models (e.g., LeDoux and Gorman, 2001), we found amygdala–striatal interactions in humans potentially underlying avoidance learning and exerting control over conditioned fears. Future studies will probe how human mechanisms of affective learning through negative

reinforcers (i.e., avoidance learning) compare to learning through positive reinforcers (i.e., approach learning) to further understand the range of involvement of regions such as the striatum in affective learning, behavioral control and decision-making.

ACKNOWLEDGMENTS

This work was supported by James S. McDonnell foundation grant to Elizabeth A. Phelps, the Beatrice and Samuel A. Seaver Foundation and a National Institute on Drug Abuse grant to Mauricio R. Delgado (DA022998). We would like to acknowledge Mike Niznikiewicz for technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/behavioralneuroscience/paper/10.3389/neuro.08/033.2009/

REFERENCES

Allen, J. D., and Davison, C. S. (1973). Effects of caudate lesions on signaled and nonsignaled Sidman avoidance in the rat. *Behav. Biol.* 8, 239–250.

Amorapanth, P., LeDoux, J. E., and Nader, K. (2000). Different lateral amygdala outputs mediate reactions and actions elicited by a fear-arousing stimulus. *Nat. Neurosci.* 3, 74–79.

Balleine, B. W., Delgado, M. R., and Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* 27, 8161–8165.

Buchel, C., Morris, J., Dolan, R. J., and Friston, K. J. (1998). Brain systems mediating aversive conditioning: an event-related fMRI study. *Neuron* 20, 947–957.

Cain, C., and LeDoux, J.E. (2008). Emotional processing and motivation: in search of brain mechanisms. In Handbook of Approach and Avoidance Motivation, A. Elliot, ed. (New York, Psychology Press), pp. 17–34.

Cardinal, R. N., Parkinson, J. A., Hall, J., and Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.* 26, 321–352.

Daw, N. D., and Doya, K. (2006). The computational neurobiology of learning and reward. *Curr. Opin. Neurobiol.* 16, 199–204.

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Ann N. Y. Acad. Sci.* 1104, 70–88.

Delgado, M. R., Li, J., Schiller, D., and Phelps, E. A. (2008a). The role of the striatum in aversive learning and aversive prediction errors. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 363, 3787–3800.

Delgado, M. R., Nearing, K. I., Ledoux, J. E., and Phelps, E. A. (2008b). Neural circuitry underlying the regulation of conditioned fear and its relation to extinction. *Neuron* 59, 829–838.

Delgado, M. R., Stenger, V. A., and Fiez, J. A. (2004). Motivation-dependent responses in the human caudate nucleus. *Cereb. Cortex* 14, 1022–1030.

Everitt, B. J., Morris, K. A., O'Brien, A., and Robbins, T. W. (1991). The basolateral amygdala–ventral striatal system and conditioned place preference: further evidence of limbic–striatal interactions underlying reward-related processes. *Neuroscience* 42, 1–18.

Everitt, B. J., Parkinson, J. A., Olmstead, M. C., Arroyo, M., Robledo, P., and Robbins, T. W. (1999). Associative processes in addiction and reward. The role of amygdala–ventral striatal subsystems. *Ann. N. Y. Acad. Sci.* 877, 412–438.

Jensen, J., McIntosh, A. R., Crawley, A. P., Mikulis, D. J., Remington, G., and Kapur, S. (2003). Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron* 40, 1251–1257.

Kalisch, R., Wiech, K., Critchley, H. D., Seymour, B., O'Doherty, J. P., Oakley, D. A., Allen, P., and Dolan, R. J. (2005). Anxiety reduction through detachment: subjective, physiological, and neural effects. *J. Cogn. Neurosci.* 17, 874–883.

Killcross, S., Robbins, T. W., and Everitt, B. J. (1997). Different types of fear-conditioned behaviour mediated by separate nuclei within amygdala. *Nature* 388, 377–380.

Kim, H., Shimojo, S., and O'Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol.* 4, e233. doi:10.1371/journal.pbio.0040233

Knight, D. C., Smith, C. N., Cheng, D. T., Stein, E. A., and Helmstetter, F. J. (2004). Amygdala and hippocampal activity during acquisition and extinction of human fear conditioning. *Cogn. Affect. Behav. Neurosci.* 4, 317–325.

Knutson, B., and Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Curr. Opin. Neurol.* 18, 411–417.

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., and Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron* 20, 937–945.

LaBar, K. S., LeDoux, J. E., Spencer, D. D., and Phelps, E. A. (1995). Impaired fear conditioning following unilateral temporal lobectomy in humans. *J. Neurosci.* 15, 6846–6855.

LeDoux, J. E., and Gorman, J. M. (2001). A call to action: overcoming anxiety through active coping. *Am. J. Psychiatry* 158, 1953–1955.

Li, M., Parkes, J., Fletcher, P. J., and Kapur, S. (2004). Evaluation of the motor initiation hypothesis of APD-induced conditioned avoidance decreases. *Pharmacol. Biochem. Behav.* 78, 811–819.

Machado, C. J., Kazama, A. M., and Bachevalier, J. (2009). Impact of amygdala, orbital frontal, or hippocampal lesions on threat avoidance and emotional reactivity in nonhuman primates. *Emotion* 9, 147–163.

McCullough, L. D., Sokolowski, J. D., and Salamone, J. D. (1993). A neurochemical and behavioral investigation of the involvement of nucleus accumbens dopamine in instrumental avoidance. *Neuroscience* 52, 919–925.

Menon, M., Jensen, J., Vitcu, I., Graff-Guerrero, A., Crawley, A., Smith, M. A., and Kapur, S. (2007). Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: effects of dopaminergic modulation. *Biol. Psychiatry* 62, 765–772.

Milad, M. R., and Quirk, G. J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature* 420, 70–74.

Mogenson, G. J., Jones, D. L., and Yim, C. Y. (1980). From motivation to action: functional interface between the limbic system and the motor system. *Prog. Neurobiol.* 14, 69–97.

Montague, P. R., and Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron* 36, 265–284.

Neill, D. B., Boggan, W. O., and Grossman, S. P. (1974). Impairment of avoidance performance by intrastriatal administration of 6-hydroxydopamine. *Pharmacol. Biochem. Behav.* 2, 97–103.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., and Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* 304, 452–454.

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776.

Ochsner, K. N., and Gross, J. J. (2005). The cognitive control of emotion. *Trends Cogn. Sci* 9, 242–249.

Phelps, E. A., Delgado, M. R., Nearing, K. I., and LeDoux, J. E. (2004). Extinction learning in humans: role of the amygdala and vmPFC. *Neuron* 43, 897–905.

Phelps, E. A., and LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* 48, 175–187.

Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556.

Robbins, T. W., Cador, M., Taylor, J. R., and Everitt, B. J. (1989). Limbic–striatal interactions in reward-related processes. *Neurosci. Biobehav. Rev.* 13, 155–162.

Seymour, B., Daw, N., Dayan, P., Singer, T., and Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *J. Neurosci.* 27, 4826–4831.

Seymour, B., O'Doherty, J. P., Dayan, P., Koltzenburg, M., Jones, A. K., Dolan, R. J., Friston, K. J., and Frackowiak, R. S. (2004). Temporal difference models describe higher-order learning in humans. *Nature* 429, 664–667.

Talairach, J., and Tournoux, P. (1988). Co-Planar Stereotaxic Atlas of the

Human Brain: An Approach to Medical Cerebral Imaging. Stuttgart, NY, Medical Publishers.

Tricomi, E. M., Delgado, M. R., and Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron* 41, 281–292.

Winocur, G., and Mills, J. A. (1969). Effects of caudate lesions on avoidance behavior in rats. *J. Comp. Physiol. Psychol.* 68, 552–557.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 June 2009; paper pending published: 28 July 2009; accepted: 12 September 2009; published online: 01 October 2009.
*Citation: Delgado MR, Jou RL, LeDoux JE and Phelps EA (2009) Avoiding negative outcomes: tracking the mechanisms of avoidance learning in humans during fear conditioning. Front. Behav. Neurosci. **3**:33. doi: 10.3389/neuro.08.033.2009*

frontiers in
BEHAVIORAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 08 February 2010
doi: 10.3389/neuro.08.057.2009

Modeling the value of strategic actions in the superior colliculus

Dhushan Thevarajah[1†], Ryan Webb[2†], Christopher Ferrall[2] and Michael C. Dorris[1]*

[1] Department of Physiology, Centre for Neuroscience Studies and Canadian Institutes of Health Research Group in Sensory-Motor Systems, Queen's University, Kingston, ON, Canada
[2] Department of Economics, Queen's University, Kingston, ON, Canada

Edited by:
Daeyeol Lee, Yale University School of Medicine, USA

Reviewed by:
Tim Behrens, University of Oxford, UK
Hiroshi Abe, Yale University School of Medicine, USA

***Correspondence:**
Michael C. Dorris, Department of Physiology, Room 440 Botterell Hall, Queen's University, Kingston, ON, Canada K7L3N6.
e-mail: dorrism@biomed.queensu.ca

[†]Dhushan Thevarajah and Ryan Webb contributed equally to this work.

In learning models of strategic game play, an agent constructs a valuation (action value) over possible future choices as a function of past actions and rewards. Choices are then stochastic functions of these action values. Our goal is to uncover a neural signal that correlates with the action value posited by behavioral learning models. We measured activity from neurons in the superior colliculus (SC), a midbrain region involved in planning saccadic eye movements, while monkeys performed two saccade tasks. In the strategic task, monkeys competed against a computer in a saccade version of the mixed-strategy game "matching-pennies". In the instructed task, saccades were elicited through explicit instruction rather than free choices. In both tasks neuronal activity and behavior were shaped by past actions and rewards with more recent events exerting a larger influence. Further, SC activity predicted upcoming choices during the strategic task and upcoming reaction times during the instructed task. Finally, we found that neuronal activity in both tasks correlated with an established learning model, the Experience Weighted Attraction model of action valuation (Camerer and Ho, 1999). Collectively, our results provide evidence that action values hypothesized by learning models are represented in the motor planning regions of the brain in a manner that could be used to select strategic actions.

Keywords: decision, macaque, mixed strategy, motor intention, saccade, reinforcement, game theory, EWA

INTRODUCTION

In reinforcement learning models, an individual's choice is a probabilistic function of the current values of possible actions, which in turn are functions of past choices and past rewards (Sutton and Barto, 1998). These learning models are based on the concept of choice reinforcement, traced back to the Law of Effect (Thorndike, 1898; Erev and Roth, 1998).

Empirical studies have supported such learning models in a variety of strategic environments with mixed strategy equilibria (Mookherjee and Sopher, 1994, 1997; Erev and Roth, 1998; Camerer and Ho, 1999; Ho et al., 2007, 2008). However, because learning models predict serial dependence in sequential choices, they conflict with independent (uncorrelated) choice predicted by repetition of the stage game Nash Equilibrium in a repeated game. For example, while laboratory studies of the matching pennies game in humans confirm the equilibrium prediction of a 50/50 ratio of choices, sequential dependencies in individual choices remain (Mookherjee and Sopher, 1994; Ochs, 1995). Similar results have been observed against a computer opponent in studies of both humans (Spiliopoulos, 2008) and monkeys (Lee et al., 2004; Thevarajah et al., 2009). Studies of a broader class of mixed strategy games also exhibit similar choice dependencies though not all the authors address learning models directly (O'Neill, 1987; Brown and Rosenthal, 1990; Rapoport and Boebel, 1992; Rapoport and Budescu, 1992; McCabe et al., 2000).

The goal of this study is to look for evidence of neuronal signals that correlate with the action values predicted by the Experience Weighted Attraction (EWA) learning model (Camerer and Ho, 1999). We use EWA because it is both empirically established and a general formulation. It incorporates simple reinforcement learning (Win/Stay-Lose/Shift), both cumulative reinforcement learning and average reinforcement learning (or Q-Learning) (Watkins, 1989; Erev and Roth, 1998), and belief-based models (Fudenberg and Levine, 1998), as special cases of its parameterization. In fact, it is entirely reasonable for behaviour to lie in some middle ground of the above model restrictions of EWA, and empirical evidence suggests it does (Camerer and Ho, 1999; Ho et al., 2008).

Evidence that learning models are instantiated by the brain has been found from measuring neural signals while humans and animals decide. Evaluative signals are encoded, in part, via dopaminergic structures which represent the difference between realized and expected reward following an action (Schultz, 2004; Caplin et al., 2010). In addition, neural signals have been found that encode the combination of actions and their associated outcomes during adaptive decision-making (Barraclough et al., 2004; Lau and Glimcher, 2007; Seo et al., 2007; Luk and Wallis, 2009). Finally, some neural signals reflect the value of potential actions. Thus they may play an important role in driving the choice process (Platt and Glimcher, 1999; Dorris and Glimcher, 2004; Rushworth et al., 2004; Sugrue et al., 2004; Samejima et al., 2005; Padoa-Schioppa and Assad, 2006; Kennerley et al., 2006; Lau and Glimcher, 2008; Jocham et al., 2009).

We build on this previous work by looking for action value signals within a brain region quite close to the motor output, the intermediate layers of the superior colliculus (SCi). The SCi has a number features that suggest it may encode action value. The

SCi is topographically organized as a map of potential saccadic eye movements (Robinson, 1972; Schiller and Stryker, 1972) and determines when and where a saccade will be directed (Glimcher and Sparks, 1992; Dorris et al., 1997). The SCi receives input signals from upstream brain regions involved in choosing saccades in both strategic environments (Barraclough et al., 2004; Dorris and Glimcher, 2004; Seo et al., 2007; Seo and Lee, 2008) and non-strategic environments (Schultz, 1998; Sugrue et al., 2004; Samejima et al., 2005; Lau and Glimcher, 2007, 2008). The topographic organization of the SCi ensures that any value-related signals we observe are closely associated with specific actions. Moreover, strong lateral inhibition between distant SCi locations could play an important role in selecting between action values associated with competing saccades (Munoz and Istvan, 1998; Dorris et al., 2007). Finally, the SCi sends commands to premotor neurons in the brainstem (Moschovakis and Highstein, 1994), as well as providing feedback to dopaminergic neurons in the ventral tegmental area and substantia nigra (Comoli et al., 2003; Dommett et al., 2005).

We measured preparatory activity in the SCi while a monkey played a simultaneous move game of matching pennies against a computer algorithm designed to exploit serial dependence in the monkey's choices. To control for any serial dependence outside of strategic competition, we also measured activity during a sequential move game with random payoffs. First, we hypothesize that SCi activity displays serial dependence based on both previous saccades and their outcomes, and that more recent events will exert a stronger influence. Second, we hypothesize that SCi activity predicts upcoming strategic choices. Finally, we hypothesize that activity in the SCi provides a signal that is correlated with the current value of actions in the EWA learning model. Collectively, our results support the conclusion that action value signals are represented in the motor planning regions of the brain in a manner suitable for selecting strategic actions.

MATERIALS AND METHODS

Electrophysiological experiments were conducted on two male rhesus monkeys (Macaca mulatta), weighing between 9–13.5 kg each, while they performed saccadic eye movement tasks. All procedures were approved by the Queen's University Animal Care Committee and complied with the guidelines of the Canadian Council on Animal Care. Animals were under the close supervision of the university veterinarian. Physiological recording techniques as well as the surgical procedures have been described previously (Munoz and Istvan, 1998; Thevarajah et al., 2009).

GENERAL METHODOLOGY

Behavioral paradigms, visual displays, delivery of liquid reward, and storage of both neuronal discharge and eye position data were under the control of a PC computer running a real-time data acquisition system (Gramalkn, Ryklin Software). Red visual stimuli (11 cd/m^2) were produced with a digital projector (Duocom InFocus SP4805, refresh rate 100 Hz) and back-projected onto a translucent screen that spanned 50° horizontal and 40° vertical of the visual space. Right eye position was recorded at 500 Hz with resolution of 0.1° using an infra-red eye tracker system (Eyelink II, SR Research). Trials were aborted online if eye position was not maintained within ±3° of the appropriate spatial location or if saccades were initiated outside the 70–300 ms temporal window following target presentation. We have further discussion of aborted trials in Section "Results".

The activity of single neurons was recorded with tungsten microelectrodes (Frederick Haer, 1–2 MΩ at 1 kHz) and sampled at 1 kHz. Data analysis was performed offline using Matlab, version 7.6.0 (Mathworks Inc.) on an Intel Core 2 Duo processor. To quantify neuronal activity, each spike train was convolved with a post-synaptic activation function with a rise time of 1 ms and a decay time of 20 ms (Thompson et al., 1996).

NEURONAL CLASSIFICATION

We recorded the activity from saccade-related neurons located between 1.0 and 3.0 mm below the surface of the SC. The center of each neuron's response field was defined as the location, relative to central fixation, associated with the most vigorous activity during target-directed saccades. One target was always placed at this location (referred to hereafter as *in*) and the other at the mirror-image location in the opposite hemi-field (*out*) except ten experiments where two neurons located in opposite colliculi were recorded simultaneously. For these dual neuron experiments, the two targets were located in opposite hemifields corresponding to the response fields of the two neurons under study. To be included in our analysis, neurons had to meet two requirements: (1) *motor burst*, a transient burst of activity that was time-locked to onset of the saccade into the response field that surpassed 100 spikes/s and (2) *preparatory activity*, neural activity during the 50 ms that followed presentation of the mixed-strategy targets that exceeded 30 spikes/s and was significantly greater than the mean activity 100 ms before fixation point offset (paired t-test, $p < 0.01$). Note that in the modelling Section "Value, SCi Activity and Actions", this preparatory activity will be designated SC^s_{it}.

BEHAVIORAL TASKS

Monkeys performed two behavioral tasks. In the strategic task, monkeys were free to choose between two saccade targets while they competed against an adaptive computer opponent playing the matching pennies game. In the instructed task, monkeys were instructed which saccade to make with the presentation of a single saccade target on each trial. The purpose of the instructed task is to characterize how SCi activity is shaped by previous choices and outcomes. The strategic task is used to emphasize this relationship between SCi activity and the history of the game, and determine whether SCi activity is predictive of choice in a strategic decision making environment.

Strategic task

Monkeys competed in a saccadic version of the repeated mixed-strategy game matching-pennies against an adaptive computer opponent (**Figure 1**). Each trial, both the subject and computer reveal a strategy *in* or *out*. The monkey, pre-designated the "matcher", wins if their strategies match, and the computer, pre-designated the "non-matcher", wins if their strategies differ. The unique Minimax/Nash Equilibrium in mixed strategies is for each player to play *in* and *out* with equal probability (von Neumann and Morgenstern, 1947; Nash, 1951), though our analysis does not require that equilibrium play is achieved. Because our experimental setup limits the ability for the monkey to suffer a loss we replaced a loss with a withholding of reward, though the equilibrium remains unchanged. The payoff

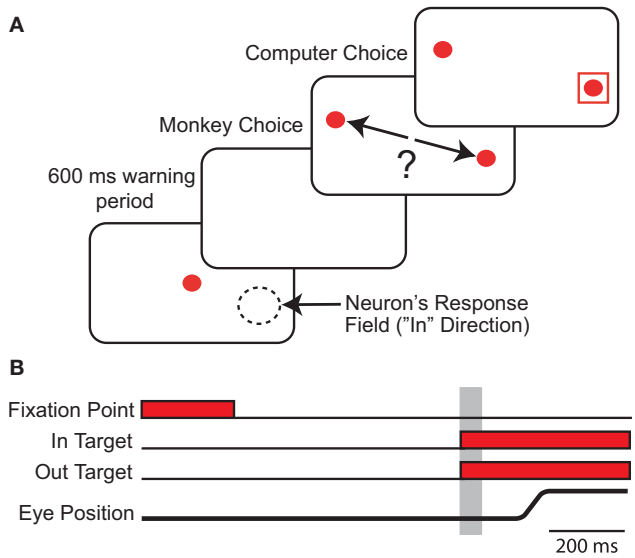


FIGURE 1 | Strategic Task. (A) Each panel represents successive visual displays presented to the monkey. Red circles represent the central fixation point and choice targets respectively. In the third panel, arrows indicate the monkeys possible saccadic choices. One of the saccade targets was always placed in the center of the neuron's response field (i.e., *in* target) as indicted with the dashed circle. The *out* target was placed in the opposite hemifield. The red square indicates the choice of the computer opponent. **(B)** Time-line of strategic task. Grey shaded region indicates the 50 ms epoch during which SCi preparatory activity was sampled for neuronal analyses. The stimuli setup and time-line were identical for the instructed task (not shown) except that only one target was presented per trial and the red square surrounded the target only on rewarded trials.

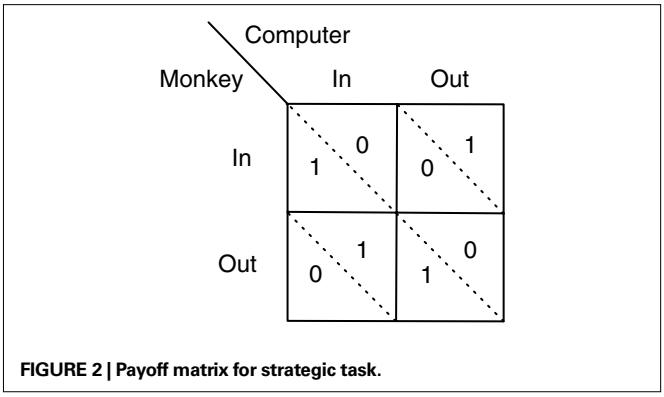


FIGURE 2 | Payoff matrix for strategic task.

matrix is given in **Figure 2** and has been previously studied experimentally in humans (Mookherjee and Sopher, 1994) and monkeys (Lee et al., 2004; Thevarajah et al., 2009).

Subjects were required to maintain central gaze fixation throughout the 800 ms presentation of the fixation point, and after its removal during a fixed 600 ms warning period. Subjects were free to saccade towards either of two simultaneously presented targets, i.e. *in* and *out* of the response field. The fixed warning period and known target locations facilitated advanced selection and preparation of saccades (Thevarajah et al., 2009). After fixating on the target stimulus for 300 ms, a red square, which indicated the computer opponent's choice, appeared around one of the targets for 500 ms. The monkey received a 0.3 mL liquid reward if both players chose the same target and nothing otherwise. The computer opponent performed statistical analyses on the subject's history of previous choices and payoffs and exploited systematic biases in their choice strategy (see algorithm 2 from Lee et al., 2004 for specific details).

Instructed task

The instructed task was identical to the strategic task with two exceptions. First, only a single saccade target was presented on each trial. This target was equally likely to be presented *in* or *out*. Second, reward was equally likely to be received or withheld for successful completion of each trial. Therefore, the expected value of the instructed task is equal to the equilibrium payoff of the strategic task, but saccadic choice was under sensory instruction in the former and under voluntary control in the latter.

DEPENDENCE ON PREVIOUS CHOICES AND REWARDS

To examine any biases exerted by previous saccades and rewards, we segregated SCi activity and saccadic responses on the current trial t based on past ($t-n$, where $1 \leq n \leq 7$) and future ($t+n$, where $1 \leq n \leq 3$) events (Maljkovic and Nakayama, 1994). Future events were examined for control purposes as these should not exert any influence on the current trial. This sequential analysis is illustrated in **Figures 5 and 6** which shows neuronal activity on the current trial segregated into four categories based on four possible events that occurred on the previous trial. (1) a rewarded saccade into the response field (*in*/*R*), (2) an unrewarded saccade into the response field (*in*/*U*), (3) a rewarded saccade out of the response field (*out*/*R*), and (4) an unrewarded saccade out of the response field (*out*/*U*).

We estimated preparatory activity from the postsynaptic spike activation function during the 50 ms following target presentation (**Figure 5**, grey bar). This represented the neuronal firing rate just before saccadic responses were made yet still uncontaminated by visual inputs related to target presentation (Dorris et al., 2000).

The same sequential analysis was performed on choice selection during the strategic task. Response biases were quantified by determining the probability of the monkey selecting the *in* target on the current trial based on past or future events.

Comparatively, for the instructed task, sequential analysis was performed on SRTs rather than saccade choice since saccade location was instructed. SRTs were defined as the time to initiate a saccade following target presentation. Computer software determined the beginning and end of each saccade using a velocity and acceleration threshold. These events were verified by an experimenter to ensure accuracy. Response biases were quantified by examining the influence of an event n trials in the past or future on trials only where saccades were instructed to *in*.

Sequences of trials were constructed from the raw data based on the following criteria. First only sequences of 5 or more consecutive non-aborted trials in length were analyzed. Second, single aborted trials were removed and the sequence was treated as continuous. Third, sequences were started anew if two or more aborted trials occurred in succession. We felt these criteria struck

a balance between providing sufficient sequential data for the analysis in this section while removing those sequences with poor continuity.

EWA LEARNING

The behavior of the subject in trial t of experiment i is coded as

$$s_{i,t} = \begin{cases} in & \text{saccade into response field in trial } t \text{ of experiment } i. \\ out & \text{otherwise.} \end{cases} \quad (1)$$

Let $s'_{i,t}$ denote the computer opponent's choice. Whether reward is received depends on both choices and the experiment being conducted. Let $\pi_{i,t} = 1$ indicate that a reward was received in trial t of experiment i and 0 otherwise.

$$\pi_{i,t} = \begin{cases} R & \text{if } s_{i,t} = s'_{i,t} \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

In both tasks, a reward is only received when the choices match, $s_{i,t} = s'_{i,t}$. During the strategic task the computer opponent makes its choice simultaneously, and if the choices match the subject is rewarded with $R = 1$. During the instructed task, $s'_{i,t}$ is chosen before $s_{i,t}$, but even if the choices agree the monkey is only rewarded half the time:

$$R = \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

Therefore the expected payoff during the instructed task equals the equilibrium expected payoff in the strategic task.

An EWA learning model posits an action value $A^s_{i,t}$ for each strategy s in trial t in experiment i, and includes free parameters which control how action value evolves. On a given trial, it yields a continuous propensity to choose each action, $s_{i,t+1}$, as a monotonic function of current action values $A_{i,t} = [A^{in}_{i,t}, A^{out}_{i,t}]$.

At the start of the experiment $A^{in}_{i0} = A^{out}_{i0}$ for each strategy so that values are equal in the first trial. In general, after trial t the current value of strategy s is updated according to a formula that depends on whether s was chosen or not. If strategy s was chosen then its updated value can be written as a combination of past value (with weight ϕ) and current reward:

$$\text{if } s = s_{i,t}, \text{then } A^s_{i,t} = \frac{\phi N_{t-1} A^s_{i,t-1} + \pi(s_{i,t}, s'_{i,t})}{N_t}. \quad (4)$$

Alternatively, if strategy s was not chosen then its updated value depends on past value (with weight ϕ) and *foregone* payoffs:

$$\text{if } s \neq s_{i,t}, \quad A^s_{i,t} = \frac{\phi N_{t-1} A^s_{i,t-1} + \delta\pi(s, s'_{i,t})}{N_t}. \quad (5)$$

The weight δ is the foregone payoff the subject would have received had it counterfactually chosen s. In both equations, N_t is a trial weight which evolves according to

$$N_t = \rho N_{t-1} + 1. \quad (6)$$

On a given trial, the probability of choosing $s_{i,t} = in$ is defined as

$$P(s_{i,t} = in) = \frac{e^{\lambda A^{in}_{t-1}}}{e^{\lambda A^{in}_{t-1}} + e^{\lambda A^{out}_{t-1}}}, \quad (7)$$

and the parameters $\lambda, \phi, \delta, \rho$, and N_0 are estimated via maximum likelihood. The estimated parameters (except λ) are then used to generate a sequence of fitted action values which we use in our analysis. Importantly, $A_{i,t}$ is computed using only choices and rewards (both actual and fictitious) through trial t, which implies that it can directly enter a model of choice for the next trial, $t+1$. For a complete definition of the EWA model and estimation procedure, see the APPENDIX.

WIN-STAY/LOSE-SWITCH LEARNING

Since EWA is based on a reinforcement premise, it includes a Win-Stay, Lose-Switch choice dependency as a special case. Relative to trial $t + 1$, a Win-Stay outcome for strategy s is coded with an indicator for $s = s_{i,t}$ and $\pi_{i,t} = 1$. A Lose-Switch outcome is $s \neq s_{i,t}$ and $\pi_{i,t} = 0$. This behavior can be captured by a different value, $\text{WSLS}^s_{i,t}$ with its own updating formula,

$$\text{WSLS}^s_{i,t} = \begin{cases} \pi_{i,t} & s = s_{it} \\ 1 - \pi_{i,t} & \text{otherwise.} \end{cases} \quad (8)$$

As in the EWA model, current reward affects the evolution of action value (here represented by $\text{WSLS}^s_{i,t}$). Similarly, the strength of the connection between $\text{WSLS}_{i,t}$ and $s_{i,t+1}$ can be modulated with additional parameters (see Eq. 7). But unlike Eqs. 4 and 5, the WSLS model of value in Eq. 8 does not account for past events before period t nor does it account for a fictitious assessment of actions not chosen (foregone payoffs).

Both Win-Stay/Lose-Switch and the more general EWA models of value predict dependence in the sequence of actions $s_{i,t}$ across adjacent trials. One method for exploring this dependence is to use the updating equations to generate predictors for actions in the following trial. First, we can rewrite Eq. 8 as the sum of two terms,

$$\text{WSLS}^{in}_{i,t} = \text{WS}^{in}_{i,t} + \text{LS}^{in}_{i,t} = 1_{[s_{i,t}=in]}\pi_{i,t} + 1_{[s_{i,t}=out]}\left(1 - \pi_{i,t}\right). \quad (9)$$

This formulation motivates a probit model for choice of the form:

$$P\left(s_{i,t} = in\right) = \Phi\left(\nu_i + \alpha_1 1_{[s_{i,t-1}=in]} + \alpha_2 \text{WS}^{in}_{i,t-1} + \alpha_3 \text{LS}^{in}_{i,t-1}\right), \quad (10)$$

for $t = 1, \ldots, T_i$ and $\Phi()$ denotes the standard normal distribution function (see Wooldridge, 2001 for a discussion of the probit and tobit model introduced below). The term ν_i is a fixed effect for experiment i. $1_{[s_{i,t}=in]}$ is the indicator function which yields 1 if $s_{i,t} = in$ and 0 otherwise. A simple Win-Stay/Lose-Switch hypothesis would predict $\alpha_1 = 0$ since the WS and LS variables would capture all the dependence in the sequence of decisions. Further, it would predict that $\alpha_2 = \alpha_3$, since the effects of winning and losing are symmetric.

VALUE, SCi ACTIVITY AND ACTIONS

To address how value is encoded in neural signals, we introduce our measurement of $SC^s_{i,t}$, defined as the SCi activity associated with saccade target s in trial t of experiment i. In 10 experiments we observe $SC^s_{i,t}$ for both choices; for the other 58 we

observe it only for one choice, $s = in$. To test whether SCi activity encodes the value of actions, in the form of a choice, we estimate the probit

$$P(s_{i,t} = in) = \Phi(\gamma_1 + \gamma_2 SC^{in}_{i,t} + \nu_i), \quad (11)$$

for $t = 1,\ldots,T_i$. Associating $SC^{in}_{i,t}$ with the value of $s = in$ is the hypothesis that $\gamma_2 > 0$. Rejecting the hypothesis $\gamma_2 = 0$ in favour of $\gamma_2 > 0$ is a necessary condition for SC_{i,t} to encode value, but is not sufficient proof that it does.

For the 10 experiments in which we measure SCi activity associated with both choices, we can also estimate a probit of the form

$$Prob(s_{i,t} = in) = \Phi(\mu_1 + \mu_2 \Delta\overline{SC}_{i,t} + \nu_i), \quad (12)$$

where $\Delta\overline{SC}_{i,t} = (SC^{in}_{i,t} - \overline{SC}^{in}_i) - (SC^{out}_{i,t} - \overline{SC}^{out}_i)$ is the difference in SCi activity across actions relative to their within-experiment means, $\overline{SC}^{s}_i$. A positive value for $\Delta\overline{SC}_{i,t}$ indicates the de-meaned activity associated with the *in* target was larger than for *out*. If choice depends on the comparative value of actions, and value is encoded in SCi activity, then choice probabilities should depend on differences in SCi activity. Thus we hypothesize that $\mu_2 > 0$.

Our final hypothesis is that SCi activity reflects the action valuation in the EWA model. To test it, we consider a random-effects regression of the form:

$$SC^{in}_{i,t} = \beta_1 + \beta_2 D_i + \beta_3 A^{in}_{i,t-1} + \beta_4 D_i A^{in}_{i,t-1} + \beta_5 A^{out}_{i,t-1} + \beta_6 D_i A^{out}_{i,t-1} + \nu_i + \varepsilon_{i,t}. \quad (13)$$

For experiments involving the strategic task, we define $D_i = 1$, with $D_i = 0$ for the instructed task. The constant term, β_1, records the conditional mean activity for the sample of neurons examined, while β_2 measures the effect of the strategic task on this baseline activity. The coefficient β_3 captures the relationship between SCi activity (*in* the response field) and the EWA action value of choosing *in*. The strength of association between SCi activity and action value in the strategic task is determined by the value of the interaction parameter β_4. To capture any relationship between SCi activity (for *in*) and the valuations of alternative actions we include $A^{out}_{i,t-1}$ as a regressor with parameter β_5. Again, this relationship in the strategic task is reflected by the interaction parameter β_6. Our hypothesis is that only EWA action value for *in* positively influences SC activity: $\beta_3 > 0$, $\beta_3 + \beta_4 > 0$, $\beta_5 \leq 0$, $\beta_5 + \beta_6 \leq 0$.

Since SCi activity varies continuously, we can estimate equation 13 as a regression. However, on some trials there is no SCi activity measured during our 50 ms preparatory epoch, thus there is left-censoring at zero of the endogenous variable $SC^{s}_{i,t}$ for a small but sizeable portion of trials. We account for this censoring by estimating equation 13 as a tobit model.

We should emphasize the timing of our regression equations 12 and 13, presented graphically in **Figure 3**. The EWA valuation A^{s}_{t-1} is a function of all observed choices and rewards through trial $t-1$ (see Appendix). SCi activity in trial t, SC^{s}_{t}, is a function of A^{s}_{t-1}, therefore is a function of all choices and rewards through trial $t-1$. Finally, the chosen action s_t is a function of the SCi activity in trial t. Importantly, A^{s}_{t-1} does not include any information from the trial t choice. Thus the maintained hypothesis is that *past* action predicts *current* SCi activity which predicts upcoming choice in the *current* trial.

ALGORITHM FOR COMPUTER OPPONENT

The computer algorithm which the monkey competes against is primarily designed to elicit equilibrium behavior from the monkey, that is, a 50/50 randomization of choices. In doing so, the algorithm does not play the Nash strategy itself. This somewhat paradoxical setup is a result of the unstable nature of mixed strategies highlighted by

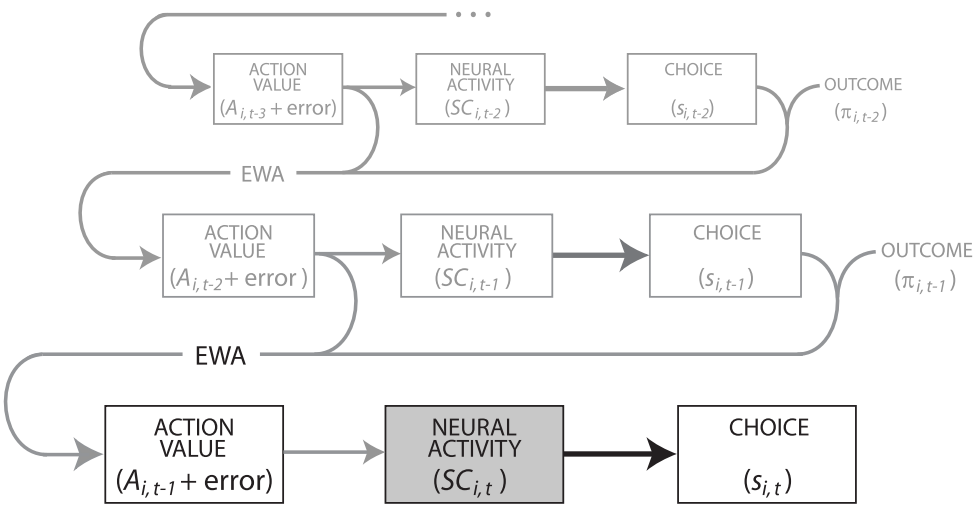


FIGURE 3 | Schematic illustrating the recursive EWA computation and its hypothesized influence on SCi activity. In our experiments, we record and analyze current neural activity ($SC_{i,t}$, gray box). $SC_{i,t}$ lies at the nexus between action value, calculated from past events, and choice on a trial t. Action value ($A_{i,t-1}$) is recursively updated based on an EWA learning model whose main inputs are past choices ($s_{i,t-1}$) and their realized or forgone outcomes ($\pi_{i,t-1}$). Action value also includes an error term to highlight that the calculated $A_{i,t-1}$ only approximates and does not fully capture underlying valuation, in addition to any error arising from noisy neural signals. As such, the relationship between hypothesized $A_{i,t-1}$ and neural activity is likely weaker (thin arrows) than between neural activity and the observable $s_{i,t}$ (thick arrows).

Harsanyi (1973). When the computer is not adaptive, but simply randomizes its choices, the monkey is indifferent between his strategies (any strategy the monkey chooses will be rewarded on half of the trials) and the monkey's choices become strongly biased in one direction (Lee et al., 2004). For this reason, the algorithm was designed to exploit the monkey's choice biases, perhaps more in line with what constitutes (approximate) equilibrium in such games. Refer to algorithm 2 from Lee et al. (2004) for additional details on the computer opponent.

RESULTS

We begin by characterizing the effects of current and previous trials on both behavior and SCi activity in Sections "Analysis of Current Trial", "Dependence of Choice on Previous Trial" and "Sequential Dependence of Choice". In section "Dependence of Choice on Previous Trial", we formally test for a Win-Stay/Lose-Switch strategy. The ability of SCi neurons to predict choice is examined in section "Neuronal Choice Prediction". In section "Behavioural EWA Estimates", we fit the EWA model to choice data and generate sequences of action values for each monkey. Finally, having established that choice is dependent on previous trials, and SCi activity predicts choice on a given trial, in Section "Encoding EWA Action Value" we test our hypothesis that SCi neurons represent the action-specific valuations posited by EWA.

We have data from 68 experiments where neurons satisfied our criteria for inclusion (See Section "Materials and Methods"). In 10 of these experiments, we were able to measure SCi activity associated with both saccades simultaneously, 20 neurons total. In the remaining 58 experiments, we were able to measure SCi activity associated with only one of the potential saccades.

The data consists of a choice, preparatory SCi activity, and a saccadic response time (SRT) for a set of $i = 1 \ldots 78$ neurons respectively with T_i ordered trials. In 38 of these experiments, data were collected for both the strategic task and the instructed task control. This sub-sample of 38 neurons is used in Sections "Analysis of Current Trial", and "Dependence of choice on previous trial", and "Sequential dependence of choice". In this sub-sample, a mean of 246 ± 11 SEM trials per neuron were analyzed during the strategic task and a mean of 146 ± 8 SEM trials per neuron were analyzed during the instructed task. The full sample is used in Section "Neuronal Choice Prediction", while Sections "Behavioural EWA Estimates" and "Encoding EWA Action Value" drop experiments in which greater than 30% of the trials were aborted. These experiments were dropped since many aborted trials within an experiment may interrupt the sequence of valuation posited by EWA learning. The cut-off 30% was set to balance choice sequence consistency and sample size.

ANALYSIS OF CURRENT TRIAL

We will briefly characterize saccade behaviors and SCi preparatory activity on the current trial before examining the effects of events on previous trials. A more detailed current trial analysis can be found in Thevarajah et al. (2009). All reported statistics are (mean ± se).

The allocation of saccade choices did not differ between the two targets during the strategic task [$p(in) = 49.8 \pm 0.6\%$; paired t-test $p > 0.05$]. Moreover, SRTs did not differ between the two targets during the instructed task (*in*: 192.9 ± 4.2 ms, *out*: 186.1 ± 3.7 ms, $p > 0.05$). These behavioral measures suggest that, on average, saccade preparation processes were not biased towards any one particular target location during either task. However, in the strategic task the monkey was rewarded on only 42.2% of the trials, whereas in the instructed task the monkey was rewarded half the time (**Table 1**).

In both tasks, neuronal activity steadily increased during the warning period in advance of choosing either target (**Figure 4**). Overall preparatory activity was greater regardless of saccade direction during the strategic task compared to the instructed task (*in*: $p < 0.05$, *out*: $p < 0.05$). Moreover, in the strategic task activity was segregated for saccades *in* (99.9 ± 8.8 spikes/s) and *out* (80.2 ± 7.2 spikes/s, paired t-test, $p < 0.001$), whereas activity was not segregated between *in* (63.5 ± 6.5 spikes/s) and *out* (64.5 ± 6.5 spikes/s) saccades during the instructed task (paired t-test, $p > 0.05$). This greater overall activation and neuronal selectivity during the strategic task may occur because saccades are under voluntary control and can be planned in advance. In the instructed task the monkey must wait for the presentation of the target.

DEPENDENCE OF CHOICE ON PREVIOUS TRIAL

We examine sequential choice dependencies by segregating behavior and neuronal activity on the events of the previous trial (i.e., previous choice and its reward outcome). Particularly, we test for the prevalence of a WS/LS strategy.

The influence of previous trials on subsequent saccadic responses

We begin by summarizing the frequencies of WS/LS choice patterns in the strategic task over all experiments (**Table 1**). Choices were repeated in a WS/LS pattern in 55.5% of the trials. A WS was observed in 62.1% of post-win trials vs. LS observed in 50.6% of post-loss trials, which suggests a WS/LS strategy is solely due to a Win-Stay rather than Lose-Shift bias. The larger percentage of losing trials suggests the computer opponent was able to exploit this tendency in choice patterns.

To further assess the influence of previous trial events, we estimate Eq. 10 which models choice as a function of lagged choice and the Win-Stay and Lose-Switch variables. Estimates of the fixed-effects probit are presented in **Table 2**. The explicit prediction of the simple WS/LS strategy is rejected because the estimated coefficients α_2 and α_3 are significantly different from each other: the tendency to repeat rewarded actions is greater than the tendency to switch from unrewarded actions. We can also note that the tendency to repeat choices is largely due to the Win-Stay bias since α_1 is not significantly different from zero.

To measure any biases in the instructed task, saccadic reaction times (SRTs) were examined in **Table 3**. Considering that target location and outcome were stochastic, therefore unpredictable, these previous events had a surprisingly large influence on SRTs. Repeating an action resulted in faster SRTs than switching actions (Stay vs. Switch, binomial test $p < 10^{-11}$). SRTs were particularly biased if a saccade direction was previously rewarded (Win-Stay vs. Lose-Stay, binomial test $p < 10^{-5}$; Win-Switch vs. Lose-Switch, binomial test $p < 10^{-3}$). This suggests preparation biases were a function of both previous choices and their outcomes.

The influence of previous trials on SCi preparatory activity

Figure 5 illustrates how SCi activity was also influenced by the previous trial. The black dashed line shows mean activity over all experiments. Each of the coloured lines depicts how current trial activity was influenced by choices and outcomes on the previous

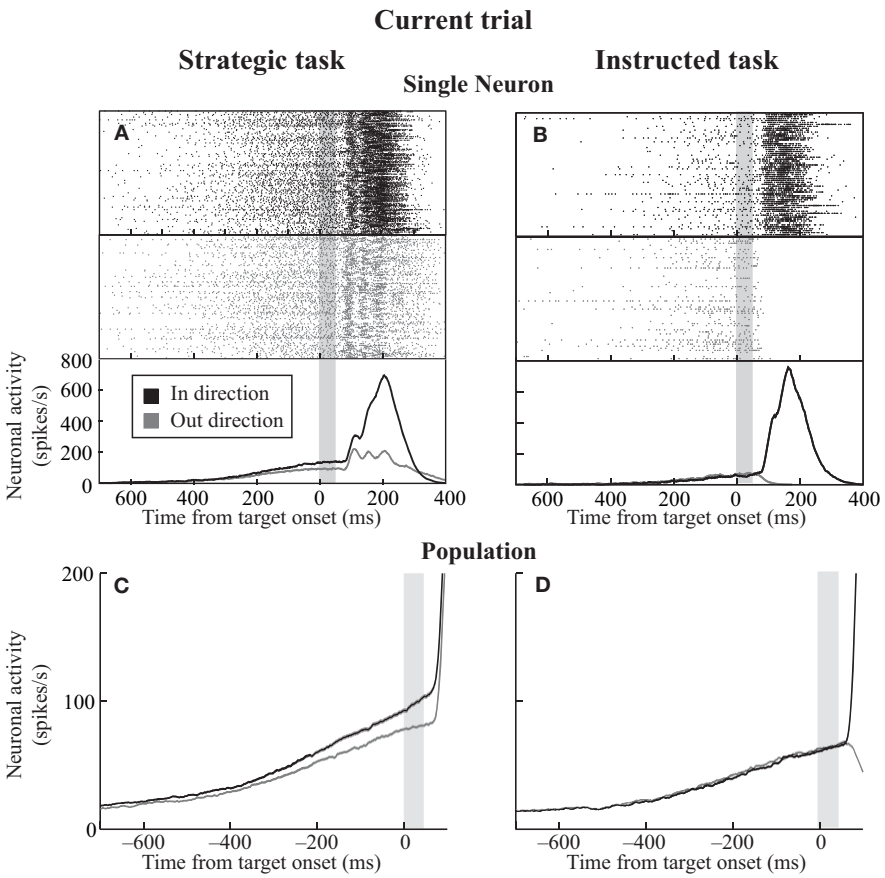


FIGURE 4 | SCi activity during the current trial. (A, B) Activity of a representative SCi neuron during the strategic **(A)** and instructed **(B)** tasks. Rasters (top panels) and post-synaptic activation functions (bottom panel) are sorted based on saccades directed *in* (black) and *out* (gray) of the neuron's response field. The shaded gray bar denotes the epoch during which preparatory activity was analyzed. **(C, D)** Mean activity of neuron sample in which both strategic **(C)** and instructed **(D)** tasks were recorded (38 neurons).

Table 1 | Frequencies of choice dependencies in strategic game.

Previous trial dependency	Proportion (%)
Win	42.2
Loss	57.8
Win-Stay/Lose-Switch	55.5
Win-Stay	62.1
Lose-Switch	50.6

Table 2 | Probit estimates of $s_{i,t+1}$ on lagged choice and Win-Stay/Lose-Switch outcomes.

Variable	Coefficient	Estimate	Standard error	*p*-Value
$I(s^{in}_{i,t-1})$	α_1	0.1415	0.1095	0.20
$WS^{in}_{i,t}$	α_2	0.5478	0.1118	0.00
$LS^{in}_{i,t}$	α_3	0.252	0.0986	0.01

Estimates of Eq. 10 on sample of 33 experiments. (5 redundant paired experiments dropped). Sample size = 8809. Estimates of the 33 experimental effects ν_i *not reported. Reported standard errors are clustered within experiments.*

Table 3 | Reaction time dependencies in instructed task.

Previous trial dependency	Reaction time (ms)	Standard error
Win-Stay	170.8	4.0
Stay	177.2	3.2
Lose-Stay	183.6	5.0
Lose-Switch	194.5	5.0
Switch	200.1	4.0
Win-Switch	205.7	6.4

trial. This influence is most prevalent at the end of the warning period (gray-shaded area). Therefore we will use SCi activity in this epoch for the sequential analysis that follows.

SEQUENTIAL DEPENDENCE OF CHOICE

Having observed a dependency in choices and outcomes in the previous trial, we will now characterize this dependency over multiple trials. Two sequential patterns were evident in both tasks (**Figure 6**). First, more recent events had the greatest influence. Second, actions that were rewarded generally had a more pronounced effect, both in terms of magnitude and duration, than

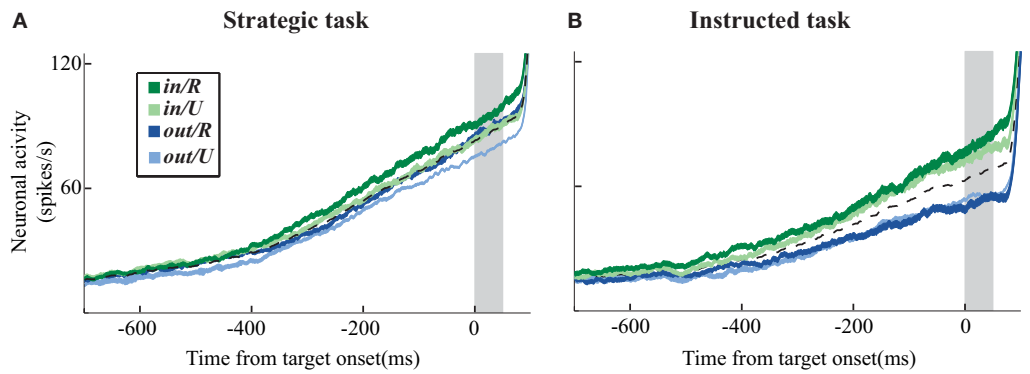


FIGURE 5 | Mean activity of neuronal sample (38 neurons) segregated on previous trials during the strategic (A) and instructed (B) tasks respectively. The black dashed line represents the mean activity for all trials. The colored lines also represent current trial activity, but this activity is segregated into four categories based on the combination of target choice (*in* or *out*) and their outcomes (*R* or *U*). Note that the mean activity is a sum of each of the four colored lines weighted by the proportion of trials in each category. Line widths represent the SEM. The shaded gray bar denotes the epoch during which preparatory activity was sampled during subsequent analyses.

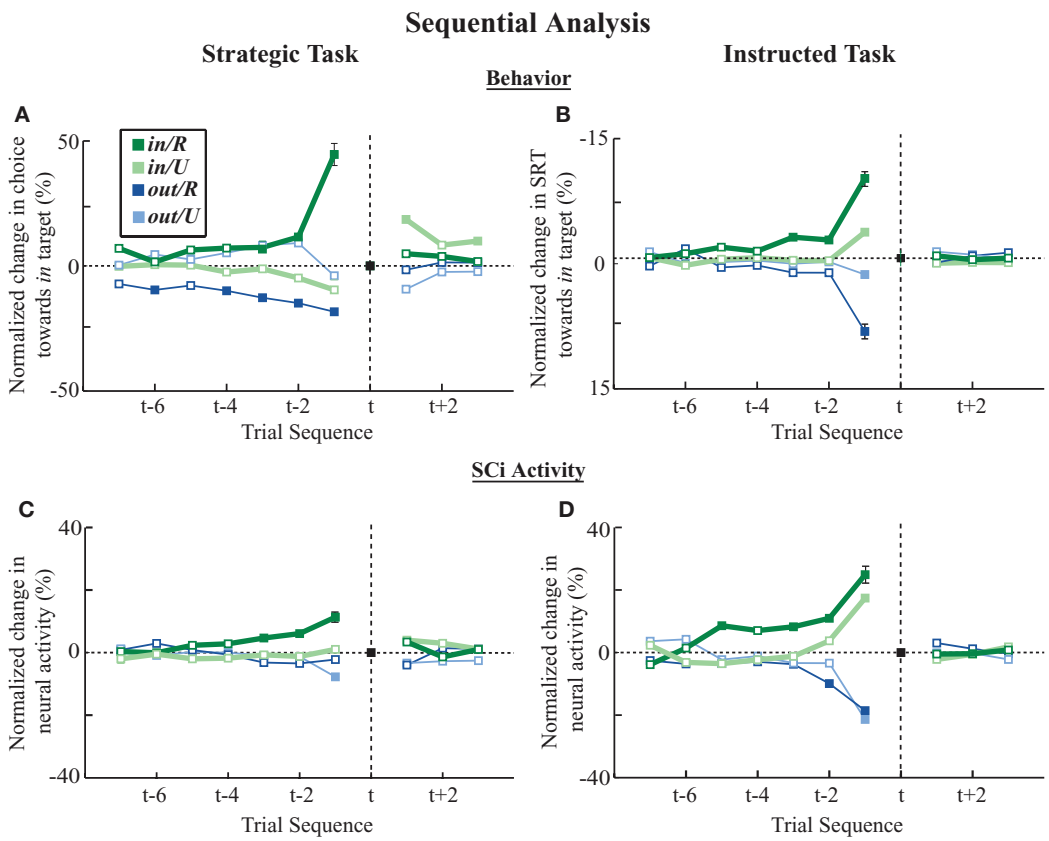


FIGURE 6 | Saccade behavior and SCi activity segregated on previous and future trials. The data for each trial sequence is presented as the percentage change from the mean data for all trials at trial *t* (black dots). Note that each data point in the trial sequence represents the influence of an earlier or later trial on the current trial. Therefore, the four colored data points at each time sequence always sum to the mean data point at time *t* when weighted by the proportion of trials in each category. **(A)** Changes in *in* target choices during the strategic task. **(B)** Changes in *in* target SRTs during the instructed task. Note that the the ordinate axis has been flipped because SRTs are negatively correlated with SCi activity. **(C)** Changes in SCi activity during the strategic task. **(D)** Changes in SCi activity during the instructed task. Filled squares indicate significant differences from the mean activity. Representative standard errors are shown for *in*/*R* data points.

unrewarded actions (**Figure 6**, dark colored lines vs. light colored lines). Whether a previous trial was rewarded or not, did not, by itself, affect SC activity or saccade behaviors. Instead, the effects of reward influenced a particular saccade location rather than providing a general motivating or alerting effect for both actions.

The strategic and instructed tasks also differed in two ways during this sequential analysis. First, future events were correlated with choice selection in the strategic but not the instructed task (**Figure 6A**). This seemingly paradoxical finding is a consequence of the computer exploiting the monkey's Win-Stay bias. That is, monkeys were more likely to lose following a rewarded trial as they tended to repeat actions. This phenomena is evident in the Lose-Stay bias observed in future choices in **Figure 6A**. Second, modulation of SCi activity by past events was greater for the instructed task than for the strategic task. For example, the change in activity imposed by the previous trial was approximately three times as large during the instructed task compared to the strategic task (compare the spread in data along the vertical axis in **Figures 6C vs. D**).

NEURONAL CHOICE PREDICTION

Having characterized serial dependency in choices, the second step in determining whether neurons in the SCi encode action value is to determine if activity predicts choice. The ten experiments where we measured two neurons simultaneously, one for each target, allows us to specify how opposing SCi activity is compared in Eq. 12. Results for the fixed-effects probit estimation are given in **Table 4**.

The parameter μ_2 measures the impact of SCi activity on the probability of an *in* saccade and is both positive and highly significant. To interpret the magnitude of the coefficient μ_2, we take the predicted probabilities from the regression and compare them to the observed choices by two methods. The first rounds the probabilities to the nearest integer and compares them to the choices, resulting in a prediction rate of 65%. The second simulates choices from the binomial distribution using the predicted probabilities, and compares the simulated choices to the actual choices, resulting in a prediction rate of 56% for 1000 simulations. Comparatively, 1000 independent draws from a 50/50 binomial distribution would predict 56% of the trials (560 matches of the monkey's choice) with probability 6.3×10^{-5}. Results did not change significantly when we estimated on half the sample and predicted out of sample.

For the entire 78 neuron sample, we can also assess how well single neurons predict choice from Eq. 11. Results are reported in **Table 5**. Again, we observe that the estimate of SCi activity, γ_2, is both positive and highly significant. As before, our assessment of the magnitude of the parameter γ_2 relies on in-sample prediction. Rounding the fitted probabilities results in a 60% prediction rate for the 78 individual neurons, while simulating the choices results in a 53% prediction rate. As expected, the single neuron is a worse predictor compared to the the paired neuron analysis, presumably because choice is based on a comparison of valuation between the two targets. Again, 1000 independent 50/50 draws would still only predict 53% with probability 0.03.

Table 4 | Probit estimates of $s_{i,t}$ based on difference in activity from neuronal pairs.

Variable	Coefficient	Estimate	Standard error	*p*-Value
Constant	μ_1	0.0435	0.0015	0.00
$\Delta \overline{SC}_{i,t}$	μ_2	0.0054	0.0005	0.00

Estimates of Eq. 12 using ten experiments with paired neuronal measures. Fixed effect estimates are not reported. Standard errors were clustered at the experiment level.

BEHAVIOURAL EWA ESTIMATES

To generate a sequence of action values which reflect each monkey's valuation on a given trial, we estimated the EWA model on choice data (see Section "EWA Learning" and APPENDIX). Estimates are reported in **Tables 6 and 7.** We observe significant heterogeneity in the fitted EWA parameters, similar to Ho et al. (2008). Estimates suggest Monkey H (54/78 experiments) is a cumulative reinforcement learner ($\delta = 0$, $\rho = 0$), while Monkey B (24/78 experiments) has a fictive learning component and averages rewards as in Q-Learning ($\delta > 0$, $\rho = \phi$). For each monkey, the estimates for ϕ, δ, ρ, and N_0 are used to generate the sequence $A^s_{i,t}$ which we use in section "Encoding EWA Action Value".

Table 5 | Probit estimates of $s_{i,t}$ based on activity from individual neurons.

Variable	Coefficient	Estimate	Standard error	*p*-Value
Constant	γ_1	0.0045	0.0008	0.00
$SC_{i,t} - \overline{SC}_i$	γ_2	0.0053	0.0005	0.00

Estimates of Eq. 11 using 78 experiments. Fixed effect estimates are not reported. Standard errors were clustered at the experiment level.

Table 6 | EWA Estimates for Monkey B.

Parameter	Estimate	Standard Error
λ	3.68	3.54
ϕ	0.78	0.08
δ	0.12	0.07
ρ	0.91	0.08
N_0	3.73	8.46

Sample of 19 experiments for monkey B (4/24 experiments dropped due to >30% aborted trials; 1/54 redundant paired experiments additionally dropped).

Table 7 | EWA Estimates for Monkey H.

Parameter	Estimate	Standard Error
λ	0.45	0.29
ϕ	0.52	0.04
δ	0.00	0.05
ρ	0.00	0.64
N_0	1	0

Sample of 27 experiments for monkey H (20/54 experiments dropped due to >30% aborted trials; 7/54 redundant paired experiments additionally dropped). The restriction $N_0 = 1$ was imposed to ensure identification of ρ (see Appendix).

ENCODING EWA ACTION VALUE

The EWA action value is a function of the observed choices and reward structure of the game. Our final hypothesis is that SCi activity reflects the fitted action values from Section "Behavioural EWA Estimates". To test this hypothesis, we estimate Eq. 13 separately for each monkey and its appropriate action value $A^{s}_{i,t}$. Results are reported in **Tables 8 and 9**.

For monkey H (**Table 8**), the instructed task relationship between EWA action value and SCi activity for target *in* is positive, significant and large in magnitude $(\beta_3 = 24.46)$. Over the observed range of the EWA action value $(0.00 < A_{i,t} < 1.96)$, this represents an 81% change in SCi activity relative to baseline activity of 59.86 spikes/s. Notably, this relationship is partially offset by the *out* EWA action value (H_0: $\beta_3 + \beta_5 = 0$, $p = 0.36$). If the action values of the two targets were equal ($A^{in}_{i,t} = A^{out}_{i,t}$), the estimates predict there would still be an increase in SCi activity for the *in* target. This suggests that a given SC neuron encodes the action value for the target it is associated with on the topographic map, but other neurons (valuable targets) can partially inhibit this valuation.

Table 8 | Estimates of $SC^{in}_{i,t}$ on EWA action values and task type for monkey H.

Variable	Coefficient	Estimate	Standard error	*p*-Value
Constant	β_1	59.86	12.94	0.00
D_i	β_2	24.69	9.46	0.01
$A^{in}_{i,t}$	β_3	24.46	3.72	0.00
$D_i A^{in}_{i,t}$	β_4	–9.07	4.58	0.05
$A^{out}_{i,t}$	β_5	–17.95	5.74	0.01
$D_i A^{out}_{i,t}$	β_6	17.25	5.91	0.00
Var(ν_i)	σ^2_ν	59.69	9.86	0.00
Var($\varepsilon_{i,t}$)	σ^2_ε	62.99	5.43	0.00

Random-effects tobit estimates of Eq. 13 on 36 neurons for monkey H (18/54 experiments were dropped due to > 30% aborted trials). Sample size=10704, 998 observations censored at 0. σ^2_ν is the variance of the random effect ν_i; σ^2_ε is the variance of $\varepsilon_{i,t}$. $A^{s}_{i,t}$ is generated using behavioural EWA estimates for monkey H (see Section "Behavioural EWA Estimates"). Standard errors are calculated by means of clustered bootstrap with 1000 bootstrap samples, re-sampling within experiment i.

Table 9 | Estimates of $SC_{i,t}$ on EWA action values and task type for monkey B.

Variable	Coefficient	Estimate	Standard Error	*p*-value
Constant	β_1	48.72	10.39	0.00
D_i	β_2	20.51	9.30	0.03
$A^{in}_{i,t}$	β_3	28.06	21.10	0.18
$D_i A^{in}_{i,t}$	β_4	–5.45	31.28	0.86
$A^{out}_{i,t}$	β_5	8.71	12.62	0.49
$D_i A^{out}_{i,t}$	β_6	–10.93	19.14	0.57
Var(ν_i)	σ^2_ν	35.29	7.28	0.00
Var($\varepsilon_{i,t}$)	σ^2_ε	39.78	4.59	0.00

Random-effects tobit estimates of Eq. 13 on 19 neurons for monkey B (5/24 experiments were dropped due to >30% aborted trials). Sample size = 5907, 258 observations censored at 0. σ^2_ε is the variance of the random effect ν_i; σ^2_ε is the variance of $\varepsilon_{i,t}$. $A^{s}_{i,t}$ is generated using behavioural EWA estimates for monkey B (see Section "Behavioural EWA Estimates"). Standard errors are calculated by means of clustered bootstrap with 1000 bootstrap samples, re-sampling within experiment i.

As expected from our sequential analysis, the relationship between SC activity and action value is attenuated in the strategic task ($\beta_4 < 0$) though it is still positive and significant (H_0: $\beta_3 + \beta_4 = 0$, $p = 0.00$). The estimates yield a 36% increase in SC activity relative to baseline ($\beta_1 + \beta_2 = 84.5$ spikes/s) over the range of $A^{in}_{i,t}$. However the *out* EWA action value now has no impact (H_0: $\beta_5 + \beta_6 = 0, p = 0.82$) suggesting no inhibition from *out* target neurons during this measurement epoch of the strategic task.

Estimation results for monkey B have considerably more variance (**Table 9**). In the instructed task, we still observe a positive coefficient for $A^{in}_{i,t}$ ($\beta_3 = 28.06$) but with a larger p-value ($p = 0.18$) and a smaller magnitude relative to baseline (33%) over the observed range of action values ($0.00 < A_{i,t} < 0.58$). While the estimate for attenuation in the strategic sample is of the correct sign ($\beta_4 < 0$), it is not significantly different from zero ($p = 0.86$). The estimates for the *out* action value are also highly variable and not significantly different from zero in either task. We should note that the sub-sample for monkey B contains half as many observations and neurons as the sub-sample for monkey H, though this efficiency loss likely does not account for all of the increased variability of the estimates.

DISCUSSION

SUMMARY OF FINDINGS

This study examined whether a valuation of future actions, constructed as a function of previous choices and rewards, is represented by the superior colliculus in a strategic environment. Our results show that SCi preparatory activity was shaped by both previous saccades and their outcomes, particularly a Win-Stay bias, and more recent events had a more pronounced effect. These sequential biases were reflected in upcoming choices during the strategic task and upcoming saccadic reaction times during the instructed task.

SCi activity was also predictive of upcoming strategic saccades on a trial-by-trial basis (**Tables 4 and 5**); at a rate of 60% for single neurons and 65% for opposing neuron pairs. Although our pool of neuron pairs was small (10 pairs), this improvement in prediction suggests that it is not the absolute level of activity, but the relative level of activity between potential actions, that is best correlated to choice.

The fact that SCi activity was both shaped by previous choices and rewards and predicted future choices suggest it as a candidate neural correlated of action values posited by behavioural learning model. Our analysis demonstrated that SCi activity was correlated on a trial-by-trial basis with the EWA learning valuation. Specifically, SCi activity was positively correlated with the action value for its response field, with some evidence that it is negatively correlated with the action value of the alternative target. Collectively, our empirical and modelling results suggest that hypothesized action value signals are represented in the motor planning regions of the brain in a manner that could be used to select strategic actions.

EFFECTS OF PREVIOUS ACTIONS AND REWARDS

Serial dependence of choices has previously been observed in strategic and non-strategic environments. Consistent with previous studies, more recent events had a greater influence on both

choices (Juttner and Wolf, 1992; Maljkovic and Nakayama, 1994; Dorris et al., 2000; Barraclough et al., 2004; Lee et al., 2004; Lau and Glimcher, 2005) and neuronal activity (Dorris et al., 2000; Bayer and Glimcher, 2005; Seo and Lee, 2007), and these influences decayed with time (**Figure 6**). Unlike the computer opponent which weighed all past events equally, monkeys gave more weight to recent events when selecting actions. This policy may be an efficient solution for using past events to predict future rewarded actions given organisms have a limited memory store (Anderson et al., 1996; Callicott et al., 1999), and it allows organisms to more readily adapt to a changing environment.

Sequential effects have been characterized previously in the SCi during a task similar to our instructed task (Dorris et al., 2000). Although target location was unpredictable in this previous study, all saccades were rewarded; therefore the contribution from repeating a motor action, or repeating a rewarded location, remained unclear. By allocating rewards unpredictably, we were able to isolate the contribution of these factors. Previously unrewarded actions had a biasing effect, but to a lesser extent than previously rewarded actions. We found no effect of previously rewarded trials when analyzed independently of actions, which suggested that reward, at least our task, did not have a generalized alerting or motivating effect. Instead, SCi activity was found to be influenced by a combination of both previous actions and rewards. These biases, in turn, were reflected in saccade behaviors (**Figure 6**).

Finally, we observed differences in how SC activity was influenced by previous events during the two tasks. First, the overall level of SC activity was greater preceding strategic than instructed saccades (i.e., compare black dashed lines in **Figures 5A vs. B**). Strategic saccades may have been more fully prepared because the locations of the two targets were known in advance whereas the location of the single target had to be identified before the saccade preparation processes could be completed in the instructed task. Second, previous events exerted less influence on SCi activity during the strategic task (i.e., compare **Figures 6C vs. D**). This was observed in the magnitude of the sequential dependencies and the number of previous trials which exerted an influence. Although having sequential biases was seemingly unnecessary in the instructed task, as the monkey could neither control nor predict saccade direction or reward, having such biases were relatively inconsequential. In the strategic task however, sequential biases led to exploitation by the computer opponent as evidenced by a reduced reward rate (**Table 1** and Barraclough et al., 2004). Our results suggest the influence of previous events, borne out in sequential dependencies, can be attenuated in strategic situations.

WIN-STAY BIAS

Though the analysis in Sections "Dependence of Choice on Previous Trial" and "Sequential Dependence of Choice" revealed notable choice tendencies in the strategic sample, many of which are incorporated in the EWA learning model, there is one in particular we wish to highlight. Although both effects were significant, subjects repeated winning choices more often than switching from losing choices controlling for repeated choices ($\alpha_2 > \alpha_3$), or a Win-Stay bias. This observation is a rejection of a strict Win-Stay/Lose-Switch model of choice in repeated games.

However, a stronger Win-Stay bias is compatible with our candidate model of action value (EWA). If unchosen winning actions are updated by a fraction $\delta < 1$ relative to chosen winning actions, the difference in the action value after a rewarded trial is larger than after an unrewarded trial:

$$\Delta A_t \equiv A_t^{s_t} - A_t^{-s_t} = \begin{cases} \frac{\phi N_{t-1}\left(A_{t-1}^{s_t} - A_{t-1}^{-s_t}\right)+1}{N_t} & \text{if } \pi_t = 1 \\ \frac{\phi N_{t-1}\left(A_{t-1}^{s_t} - A_{t-1}^{-s_t}\right)-\delta}{N_t} & \text{if } \pi_t = 0 \end{cases} \quad (14)$$

Therefore

$$\left(\Delta A_t\right)_{\pi_t=1} - \left(\Delta A_t\right)_{\pi_t=0} > \frac{1-\delta}{N_t} > 0. \quad (15)$$

This result holds generally for all models nested by EWA, as long as $\delta < 1$. A Win-Stay bias may be exacerbated in our experiment because our payoff matrix is not zero-sum (**Figure 2**); not matching the opponent constituted a withholding of reward rather than a loss of reward. This asymmetry in payoffs may bias the subject's responses in favour of rewarded trials.

PREDICTING CHOICE

Our results indicate that the activity of individual SCi neurons can predict upcoming choices with 60% reliability. Although significantly better than chance, the SCi may not appear to be a particularly impressive predictor. However, a number of issues must be taken under consideration to make this judgment.

The predictive capability of SCi neurons depends on the number of neurons in the population, the correlation in their firing patterns, and the manner in which downstream structures read-out these predictive signals. Although we only had a sample of 10 neuronal pairs, our results demonstrate that simply comparing the relative firing of two opposing neurons increases prediction from 60% to 65%. Moreover, while the predictive capability of any one (or two) neuron(s) may be weak, this is a very consistent prediction across the neuronal population (see **Figure 5D** from Thevarajah et al., 2009). Therefore, these small individual biases can be amplified to provide a strong signal for selecting strategic actions.

Although the SCi is required for generating saccades (Hanes and Wurtz, 2001) and manipulating SCi activity alters saccadic choices (Carello and Krauzlis, 2004; McPeek and Keller, 2004; Dorris et al., 2007; Thevarajah et al., 2009), the robust activity for *out* direction saccades (**Figure 4**) demonstrates that the reverse is not true; executing a saccade is not a pre-requisite for preparatory SCi activity. This evidence strongly suggests that a causal arrow passes from SCi to choice uni-directionally (**Figure 3**). Similarly, if action value is indeed a function of past choices, then it must be action value that influences SCi activity. If these arrows were not uni-directional then current activity or choices would paradoxically cause past choices.

NEURONAL CORRELATES OF EXPERIENCE WEIGHTED ATTRACTION

Our preliminary analysis has shown that both behaviour and SCi activity are correlated with previous choices and rewards, particularly through a reinforcement of rewarded choices (Win-Stay). To formalize this result, we found a neural correlate of a general learning model based upon this reinforcement premise. This model calculates an action value on each trial as a function of the history of observed

choices and payoff structure of the game. Therefore, our results in Section "Encoding EWA Action Value" are consistent with the hypothesis that neurons in the SCi encode the history of the two tasks in the form of learned action values for each potential action. A given neuron in the SCi is correlated with the action value of its target in both tasks, though the magnitude of this relationship is attenuated in the strategic task. Further, SCi activity is negatively correlated with the action values of competing targets in the instructed task, but not in the strategic task during the period we measure. This suggests that both the attenuation of the value/SCi relationship, and the lack of inhibition from competing neurons *within* the preparatory period we measure, may serve a strategic purpose.

The EWA model we use in this study (Camerer and Ho, 1999) is a general learning model that has proven successful in predicting play both in and out of sample in a wide variety of games. The role EWA plays in our analysis is akin to an objective valuation. It is a function of past choices and rewards which reflects a component of the relative value of each strategy. As such, there remain unaddressed components of value. Learning models do not assess the forward-looking value of an action. That is, there is no consideration of repeated game strategies such as "leading" an opponent in order to exploit him in later periods (though we should emphasize the only unique repeated game equilibrium in matching pennies is the stage game equilibrium). Our analysis also does not address satiation in the experiment nor learning between experiments. However, the relative success of EWA in predicting choice in a strategic environment suggests that its historical, objective component is important in the ultimate valuation of an action.

As a theoretical construct of valuation, both the simplifying assumptions mentioned above and additional neural and/or behavioural factors will combine to limit the explanatory power of EWA (referred to in **Figure 3**). But even if the SC is not coding action value as specified by EWA, the fact that EWA action value significantly predicts SC activity suggests that the correct model will share many features of the EWA formulation. Whether a complete model actually nests EWA as a special case remains an open question that is beyond the scope of this paper.

There has been some progress in identifying the neural correlates of the functional elements of EWA. It has been previously observed that the striatum encodes the difference between realized and expected reward, suggesting the striatum may form part of a learning system in the brain (Schultz, 1998; Caplin et al., 2010). Rewriting Eq. 18 for only the chosen strategy s_{it} highlights the role the striatum may play in a general EWA formulation:

$$A_{it}^{s_{it}} = \frac{(1+\phi)N_{i,t-1}A_{i,t-1}^{s_{it}} + \Delta_{it}}{N_t}, \tag{16}$$

where

$$\Delta_{it} = \left[\pi\left(s_{it}, s'_{it}\right) - A_{i,t-1}^{s_{it}}\right], \tag{17}$$

and Δ_{it} is the dopaminergenic response system analyzed in Caplin et al. (2010). Left unspecified here is the means by which all action values for unchosen actions, $s \neq s_{it}$, are updated (see Lohrenz et al., 2007).

Other important components associated with reinforcement learning models are also encoded in a network of cortical structures that send projections to the SCi. In contrast to the SCi, the signals carried by these cortical structures are much more heterogeneous across individual neurons. A proportion of neurons in the dorsolateral prefrontal cortex (Barraclough et al., 2004), dorsal anterior cingulated cortex (Seo and Lee, 2008) and lateral intraparietal cortex (Platt and Glimcher, 1999; Dorris and Glimcher, 2004; Seo et al., 2009) encode relevant information necessary to construct action value such as past choices, opponent's choices, the animal's reward history, as well as functions of action value. Like the SCi, some cortical signals display serial dependencies over trials (Seo and Lee, 2007).

ROLE OF THE SCi WITHIN THE SACCADE DECISION CIRCUIT

We propose that the SCi is involved in three important aspects of selecting strategic saccades:

1. *integrating* value related inputs and tagging action values to particular saccade vectors;
2. *selecting* a saccade in a process where action value representations are compared;
3. *providing feedback* of choices to dopaminergic centres.

First, as outlined in Section "Neuronal Correlates of Experience Weighted Attraction", the SCi receives inputs from regions that encode functional elements of action value learning models. Because the SCi integrates many inputs, and outputs to pre-motor neurons, its representations of action value may be particularly suited for choosing final actions. Moreover, the topographic organization of the SCi allows value representations to be tagged to particular saccade vectors.

Second, the SCi provides a platform where multiple action value representations can compete and ultimately be resolved to choose a particular action. The topographic map within the SCi is organized based on the principle of local excitation and distant inhibition (Munoz and Istvan, 1998; Trappenberg et al., 2001; Dorris et al., 2007). Once activity reaches a certain threshold level on this map, a saccade command is sent to pre-motor neurons in the brainstem (see Moschovakis and Highstein, 1994 for review). Therefore, the SCi is perhaps the last site within the visuosaccadic circuit where action value can be represented to influence saccade selection without directly triggering (or necessarily resulting in) saccades.

Third, the SCi sends direct mono-synaptic projections to dopaminergic neurons in the substantia nigra and ventral tegmental area (Comoli et al., 2003; Dommett et al., 2005). Therefore, the SCi may provide feedback on selected actions, thus providing a critical component for the reinforcement learning circuitry of the striatum.

CONCLUSION

Our results suggest that the evolutionarily old SCi does not simply execute sensory-driven reflexive saccades but also encodes action value signals that can be used to select voluntary, strategic saccades. As would be expected from a brain region involved in the decision process, SCi activity simultaneously reflects past choices and their outcomes, and predicts future choice. Similarly, learning models, such as EWA, recursively compute action values from past events to probabilistically choose future actions. We demonstrate that these small trial-to-trial fluctuations in SCi activity are not entirely random but have serial dependencies which can be captured, in part, by the EWA learning model.

APPENDIX

The goal of EWA learning is to construct a model that predicts play across a wide variety of games yet retains a framework that is psychologically sound. In an EWA learning model, each strategy has an attraction (which we re-labelled action value) that is updated based on observed choices and the payoff structure.

We introduce EWA in the context of a player who faces a single opponent. Each period the player chooses s from one of two alternatives, $s \in \{in, out\}$. For each trial t, the subject makes a choice s_t, the opponent chooses $s_t' \in \{in, out\}$, and the subject receives a payoff $\pi_t(s_t, s_t')$ as defined in Section "EWA Learning". We drop the experiment subscript i here for illustration.

Once a choice is made and payoff received in trial t, the attraction of strategy s in trial t is defined as a recursive function of past attractions, choices, and rewards by means of

$$A_t^s = \frac{\phi_t N_{t-1} A_{t-1}^s + [\delta + (1-\delta) 1_{[s=s_t]}]\pi(s, s_t')}{N_t} \tag{18}$$

where $A_t^s \in A_t$ and $A_1^s = 0.5$ at the beginning of each experiment.

The first component of Eq. 18, $\phi N_t A_{t-1}^s$ is a depreciation of the previous period's action value. The second component, $[\delta + (1-\delta)1_{[s=s_t]}]\pi(s, s_t')$, is determined by the choice and reward of the current trial. The experience weight N_t is given by

$$N_t = \rho N_{t-1} + 1 \tag{19}$$

with $N_0 = 0$. Assuming $\rho < 1$, $\lim_{t \to \infty} N_t = \frac{1}{1-\rho}$.

Finally, after a choice is made and a reward is determined in trial t, A_t^s is updated to reflect the valuation of every candidate choice in trial t. On a given trial, the probability of choosing $s_{i,t} = in$ is defined as

$$P(s_{i,t} = in) = \frac{e^{\lambda A_{t-1}^{in}}}{e^{\lambda A_{t-1}^{in}} + e^{\lambda A_{t-1}^{out}}}, \tag{20}$$

which yields a likelihood function for our observed choices

$$L = \prod_{i=1}^{I} \prod_{t=1}^{T_i} P(s_{i,t} = in) \tag{21}$$

which is estimated via maximum likelihood using the log-likelihood function

$$LL = \sum_{i=1}^{I} \sum_{t=1}^{T_i} 1_{[s_t = in]} P(s_{i,t} = in) + 1_{[s_t = out]}(1 - P(s_{i,t} = in)). \tag{22}$$

In addition to the identification restrictions detailed in Ho et al. (2008), we had to make an additional identification assumption for monkey H. We found that the restriction $N_0 = \frac{1}{1-\rho}$ was always binding, so we restricted $N_0 = 1$ for this monkey to ensure identification of ρ, although the estimates are robust to $N_0 \leq 1$.

ACKNOWLEDGMENTS

This work was supported by a Career Development Award from the Human Frontier Science Program (HFSP), a Discovery Grant from the National Science and Engineering Research Council (NSERC) of Canada and a group grant from the Canadian Institutes of Health Research (CIHR) awarded to MCD. We thank J. Green, S. Hickman, M. Lewis, F. Paquin and R. Pengelly for technical assistance. J. Turner provided programming expertise and E. Ryklin customized the data acquisition program. D. Byrne and D. Standage provided constructive feedback regarding the manuscript.

REFERENCES

Anderson, J. R., Reder, L. M., and Lebiere, C. (1996). Working memory: activation limitations on retrieval. *Cogn. Psychol.* 30, 221–256.

Barraclough, D. J., Conroy, M. L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* 7, 404–410.

Bayer, H. M., and Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141.

Brown, J., and Rosenthal, R. (1990). Testing the minimax hypothesis: a re-examination of O'Neill's game experiment. *Econometrica* 58, 1065–1081.

Callicott, J. H., Mattay, V. S., Bertolino, A., Finn, K., Coppola, R., Frank, J. A., Goldberg, T. E., and Weinberger, D. R. (1999). Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cereb. Cortex* 9, 20–26.

Camerer, C., and Ho, T. (1999). Experienced-weighted attraction learning in normal form games. *Econometrica* 67, 827–874.

Caplin, A., Dean, M., Glimcher, P. W., and Rutledge, R. (2010). Measuring beliefs and rewards: a neuroeconomic approach. *Q. J. Econ.* (forthcoming).

Carello, C. D., and Krauzlis, R. J. (2004). Manipulating intent: evidence for a causal role of the superior colliculus in target selection. *Neuron* 43, 575–583.

Comoli, E., Coizet, V., Boyes, J., Bolam, J. P., Canteras, N. S., Quirk, R. H., Overton, P. G., and Redgrave, P. (2003). A direct projection from superior colliculus to substantia nigra for detecting salient visual events. *Nat. Neurosci.* 6, 974–980.

Dommett, E., Coizet, V., Blaha, C. D., Martindale, J., Lefebvre, V., Walton, N., Mayhew, J. E., Overton, P. G., and Redgrave, P. (2005). How visual stimuli activate dopaminergic neurons at short latency. *Science* 307, 1476–1479.

Dorris, M. C., and Glimcher, P. W. (2004). Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44, 365–378.

Dorris, M. C., Olivier, E., and Munoz, D. P. (2007). Competetive integration of visual and preparatory signals in the superior colliculus during saccadic programming. *J. Neurosci.* 27, 5053–5062.

Dorris, M. C., Pare, M., and Munoz, D. P. (1997). Neuronal activity in monkey superior colliculus related to the initiation of saccadic eye movements. *J. Neurosci.* 17, 8566–8579.

Dorris, M. C., Pare, M., and Munoz, D. P. (2000). Immediate neural plasticity shapes motor performance. *J. Neurosci.* 20, 1–5.

Erev, I., and Roth, A. (1998). Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am. Econ. Rev.* 88, 848–881.

Fudenberg, D., and Levine, D. (1998). The Theory of Learning in Games. Cambridge, MIT Press.

Glimcher, P. W., and Sparks, D. L. (1992). Movement selection in advance of action in the superior colliculus. *Nature* 355, 542–545.

Hanes, D. P., and Wurtz, R. H. (2001). Interaction of the frontal eye field and superior colliculus for saccade generation. *J. Neurophysiol.* 85, 804–815.

Harsanyi, J. (1973). Games with randomly disturbed payoffs: a new rationale for mixed-strategy equilibrium points. *Int. J. Game Theory,* 2, 1–23.

Ho, T., Camerer, C., and Chong, J. (2007). Self-tuning experience weighted attraction learning in games. *J. Econ. Theory* 133, 177–198.

Ho, T., Wang, X., and Camerer, C. (2008). Individual differences in EWA learning with partial payoff information. *Econ. J.* 118, 37–59.

Jocham, G., Neumann, J., Klein, T. A., Danielmeier, C., and Ullsperger, M. (2009). Adaptive coding of action values in the human rostral cingulate zone. *J. Neurosci.* 29, 7489–7496.

Juttner, M., and Wolf, W. (1992). Occurrence of human express saccades depends on stimulus uncertainty and stimulus sequence. *Exp. Brain Res.* 89, 678–681.

Kennerley, S. W., Walton, M. E., Behrens, T. E. J., Buckley, M. J., and Rushworth M. F. S. (2006). Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* 9, 940–947.

Lau, B., and Glimcher, P. W. (2005). Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* 84, 555–579.

Lau, B., and Glimcher, P. W. (2007). Action and outcome encoding in the

primate caudate nucleus. *J. Neurosci.* 27, 14502–14514.
Lau, B., and Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron* 58, 451–463.
Lee, D., Conroy, M. L., McGreevy, B. P., and Barraclough, D. J. (2004). Reinforcement learning and decision making in monkeys during a competitive game. *Cogn. Brain Res.* 22, 45–58.
Lohrenz, T., McCabe, K., Camerer, C. F., and Montague, P. R. (2007). Neural signature of fictive learning signals in a sequential investment task. *Proc. Natl. Acad. Sci. U.S.A.* 104, 9493–9498.
Luk, C. H., and Wallis, J. D. (2009). Dynamic encoding of responses and outcomes by neurons in medial prefrontal cortex. *J. Neurosci.* 29, 7526–7539.
Maljkovic, V., and Nakayama, K. (1994). Priming of pop-out.1. Role of features. *Mem. Cognit.* 22, 657–672.
McCabe, K., Mukherji, A., and Runkle, D. (2000). An experimental study of information and mixed-strategy play in the three-person matching-pennies .... *Econ. Theory*, 15, 421–462.
McPeek, R. M., and Keller, E. L. (2004). Deficits in saccade target selection after inactivation of superior colliculus. *Nat. Neurosci.* 7, 757–763.
Mookherjee, D., and Sopher, B. (1994). Learning behavior in an experimental matching pennies game. *Games Econ. Behav.* 7, 62–91.
Mookherjee, D., and Sopher, B. (1997). Learning and decision costs in experimental constant sum games. *Games Econ. Behav*. 19, 97–132.
Moschovakis, A. K., and Highstein, S. M. (1994). The anatomy and physiology of primate neurons that control rapid eye movements. *Annu. Rev. Neurosci.* 17, 465–488.
Munoz, D. P., and Istvan, P. J. (1998). Lateral inhibitory interactions in the intermediate layers of the monkey superior colliculus. *J. Neurophysiol.* 79, 1193–1209.
Nash, J. (1951). Non-cooperative games. *Ann. Math. Second Series* 54, 286–295.
Ochs, J. (1995). Games with unique, mixed strategy equilibria: an experimental study. *Games Econ. Behav.* 10, 202–217.
O'Neill, B. (1987). Nonmetric test of the minimax theory of two-person zerosum games. *Proc. Natl. Acad. Sci. U.S.A.* 84, 2106–2109.
Padoa-Schioppa, C., and Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* 441, 223–226.
Platt, M. L., and Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.
Rapoport, A., and Boebel, R. (1992). Mixed strategies in strictly competitive games: a further test of the minimax hypothesis. *Games Econ. Behav.* 4, 261–283.
Rapoport, A., and Budescu, D. (1992). Generation of random series in two-person strictly competitive games. *J. Exp. Psychol. Gen.* 121, 352–363.
Robinson, D. A. (1972). Eye-movements evoked by collicular stimulation in alert monkey. *Vision Res.* 12, 1795–1808.
Rushworth, M. F., Walton, M. E., Kennerley, S. W., and Bannerman, D. M. (2004). Action sets and decisions in the medial frontal cortex. *Trends Cogn. Sci.* 8, 410–417.
Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.
Schiller, P. H., and Stryker, M. (1972). Single-unit recording and stimulation in superior colliculus of alert rhesus-monkey. *J. Neurophysiol.* 35, 915–924.
Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.
Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Curr. Opin. Neurobiol.* 14, 139–147.
Seo, H., Barraclough, D. J., and Lee, D. (2007). Dynamic signals related to choices and outcomes in the dorslateral prefrontal cortex. *Cereb. Cortex* 17, I110–I117.
Seo, H., Barraclough, D. J., and Lee, D. (2009). Lateral intraparietal cortex and reinforcement learning during a mixed-strategy game. *J. Neurosci.* 29, 7278–7289.
Seo, H., and Lee, D. (2007). Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J. Neurosci.* 27, 8366–8377.
Seo, H., and Lee, D. (2008). Cortical mechanisms for reinforcement learning in competitive games. *Philos. Trans. R. Soc.* 363, 3845–3857.
Spiliopoulos, L. (2008). Humans versus computer algorithms in repeated mixed strategy games. *MPRA Paper No.* 6672, 68.
Sugrue, L. P., Corrado, G. S., and Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science* 304, 1782–1787.
Sutton, R., and Barto, A. G. (1998). Reinforcment Learning: An Introduction. Cambridge, MIT Press.
Thevarajah, D., Mikulic, A., and Dorris, M. (2009). Role of the superior colliculus in choosing mixed-strategy saccades. *J. Neurosci.* 89, 1998–2008.
Thompson, K. G., Hanes, D. P., Bichot, N. P., and Schall, J. D. (1996). Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *J. Neurophysiol.* 76, 4040–4055.
Thorndike, E. L. (1898). Animal intelligence: An Experimental Study of the Associative Processes in Animals. New York, Macmillan.
Trappenberg, T., Dorris, M., and Munoz, D. (2001). A model of saccade initiation based on the competitive integration of exogenous and .... *J. Cogn. Neurosci.* 13, 256–271.
von Neumann, J., and Morgenstern, O. (1947). Theory of Games and Economic Behaviour. Princeton University Press, Princiton.
Watkins, C. (1989). Learning from Delayed Rewards. Ph.D. thesis, Cambridge University, Cambridge.
Wooldridge, J. M. (2001). Econometric Analysis of Cross Section and Panel Data. Cambridge, MIT Press.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 07 July 2009; paper pending published: 04 August 2009; accepted: 01 December 2009; published online: 08 February 2010.
*Citation: Thevarajah D, Webb R, Ferrall C and Dorris MC (2010) Modeling the value of strategic actions in the superior colliculus. Front. Behav. Neurosci. **3**:57. doi: 10.3389/neuro.08.057.2009*

frontiers in
BEHAVIORAL NEUROSCIENCE

REVIEW ARTICLE
published: 11 August 2009
doi: 10.3389/neuro.08.015.2009

Dopamine regulation of social choice in a monogamous rodent species

Brandon J. Aragona[1]* and Zuoxin Wang[2]

[1] *Department of Psychology and Program in Neuroscience, University of Michigan, Ann Arbor, MI, USA*
[2] *Department of Psychology and Program in Neuroscience, Florida State University, Tallahassee, FL, USA*

Edited by:
Paul Phillips,
University of Washington, USA

Reviewed by:
Larry J. Young, Yerkes National Primate Research Center, USA; Emory University School of Medicine, USA
Jeansok J. Kim,
University of Washington, USA
Scott A. Huettel, Duke University, USA

****Correspondence:***
Brandon J. Aragona, Department of Psychology and Program in Neuroscience, University of Michigan, 530 Church Street, Ann Arbor, MI 48109, USA.
e-mail: aragona@umich.edu

There is growing appreciation that social decision making in humans is strongly influenced by hedonic and emotional processing. The field of social neuroeconomics has shown that neural systems important for reward are associated with social choice and social preferences in humans. Here, we show that the neurobiology of social preferences in a monogamous rodent species, the prairie vole, is also regulated by neural systems involved in reward and emotional processing. Specifically, we describe how mesolimbic dopamine transmission differentially mediates the formation and maintenance of monogamous pair bonds in this species. Thus, reward processing exerts tremendous regulation over social choice behaviors that serve as the foundation of a rather complex social organization. We conclude that prairie voles are an excellent model system for the neuroscience of social choice and that complex social decision-making can be robustly explained by reward and hedonic processing.

Keywords: nucleus accumbens, prairie vole, monogamy, social attachment, pair bond, social decision making, social neuroeconomics

INTRODUCTION

In social contexts, decision-making is significantly influenced by positive or negative concern for the welfare of others (Fehr and Camerer, 2007). Humans display strong social preferences that are revealed through choice behavior in which people behave altruistically, act on a strong sense of fairness, and have tremendous capacities to trust (Krueger et al., 2007; Sanfey, 2007; Tankersley et al., 2007; Zak et al., 2004). Indeed, social decision-making in humans is so complex that it can appear to be the result of social cognition that is exclusive to our species (Skuse and Gallagher, 2009). However, from an evolutionary perspective, pro-social behaviors such as cooperation and trust are only ostensibly irrational or selfless (Rilling et al., 2002; Sanfey, 2007). Such behaviors are the result of selection processes that favored reciprocity among close social groups, in which it was adaptive for individuals to spend relatively small amounts of energy to help unrelated members of the group in order to receive relatively large benefits of the resulting social organization (Pfeiffer et al., 2005; Rutte and Taborsky, 2007; Trivers, 1971). From this perspective, we can expect analogous pro-social behaviors to be expressed by other species that can serve as effective laboratory models and thus allow the investigation of the neural mechanisms of social choice behavior and decision-making.

Here, we describe how the use of one such model system, the socially monogamous prairie vole (*Microtus ochrogaster*), has significantly advanced our understanding of the neural regulation of social choice behavior (Carter et al., 1995; Dewsbury, 1987; Getz and Carter, 1996; Young and Wang, 2004). We first provide a brief overview of prairie vole behavior and suggest that the complex social organization of this species can be largely achieved by two 'choice' behaviors: the initial preference of a familiar mate and the decision to avoid or aggressively reject potentially new mates (Carter and Getz, 1993; Getz and Hofmann, 1986; Insel and Young, 2001). We then highlight data from several recent studies that describe the regulation of prairie vole social behavior by neural transmission important for emotion and reward processing, dopamine (DA) signaling within the nucleus accumbens (NAc) (Aragona and Wang, 2007; Aragona et al., 2003, 2006). Finally, we compare these findings to studies that have examined the neural regulation of social decision-making in humans (Fisher et al., 2005; Kosfeld et al., 2005; Rilling et al., 2002). These comparisons reveal striking similarities between the neuroscience of social choice behaviors between humans and prairie voles, suggesting that prairie voles are an excellent model system for the study of social decision-making. Moreover, the fact that a rather large extent of the social organization of prairie voles can be largely explained by rather simple choice behaviors regulated by emotional processing may have very interesting implications for the study of social neuroeconomics (Cacioppo et al., 2000; Lee, 2008).

THE PRAIRIE VOLE MODEL

Prairie voles are small rodents (~40 g) (**Figure 1A**) distributed primarily in the grasslands of the central United States (Cushing et al., 2001; Hall, 1981; Hoffmann and Koeppl, 1985). These rodents are among the minority of mammalian species (3–5%) that show a monogamous social organization (Dewsbury, 1987). The foundation of this social organization is the 'pair bond', which is defined as the stable relationship between members of a breeder pair that share common territory and parental duties (Aragona and Wang, 2004). This species was initially identified as monogamous by field studies which showed that male–female pairs travel together (Getz et al., 1981), share a nest with one or more litters of pups (Getz and Hofmann, 1986), and aggressively repel unrelated intruders from

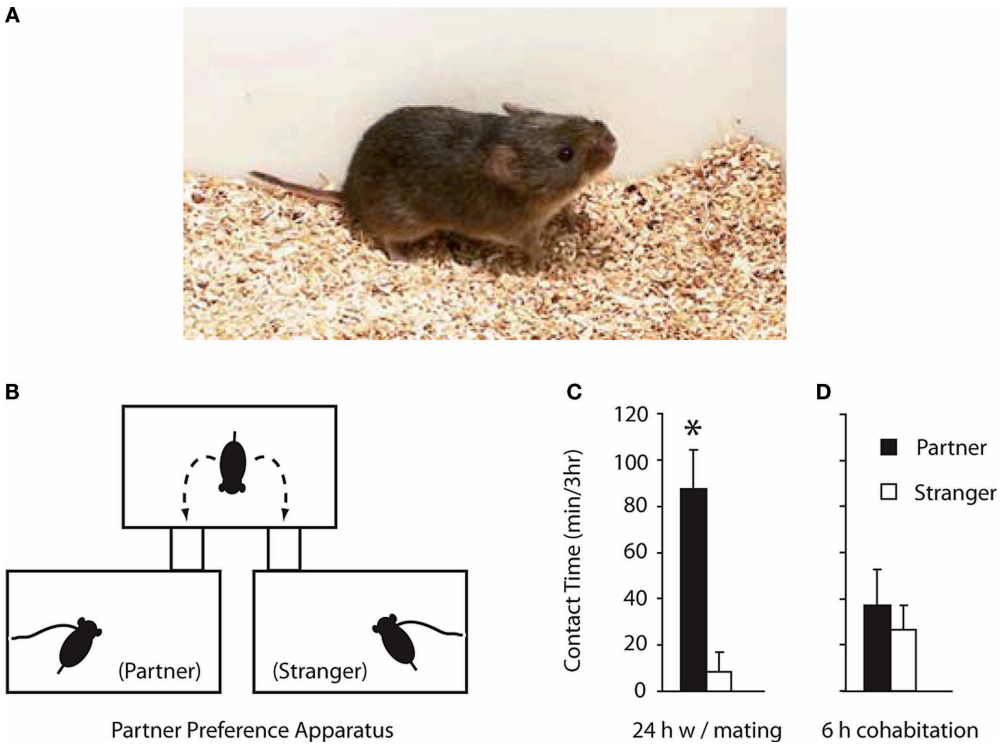


FIGURE 1 |The prairie vole model. (A) Photo of an adult male prairie vole. **(B)** Cartoon of partner preference apparatus. Each cages is identical and food and water are available *ad libitum* throughout the 3-h test. **(C)** Male prairie voles paired with an estrogen-primed female for 24 h show a robust partner preference, i.e. spend significantly more time in side-by-side contact with their familiar mates (partners) compared to novel females that are also estrogen primed (strangers). **(D)** Male prairie voles paired with an ovariectomized female that is not estrogen primed for only 6 h do not show partner preferences; i.e. they display non-selective side-by-side contact. Error bars = standard error and * indicates groups are significantly different as determined by a *t*-test.

their territory (Getz, 1978). Further, male prairie voles show high levels of parental care (Getz and Carter, 1996; Thomas and Birney, 1979) and it has been suggested that both parents are necessary for pup survival which selected for highly enduring pair bonds (Emlen and Oring, 1977; Kleiman, 1977; McGuire et al., 1993; Wang and Novak, 1992). Indeed, the pair bond is so stable that a surviving member of the pair will not accept a new mate even if the other member of the bond is lost (Getz and Carter, 1996; Thomas and Wolff, 2004). This represents a strong example of behavior that is not in the self-interest of the animal and is therefore in conflict with classic economic models of rational decision-making.

Importantly, the monogamous behaviors observed in nature are also reliably expressed under laboratory conditions (Carter and Getz, 1993; Carter et al., 1995). For instance, prairie voles preferentially mate with a familiar partner versus a novel conspecific (Dewsbury, 1975, 1987; Gray and Dewsbury, 1973). After mating, prairie voles remain together during gestation (McGuire and Novak, 1984; Thomas and Birney, 1979) and this facilitates a successful pregnancy (McGuire et al., 1992). As in their natural environment, male prairie voles show very high levels of parental care in the lab (Oliveras and Novak, 1986). Most importantly, pair bonding can be reliably assessed in the lab by measuring social preferences inferred from choice behaviors associated with the formation and maintenance of the pair bond (Williams et al., 1992; Winslow et al., 1993; Young and Wang, 2004).

LABORATORY TESTS OF PAIR BOND FORMATION AND MAINTENANCE

This review will focus on data collected from male subjects (Aragona and Wang, 2007; Aragona et al., 2003, 2006). However, there has been extensive work conducted on female prairie voles (Cho et al., 1999; Fowler et al., 2002; Insel and Hulihan, 1995; Williams et al., 1992; Witt et al., 1991) and it will be noted when data were collected using female subjects. A necessary first step in pair bond formation is that males must prefer their familiar partner over new mates, which is very unusual for males in most mammalian species since they reliably prefer to mate with novel females (Fiorino et al., 1997). However, male prairie voles prefer to mate with a familiar female (Dewsbury, 1987) and the presentation of new females does not induce copulation in sexually satiated male prairie voles (Gray and Dewsbury, 1973).

In addition to choosing to mate with a familiar female, pair bonding also requires that males choose to cohabitate with their familiar partners. This is determined in the lab by a simple social choice test referred to as the 'partner preference test' (Williams et al., 1992). For this test, a subject is placed into a three-chambered apparatus and is free to move about the chambers (**Figure 1B**). The familiar mate (partner) and an unfamiliar female (stranger) serve as stimulus animals that are tethered in separate cages (**Figure 1B**). Subjects initially explore the apparatus and interact with both stimulus animals and then lay down beside either the partner or the

stranger (Williams et al., 1992; Winslow et al., 1993). If subjects spend significantly more time in side-by-side contact with partners over strangers (assessed by a *t*-test) then the group is said to show a partner preference (Aragona and Wang, 2004; Curtis and Wang, 2005; Liu et al., 2001).

Many studies have demonstrated that male prairie voles paired with an estrogen-primed female for 24 h of mating reliably show partner preferences (Aragona et al., 2003; Lim and Young, 2004; Liu et al., 2001) (**Figure 1C**). However, if male subjects cohabitate with females for only 6 h without mating, subjects show non-selective side-by-side contact and thus fail to show partner preferences (Aragona and Wang, 2007; Curtis and Wang, 2005; Liu et al., 2001) (**Figure 1D**). Thus, we utilize the '24 h mating' paradigm to reliably induce partner preferences in control conditions and examine if pharmacological manipulations can prevent mating-induced pair bond formation. Additionally, we use the '6-h cohabitation' paradigm to examine if pharmacological manipulations can induce partner preferences in the absence of mating (Wang and Aragona, 2004; Young and Wang, 2004).

While a partner preference is necessary for a pair bond, it is not sufficient for its long-term maintenance. Pair bonded males also choose to aggressively reject potentially new mates (Aragona et al., 2006; Gobrogge et al., 2007). This is referred to as 'selective aggression' and is studied in the lab using a resident-intruder test in which the subject is exposed to novel conspecifics and aggressive behavior is quantified (Wang et al., 1997; Winslow et al., 1993). While 24 h of mating increases selective aggression (Wang et al., 1997; Winslow et al., 1993), aggressive behavior is increased much more toward male intruders (compared to novel females) and male subjects do not chase or bite female intruders following 24 h of mating (Wang et al., 1997). Conversely, following an extended cohabitation (2 weeks) in which females become pregnant, males become extremely aggressive toward novel females (showing high levels of chasing and biting) (Aragona et al., 2006; Gobrogge et al., 2007) and this decision to aggressively reject potentially new mates is critical for the stable maintenance of the pair bond.

In this review, we will consider the extent to which the monogamous social organization of prairie voles can be explained by (1) the initial choice to breed with a single female, the 'partner preference' and (2) the subsequent choice to reject potential new mates, selective aggression. Having these well-established laboratory indices allows detailed examination of the neurobiology underlying these behaviors. As pair bonding involves a myriad of cognitive and psychological processes, it is not surprising that a wide range of neural systems are important for its regulation including: oxytocin (Bales et al., 2007; Bamshad et al., 1993; Insel and Shapiro, 1992; Liu and Wang, 2003; Witt et al., 1990), vasopressin (Bamshad et al., 1994; Hammock and Young, 2005; Lim et al., 2004b; Liu et al., 2001; Winslow et al., 1993), corticosterone (DeVries et al., 1995, 1996; Lim et al., 2007), estrogen (Cushing and Wynne-Edwards, 2006), glutamate and GABA (Curtis and Wang, 2005). This list will certainly grow as more experiments are conducted and almost nothing is known about how these systems interact to regulate pair bonding. Thus, an extraordinary amount of work remains. However, we have recently conducted a series of studies demonstrating the significant involvement of mesolimbic DA transmission in pair bond formation and maintenance in male prairie voles (Aragona and Wang, 2007; Aragona et al., 2003, 2006).

NUCLEUS ACCUMBENS DOPAMINE AND PAIR BOND FORMATION

Pair bond formation is a naturally occurring association formed between monogamous mates (Aragona et al., 2006; Wang and Aragona, 2004; Young and Wang, 2004) and associative learning is significantly regulated by mesolimbic DA transmission (Di Chiara and Bassareo, 2007; Kelley, 2004; Wise, 2004). In particular, DA transmission within the NAc is critical for important aspects of reward processing (Berridge and Robinson, 2003; Everitt and Robbins, 2005; Roitman et al., 2005, 2008; Salamone and Correa, 2002; Wheeler et al., 2008) that may underlie cost-benefit analyses related to choice behavior and decision-making (Phillips et al., 2007). Therefore, we conducted a series of studies that investigated the regulation of partner preference formation by DA transmission within the NAc (Aragona and Wang, 2007; Aragona et al., 2003, 2006).

Similar to other rodent species (Jansson et al., 1999), prairie vole NAc is densely innervated by dopaminergic terminals arising from the ventral midbrain (**Figure 2A**) (Aragona et al., 2003; Curtis and Wang, 2005; Gobrogge et al., 2007). Also consistent with studies conducted in rats (Becker et al., 2001; Pfaus et al., 1995; Robinson et al., 2002), microdialysis measures indicate that mating increases extracellular DA concentration within the NAc of female prairie voles (Gingrich et al., 2000) and tissue extraction studies show that mating also increases dopamine transmission (as indicated by dopamine turnover) in male prairie voles (**Figure 2B**) (Aragona et al., 2003). These studies suggest that mating evokes modest increases in DA concentration within the NAc during copulation in prairie voles.

We hypothesized that mating-evoked increases in DA transmission were necessary for partner preference formation (Aragona et al., 2003). To test this, we first examined if blockade of DA receptors within the NAc prevented mating-induced partner preferences (**Figure 2C**). Consistent with previous studies (Williams et al., 1992; Winslow et al., 1993), control animals that received microinfusions of artificial cerebrospinal fluid (CSF) within the NAc prior to the 24-h cohabitation period (with mating) showed robust mating-induced partner preferences (**Figure 2C**). However, blockade of DA receptors with the non-selective DA receptor antagonist (haloperidol) prior to the mating period, abolished mating-induced partner preference formation (**Figure 2C**). Importantly, DA receptor blockade did not alter locomotor activity or mating behavior, indicating that DA transmission within the NAc during mating directly influenced social choice that was a consequence of mating (Aragona et al., 2003).

We next tested if pharmacological activation of DA receptors within the NAc was sufficient to induce partner preference formation in the absence of mating (Aragona et al., 2003). As previously described (Williams et al., 1992; Winslow et al., 1993), control subjects that received CSF infusions into the NAc prior to the 6-h cohabitation period did not show partner preferences (**Figure 2D**). However, low dose infusion of the non-selective DA agonist (apomorphine) induced a significant partner preference, whereas high dose infusion of apomorphine did not (**Figure 2D**). These data

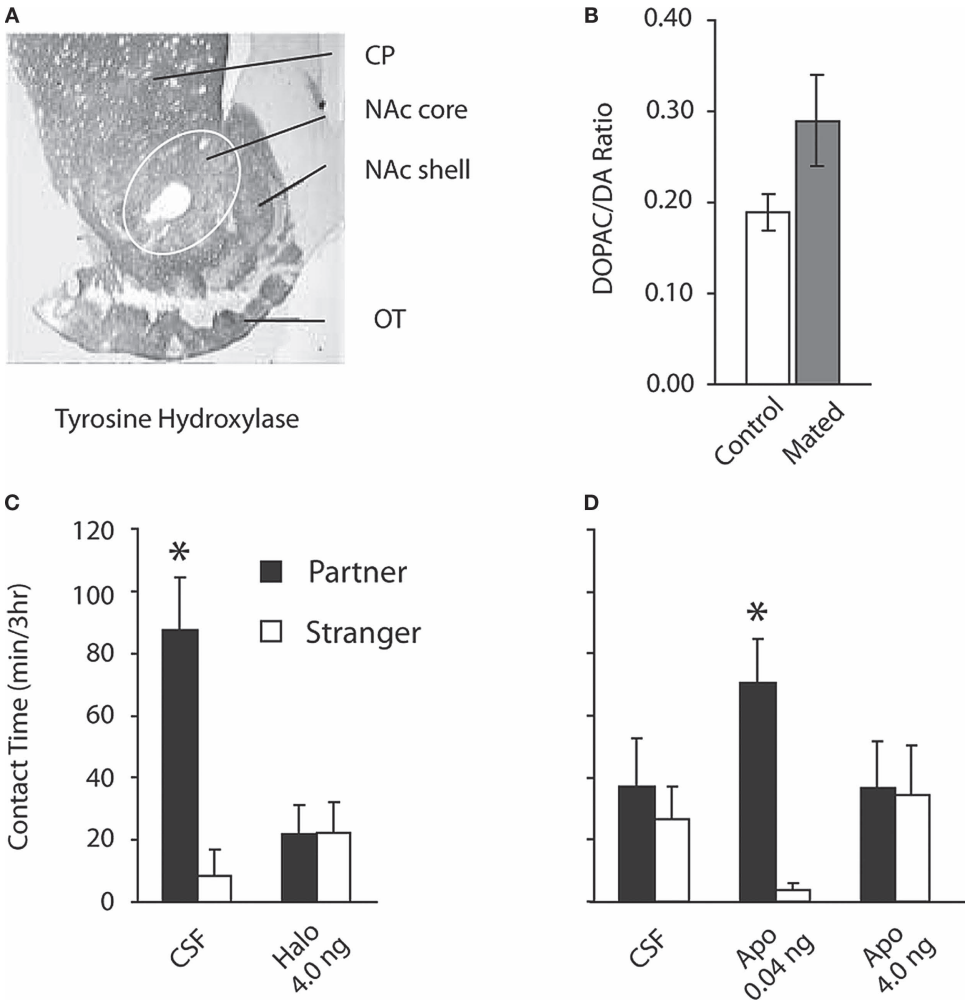


FIGURE 2 | Dopamine regulation of pair bond formation. (A) Coronal section showing tyrosine hydroxylase immunocytochemical labeling of dorsal and ventral striatum from an adult male prairie vole. CP = caudate putamen, NAc = nucleus accumbens NAc shell, OT = olfactory tuberacle. **(B)** Dopamine turnover as indicated by increased concentration of the dopamine metabolite DOPAC and decreased concentration of DA from micro-dissected of NAc tissue, chemical extraction, and measurement using HPLC-ED. Male prairie voles show increased mean DA turnover 30 min after mating onset with an estrogen-primed female. **(C)** Blockade of DA receptors within the NAc by micro-infusion of haloperidol (Halo) prevented mating-induced partner preference formation. **(D)** Micro-infusion of low (0.04 ng) but not high (4.0 ng) dose of apomorphine (Apo) induced partner preferences in the absence of mating.

show that pharmacological activation of DA receptors within the NAc is sufficient to facilitate choice of familiar partners.

OPPOSING REGULATION OF PAIR BOND FORMATION BY D1 AND D2 RECEPTOR SIGNALING PATHWAYS IN THE NAc SHELL

Facilitation of partner preferences by low dose apomorphine is indicative of the receptor specific mechanism underlying DA regulation of this behavior. There are two families of DA receptors: D1-like (D1 and D5 receptors) and D2-like (D2, D3, and D4 receptors) (Neve et al., 2004). While apomorphine binds both D1 and D2-like receptors, it binds D2-like receptors with a much greater affinity (Missale et al., 1998). Thus, we hypothesized that low dose apomorphine preferentially activated D2- but not D1-like receptors and therefore induced partner preference formation via a D2-mediated mechanism in male prairie voles (Aragona et al., 2006). Additionally, the failure of high dose apomorphine to induce partner preferences suggests that activation of D1-like receptors within the NAc actually prevents pair bond formation. These hypotheses were evaluated by testing the effects of receptor specific dopaminergic drugs on our two established paradigms to examine partner preference formation.

Consistent with data from female prairie voles (Gingrich et al., 2000; Wang et al., 1999), specific activation of D2-like receptors within the NAc shell (but not the NAc core) induced partner preferences in the absence of mating (**Figure 3A**). Activation of D1-like receptors within the NAc shell not only failed to induce partner preferences, but also prevented partner preferences induced by D2-like activation (i.e. when D1 and D2 agonists were co-infused) (**Figure 3A**). Importantly, D1-like activation within the NAc shell also blocked mating-induced partner preferences (**Figure 3B**). Together, these data demonstrate that activation of

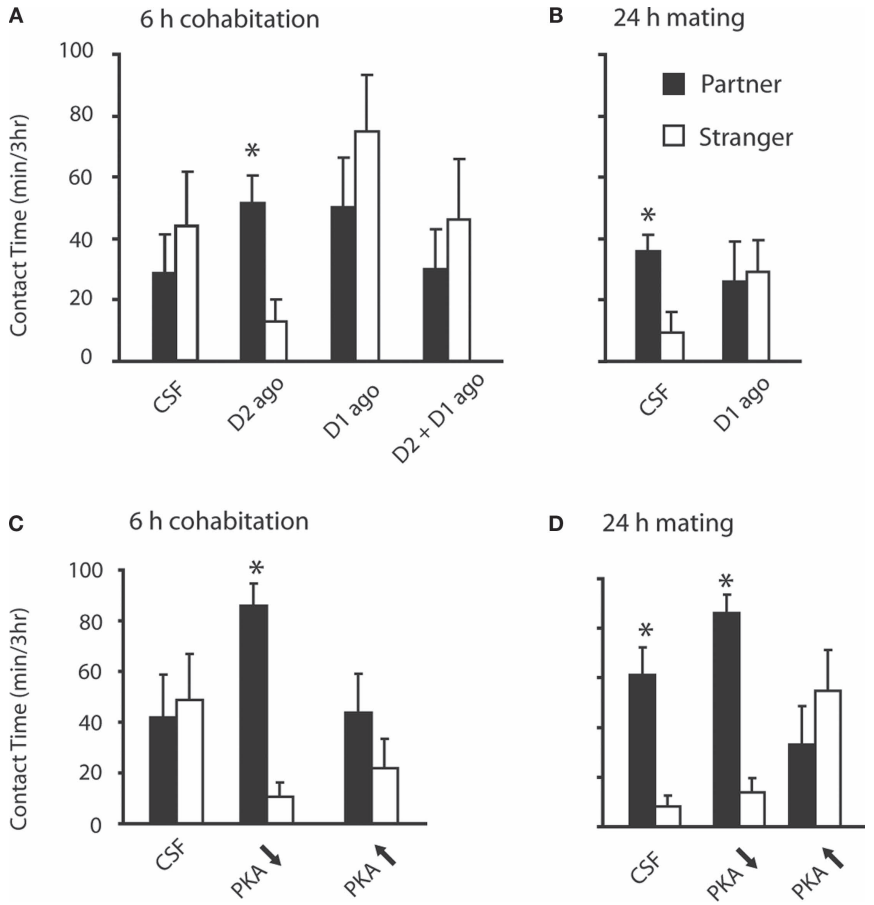


FIGURE 3 | Opposing regulation of pair bond formation by D2- and D1-like dopamine signaling systems within the NAc shell. (A) Activation of D2-like receptors within the NAc shell by micro-infusion of the D2-specific agonist quinpirole (D2 ago) induced partner preferences in the absence of mating. Activation of D1-like receptors within the shell using the D1-specific agonist SKF 38393 (D1 ago) failed to induced partner preference formation and prevented quinpirole-induced partner preferences. **(B)** Activation of D1-like receptors also prevented partner preferences induced by mating. **(C)** Decreased activation of protein kinase A (PKA) using Rp-cAMPS (PKA ↓) induced partner preferences in the absence of mating, whereas activation of PKA using Sp-cAMPS (PKA ↑) did not. **(D)** While decreased activation of PKA using Rp-cAMPS (PKA ↓) did not interfere with mating-induced partner preference formation, activation of PKA using Sp-cAMPS (PKA ↑) interfered with this behavior.

D1-like receptors within the NAc shell prevents the formation of partner preferences.

D1 and D2-like receptors have the opposite effects over cAMP signaling (Neve et al., 2004). D2-like receptors activate inhibitory G-proteins which prevents conversion of ATP to cAMP by adenyl cyclase (Missale et al., 1998). Conversely, activation of D1-like receptors activates stimulatory G-proteins which increases cAMP production and thus activation of protein kinase A (PKA) (Missale et al., 1998). Decreased cAMP production can be studied by pharmacological blockade of cAMP binding sites on PKA using a cAMP analogue (Rp-cAMPS) whereas increased cAMP production is assessed using a cAMP analogue that binds PKA and releases its regulatory subunits (Sp-cAMPS) (Lynch and Taylor, 2005; Self et al., 1998).

Given that D2-like activation within the NAc shell mediates partner preference formation, we hypothesized that reduced PKA activity would also facilitate this behavior. Consistent with D2 regulation of pair bond formation, decreasing the activity of PKA (using Rp-cAMPS) induced partner preferences in the absence of mating (**Figure 3C**). Conversely, increasing activation of PKA (using Sp-cAMPS) failed to induce partner preferences (**Figure 3C**). As expected, decreased PKA activity did not alter mating-induced pair bond formation (**Figure 3D**). However, consistent with D1-like activation preventing pair bond formation, increased activation of PKA prevented mating-induced pair bond formation (**Figure 3D**). Together, these data indicate that pair bond formation is facilitated by D2-like activation and subsequent decreased activity of the cAMP-signaling pathway. Conversely, D1-like activation and subsequent increased activation of PKA prevent pair bond formation.

UP-REGULATION OF D1-LIKE DA RECEPTORS WITHIN THE NAc OF PAIR BONDED ANIMALS

There are dramatic behavioral alterations as male prairie voles transition from sexually naive to fully pair bonded (Carter et al., 1995). Specifically, sexually naive males primarily show pro-social behaviors toward novel females, whereas pair bonded males avoid or

attack novel females. Given the significant role of DA transmission within the NAc in partner preference formation, we expected that alterations in this DA signaling system were associated with behavioral alterations associated with pair bonding (Aragona et al., 2006). We used receptor autoradiography to compare DA receptor density between sexually naive male prairie voles and males that were paired with a female for 2 weeks. During this extended cohabitation males and females shared a nest and the females became pregnant (Aragona et al., 2006). Representative examples of receptor binding clearly demonstrate that D1-like receptors (**Figure 4A**) but not D2-like receptors (**Figure 4B**) are substantially increased within the NAc in pair bonded males. Quantitative data show that D1-like receptor binding was significantly increased within the NAc in pair bonded males compared to that of sibling-paired controls (**Figure 4C**). A separate control group showed that mating alone was not sufficient to increase D1-like receptor binding (Aragona et al., 2006). Thus, pair bonded animals have an enhanced D1-like signaling system within the NAc and since this system is antagonistic to partner preference formation, we next tested if this neural restructuring is responsible for pair bond maintenance.

NEURAL REORGANIZATION WITHIN THE NAc UNDERLIES PAIR BOND MAINTENANCE

Given that pair bonded animals have increased D1-like receptor expression within the NAc and show high levels of aggression toward novel females, we tested if this neural restructuring was associated with increased aggression. Specifically, we used a resident-intruder test to determine if up-regulation of D1-like receptors within the NAc mediates the aggressive rejection of potentially new mates, i.e. selective aggression (Gobrogge et al., 2007; Wang et al., 1997; Winslow et al., 1993). In this test, the female partner was removed from the home cage and both affiliative (**Figure 5A**) and aggressive (**Figure 5B**) behavior of the male subject was examined following introduction of an 'intruder' female (Wang et al., 1997; Winslow et al., 1993). Pair bonded males showed significantly higher levels of affiliative behavior toward their familiar partners compared to that shown by sexually naive males presented with a novel female (**Figure 5C**). While pair bonded males show almost no affiliative behavior toward novel females (strangers) (**Figure 5C**), affiliative behavior is returned to levels expressed by sexually naive subjects if either D2 or D1-like receptors were blocked within the NAc (**Figure 5B**).

Neither sexually naive males presented with a novel female nor pair bonded males presented with their partner showed aggressive behavior (**Figure 5D**). However, pair bonded males were extremely aggressive when presented with novel females (strangers), showing a significant increase in the numbers of attacks (**Figure 5D**). Aggressive behavior was abolished by blockade of D1-like (but not D2-like) receptors within the NAc (**Figure 5D**). These data show that the up-regulation of D1-like receptors described above (**Figure 4**) mediates selective aggression. Thus, plasticity within the mesolimbic DA system underlies the decision to reject potentially new mates and thus maintains the initial pair bond.

SUMMARY OF DOPAMINE REGULATION OVER PAIR BOND FORMATION AND MAINTENANCE

Mesolimbic DA regulation of pair bonding may have implications for cognitive and psychological processes associated with social choice and decision-making. DA transmission that mediates partner preference formation occurs specifically within the rostral portion of the NAc shell (Aragona et al., 2006) (**Figure 6A**). This sub-region is critical for processing positive affect and unconditioned aspects of associative learning (Di Chiara and Bassareo, 2007; Ikemoto, 2007; Pecina et al., 2006). Thus, DA transmission within the NAc shell may regulate partner preference formation through enhanced reward processing or incentive motivation (Berridge, 2007; Di Chiara and Bassareo, 2007). Additionally, DA transmission within the NAc shell is also important for mother–offspring bonds, which is an inherently rewarding social attachment (Champagne et al., 2004; Li and Fleming, 2003; Numan et al., 2005). Together,

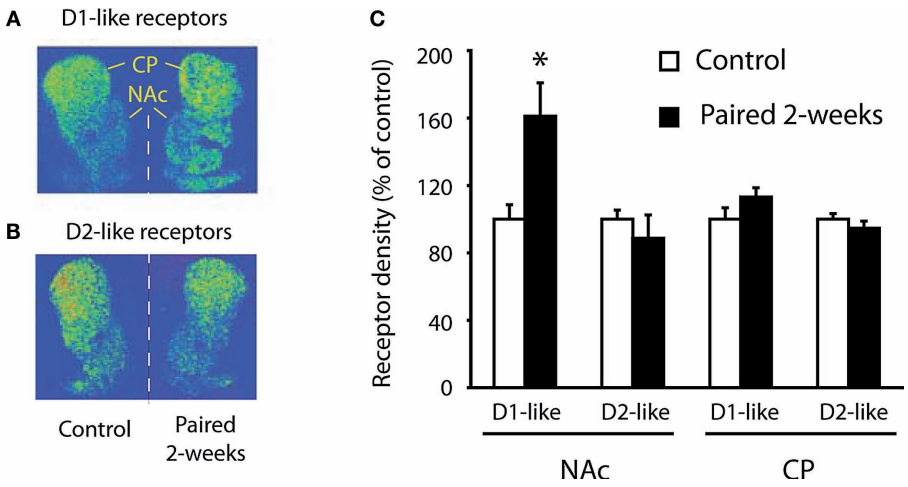


FIGURE 4 | Up-regulation of D1-like receptors within the NAc of pair bonded animals. (A) Representative examples of D1-like receptor binding within the dorsal and ventral striatum of sexually naive adult male prairie voles (left) and pair bonded males (paired with a female for 2 weeks; right). **(B)** Representative examples for D2-like receptor binding. **(C)** Quantification of receptor binding expressed as percent of control subjects. Pair bonded males show a significant increase in D1-like receptor binding within the NAc but not the CP (caudate-putamen). There is no significant change in D2-like receptor binding within either striatal region.

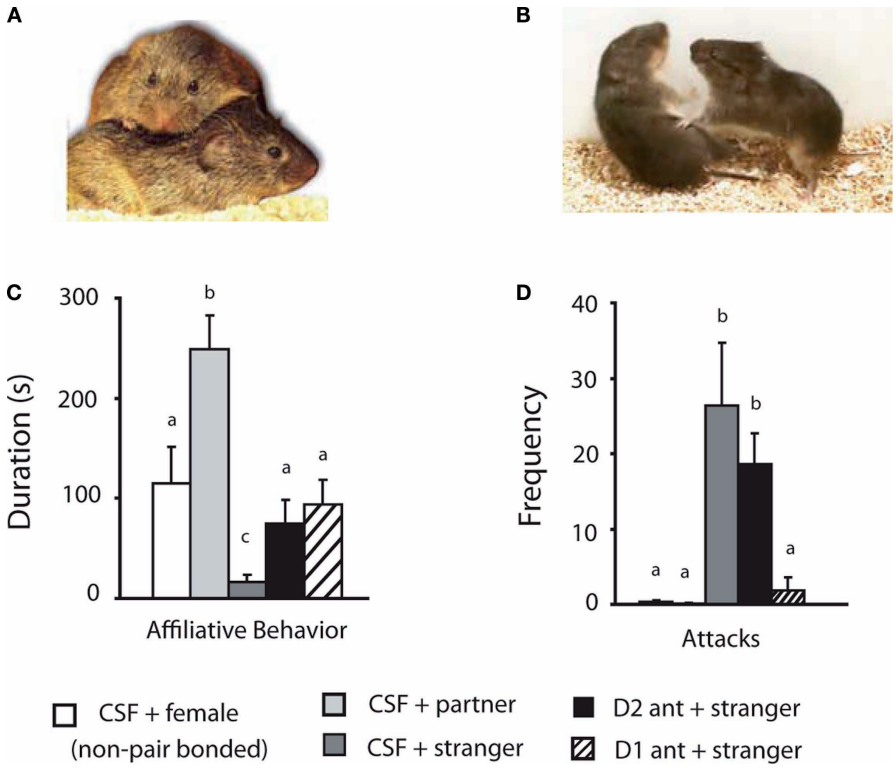


FIGURE 5 | Dopamine regulation of pair bond maintenance as indicated by selective aggression toward novel females. (A) Photo of pair bonded mates engaged in affiliative behavior (typically huddling or side-by-side contact). **(B)** Pair bonded male (right) showing aggressive behavior toward an unfamiliar/novel female (stranger; left). **(C)** Quantification of affiliative behavior during a 6-min resident intruder test of selective aggression. Pair bonded males show significantly more affiliative behavior than other group when presented with their familiar partner, but significantly less affiliation when presented with unfamiliar females (strangers). Blockade of either D1- or D2-like receptors restores affiliative behavior in pair bonded males to levels expressed by sexually naive males being exposed to a female for the first time. **(D)** While sexually naive (presented with a female) and pair bonded males (presented with their partners) show no aggressive behavior, pair bonded males show significantly greater levels of aggression when presented with a novel female (stranger). Selective aggression is blocked by D1-like (but not D2-like) receptors within the NAc.

these data suggest that reward processing is a critical component of partner preference formation in prairie voles.

Within the NAc shell, DA regulation of partner preference formation is highly specific. Mating-induced DA release selectively activates D2-like receptors and decreases cAMP signaling to promote pair bond formation (**Figure 6B**). Conversely, activation of D1-like receptors and increased activation of cAMP signaling prevents pair bond formation (**Figure 6C**). These data indicate that, under natural circumstances, DA transmission is not uniformly increased as it is under certainly laboratory conditionings (Schultz, 2002). Rather, the pair bonding studies suggest that prairie vole social interactions result in modest increases in extracellular DA concentration that selectively activate high affinity D2-like receptors while not activating low affinity D1-like receptors (Richfield et al., 1989). However, it will be necessary for future studies to test this by measuring real-time DA transmission (Aragona et al., 2008; Day et al., 2007; Phillips et al., 2003) during prairie vole social interactions to determine if *in vivo* DA transmission is consistent with the behavioral pharmacology described in this review.

Compared to their basal state (**Figure 6D**), pair bonded males show a robust increase in the surface expression of D1-like receptors within the NAc (**Figure 6E**). We have suggested this may be a compensatory increase following the lack D1-like receptor activation during social interactions that promote pair bond formation (Aragona et al., 2006). Since pair bonded males show an up-regulation in D1-like receptors within the NAc and activation of these receptors prevents pair bond formation, we have suggested that when pair bonded males in their natural environment encounter a novel female, DA is released in very high concentration (Robinson et al., 2002) sufficient to activate low affinity D1-like receptors (Richfield et al., 1989), especially since there appear to be a greater number of antagonistic D1-like receptors in pair bonded voles. This promotes the aggressive rejection of potentially new mates and thus represents an elegant mechanism for maintenance of the initial pair bond. Taken together, these data demonstrate that DA transmission with the NAc differentially mediates initial partner preference formation and the subsequent rejection of potentially novel mates. This is achieved, at least in part, by neuroplasticity (up-regulation of D1-like receptors) within this mesolimbic DA signaling system. This represents a powerful example in which a complex monogamous social organization can be significantly accounted for by two rather straightforward choice behaviors that are both mediated by emotional/reward processing by mesolimbic DA signaling.

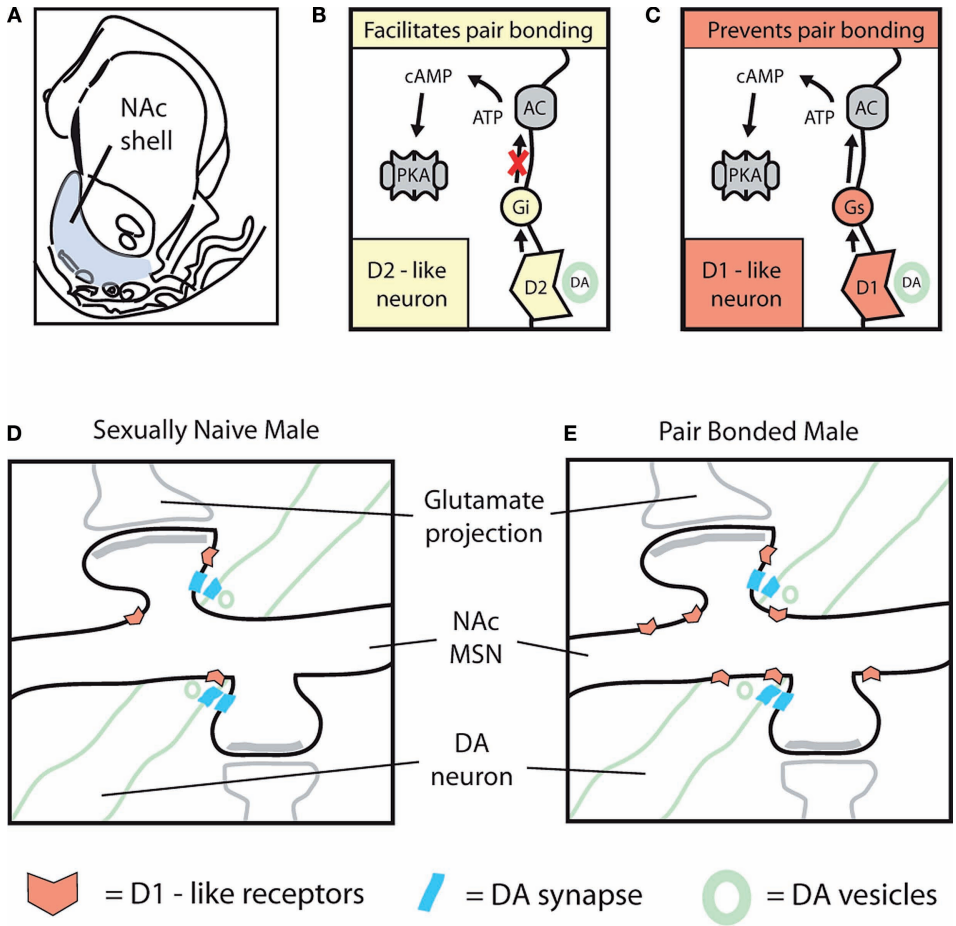


FIGURE 6 | Differential regulation of pair bond formation and maintenance by dopamine transmission within the NAc. (A) Cartoon based on (Arbuthnott and Wickens, 2007) showing the portion of the NAc shell where DA manipulations effect pair bond formation. **(B)** Diagram of D2-like signaling pathway involved in partner preference formation. **(C)** Diagram of D1-like signaling pathway that prevents partner preference formation. **(D)** Cartoon of medium spiny neuron (MSN) within the NAc receive glutamate projections to the heads of spines and dopaminergic projections to the neck of spines. This diagram represents D1-like receptor expression in sexually naive males. **(E)** A cartoon depicting the up-regulation of D1-like receptors in pair bonded male prairie voles.

DOPAMINE-OXYTOCIN INTERACTIONS AND PARTNER PREFERENCE FORMATION

Despite the critical role of DA in pair bonding, DA interacts with multiple neuropeptide systems in its regulation of this behavior (Lim et al., 2004b, 2007; Young and Wang, 2004). In particular, DA interactions with oxytocin receptors within the NAc are essential for pair bond formation (Liu and Wang, 2003). Activation of D2-like receptors within the NAc facilitates partner preference formation in the absence of mating, however, blockade of oxytocin receptors within this region (by co-infusion of an oxytocin receptor antagonist and a D2-like receptor agonist) prevents partner preferences induced by D2 activation (Liu and Wang, 2003). Further, facilitation of partner preference formation by activation of oxytocin receptors is not effective if D2-like receptors are blocked (Liu and Wang, 2003). Importantly, this study was conducted in female prairie voles (Liu and Wang, 2003), however, we have also shown that oxytocin receptors within the NAc are critical for partner preference formation in males (M. Smeltzer and Z. Wang, unpublished observations). While the mechanism of DA-oxytocin interactions is unknown, selective lesions of dopaminergic terminals in prairie voles did not reduce oxytocin receptor expression within the NAc (Lim et al., 2004a). This indicates that oxytocin receptors in this region are post-synaptic. Further, since oxytocin and D2-like receptors are both coupled to inhibitory G-protein signaling molecules (Burns et al., 2001), activation of both types of receptors may facilitate partner preference formation by inhibition of cAMP signaling pathways (Aragona and Wang, 2007). While existing data suggest that pair bond formation is mediated by co-activation of both oxytocin and D2-like DA receptors (Gingrich et al., 2000; Liu and Wang, 2003; Young et al., 2001), it is possible that they represent parallel systems that co-exist within the NAc. Future studies are needed to understand if DA and oxytocin receptor systems directly interact, and if so, determine if these interactions occur on the same or connected cells. Still, additional studies are required to understand DA interactions with the signaling systems critical for pair bonding but located outside of the NAc (such as vasopressin within the ventral pallidum; Lim et al., 2004b).

COMPARISON BETWEEN NEURAL REGULATION OF SOCIAL REWARD IN PRAIRIE VOLES AND HUMANS

Interestingly, the neural regulation of mate choice in humans also involves DA signaling systems (Fisher et al., 2005). Specifically, presentation of a picture of one's partner increases activation of dopaminergic circuitry in a similar manner as that caused by monetary reward (Aron et al., 2005; Zald et al., 2004). Thus, mate choice in humans may involve primary motivational or rewarding processes (Fisher et al., 2005) that are consistent with those observed in prairie voles. As such, the neural basis of partner preferences in prairie voles represents an excellent model for these aspects of mate choice in humans. Moreover, these findings suggest that understanding the neurobiology of reward processing is critical for understanding the neurobiology of social choice and decision-making (Loewenstein et al., 2008; Sanfey, 2007; Zak, 2004). Indeed, it has been suggested that pro-social behaviors may be achieved by activation of reward circuitry that promote cooperative behavior, in part, by facilitating positive emotions (Harbaugh et al., 2007), including feelings of trust (Rilling et al., 2002).

Trust is an essential component of human social organization and recent studies have shown that one neuropeptide critical for NAc regulation of pair bonding in voles, oxytocin, is critical for trust behavior in humans (Zak et al., 2004). The involvement of oxytocin in trust behavior was examined using a trust game, in which one player acts as an 'investor' that must choose whether or not to give money to a second player. If the 'investor' gives money to the second player, the amount of money in the game is increased and the 'investor' hopes that (during the second player's turn) the second player will reciprocate, giving the investor back more money than originally invested (Kosfeld et al., 2005). This is a one trial game so there is nothing to stop the second player from simply keeping all of the money. Thus, there is significant cost for the first player to trust that the second player will reciprocate. Interestingly, intra-nasal administration of oxytocin increased the ability of the 'investor' to overcome the risk associated with trust and increased the amount of money that the 'investor' gives to the second player (Kosfeld et al., 2005). Therefore, oxytocin appears to play a critical role in pro-social behavior in both humans and prairie voles.

CONCLUSION

The current review emphasizes some striking similarities between the neurobiology underlying pro-social behaviors in humans and prairie voles. As such, the prairie vole model is likely to be a powerful tool to investigate the neural regulation of social choice in more invasive ways that are not possible when using human subjects. While the prairie vole field is still in its infancy, experiments using this species clearly demonstrate that mesolimbic DA transmission is essential for social choice. Given that this system mediates aspects of reward and emotional processing, its involvement in social decision-making among humans may explain why humans often display strong social preferences rather than always acting out of pure self-regard (Camerer and Fehr, 2006; Fehr and Camerer, 2007; Sanfey, 2007). As the field of social neuroeconomics advances, it continues to consider whether social decision-making is best conceptualized as rational decision-making that is complicated because it involves more than one agent and thus requires more sophisticated learning alorithms (Lee, 2008), or if it is more informative to regard social decision-making as largely guided by emotional social motivation and hedonic processing (Sanfey et al., 2003; Skuse and Gallagher, 2009). While social decision-making certainly involves both reasoning as well as emotional processing, data from the prairie vole model demonstrate how a complex social organization can be achieved by a relatively small number of rather simplistic choice behaviors that are significantly mediated by reward processing. This supports the view that selection favored organisms that dealt with complex decisions by acting according to the degree of pleasure or displeasure likely to be associated with their behavioral response (Cabanac et al., 2009). Thus, while brains appear to be capable of an impressive capacity for logic and reason, very complex phenomena, such as social decision-making and cognition, can be also be robustly explained by hedonic and emotional processing.

ACKNOWLEDGEMENTS

The authors wish to thank Shanna L. Harkey for photos and Jeremy H. Day, Joshua L. Jones, and Bobby W. Pastrami for reading the manuscript. This work was supported by National Institutes of Health grants MHR01-58616, DAR01-19627, and DAK02-23048 to ZXW.

REFERENCES

Aragona, B. J., Cleaveland, N. A., Stuber, G. D., Day, J. J., Carelli, R. M., and Wightman, R. M. (2008). Preferential enhancement of dopamine transmission within the nucleus accumbens shell by cocaine is attributable to a direct increase in phasic dopamine release events. *J. Neurosci.* 28, 8821–8831.

Aragona, B. J., Liu, Y., Curtis, J. T., Stephan, F. K., and Wang, Z. (2003). A critical role for nucleus accumbens dopamine in partner-preference formation in male prairie voles. *J. Neurosci.* 23, 3483–3490.

Aragona, B. J., Liu, Y., Yu, Y. J., Curtis, J. T., Detwiler, J. M., Insel, T. R., and Wang, Z. (2006). Nucleus accumbens dopamine differentially mediates the formation and maintenance of monogamous pair bonds. *Nat. Neurosci.* 9, 133–139.

Aragona, B. J., and Wang, Z. (2004). The prairie vole (*Microtus ochrogaster*): an animal model for behavioral neuroendocrine research on pair bonding. *ILAR J.* 45, 35–45.

Aragona, B. J., and Wang, Z. (2007). Opposing regulation of pair bond formation by cAMP signaling within the nucleus accumbens shell. *J. Neurosci.* 27, 13352–13356.

Arbuthnott, G. W., and Wickens, J. (2007). Space, time and dopamine. *Trends Neurosci.* 30, 62–69.

Aron, A., Fisher, H., Mashek, D. J., Strong, G., Li, H., and Brown, L. L. (2005). Reward, motivation, and emotion systems associated with early-stage intense romantic love. *J. Neurophysiol.* 94, 327–337.

Bales, K. L., Plotsky, P. M., Young, L. J., Lim, M. M., Grotte, N., Ferrer, E., and Carter, C. S. (2007). Neonatal oxytocin manipulations have long-lasting, sexually dimorphic effects on vasopressin receptors. *Neuroscience* 144, 38–45.

Bamshad, M., Novak, M. A., and de Vries, G. J. (1993). Sex and species differences in the vasopressin innervation of sexually naive and parental prairie voles, *Microtus ochrogaster* and meadow voles, *Microtus pennsylvanicus*. *J. Neuroendocrinol.* 5, 247–255.

Bamshad, M., Novak, M. A., and de Vries, G. J. (1994). Cohabitation alters vasopressin innervation and paternal behavior in prairie voles (*Microtus ochrogaster*). *Physiol. Behav.* 56, 751–758.

Becker, J. B., Rudick, C. N., and Jenkins, W. J. (2001). The role of dopamine in the nucleus accumbens and striatum during sexual behavior in the female rat. *J. Neurosci.* 21, 3236–3241.

Berridge, K. C. (2007). The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology (Berl.)* 191, 391–431.

Berridge, K. C., and Robinson, T. E. (2003). Parsing reward. *Trends Neurosci.* 26, 507–513.

Burns, P. D., Mendes, J. O., Jr., Yemm, R. S., Clay, C. M., Nelson, S. E., Hayes, S. H., and Silvia, W. J. (2001). Cellular

mechanisms by which oxytocin mediates ovine endometrial prostaglandin F2alpha synthesis: role of G(i) proteins and mitogen-activated protein kinases. *Biol. Reprod.* 65, 1150–1155.

Cabanac, M., Cabanac, A. J., and Parent, A. (2009). The emergence of consciousness in phylogeny. *Behav. Brain Res.* 198, 267–272.

Cacioppo, J. T., Berntson, G. G., Sheridan, J. F., and McClintock, M. K. (2000). Multilevel integrative analyses of human behavior: social neuroscience and the complementing nature of social and biological approaches. *Psychol. Bull.* 126, 829–843.

Camerer, C. F., and Fehr, E. (2006). When does "economic man" dominate social behavior? *Science* 311, 47–52.

Carter, C. S., DeVries, A. C., and Getz, L. L. (1995). Physiological substrates of mammalian monogamy: the prairie vole model. *Neurosci. Biobehav. Rev.* 19, 303–314.

Carter, C. S., and Getz, L. L. (1993). Monogamy and the prairie vole. *Sci. Am.* 268, 100–106.

Champagne, F. A., Chretien, P., Stevenson, C. W., Zhang, T. Y., Gratton, A., and Meaney, M. J. (2004). Variations in nucleus accumbens dopamine associated with individual differences in maternal behavior in the rat. *J. Neurosci.* 24, 4113–4123.

Cho, M. M., DeVries, A. C., Williams, J. R., and Carter, C. S. (1999). The effects of oxytocin and vasopressin on partner preferences in male and female prairie voles (*Microtus ochrogaster*). *Behav. Neurosci.* 113, 1071–1079.

Curtis, J. T., and Wang, Z. (2005). Ventral tegmental area involvement in pair bonding in male prairie voles. *Physiol. Behav.* 86, 338–346.

Cushing, B. S., Martin, J. O., Young, L. J., and Carter, C. S. (2001). The effects of peptides on partner preference formation are predicted by habitat in prairie voles. *Horm. Behav.* 39, 48–58.

Cushing, B. S., and Wynne-Edwards, K. E. (2006). Estrogen receptor-alpha distribution in male rodents is associated with social organization. *J. Comp. Neurol.* 494, 595–605.

Day, J. J., Roitman, M. F., Wightman, R. M., and Carelli, R. M. (2007). Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat. Neurosci.* 10, 1020–1028.

DeVries, A. C., DeVries, M. B., Taymans, S., and Carter, C. S. (1995). Modulation of pair bonding in female prairie voles (*Microtus ochrogaster*) by corticosterone. *Proc. Natl. Acad. Sci. U.S.A.* 92, 7744–7748.

DeVries, A. C., DeVries, M. B., Taymans, S. E., and Carter, C. S. (1996). The effects of stress on social preferences are sexually dimorphic in prairie voles. *Proc. Natl. Acad. Sci. U.S.A.* 93, 11980–11984.

Dewsbury, D. A. (1975). Diversity and adaptation in rodent copulatory behavior. *Science* 190, 947–954.

Dewsbury, D. A. (1987). The comparative psychology of monogamy. *Nebr. Symp. Motiv.* 35, 1–50.

Di Chiara, G., and Bassareo, V. (2007). Reward system and addiction: what dopamine does and doesn't do. *Curr. Opin. Pharmacol.* 7, 69–76.

Emlen, S. T., and Oring, L. W. (1977). Ecology, sexual selection, and the evolution of mating systems. *Science* 197, 215–223.

Everitt, B. J., and Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat. Neurosci.* 8, 1481–1489.

Fehr, E., and Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* 11, 419–427.

Fiorino, D. F., Coury, A., and Phillips, A. G. (1997). Dynamic changes in nucleus accumbens dopamine efflux during the Coolidge effect in male rats. *J. Neurosci.* 17, 4849–4855.

Fisher, H., Aron, A., and Brown, L. L. (2005). Romantic love: an fMRI study of a neural mechanism for mate choice. *J. Comp. Neurol.* 493, 58–62.

Fowler, C. D., Liu, Y., Ouimet, C., and Wang, Z. (2002). The effects of social environment on adult neurogenesis in the female prairie vole. *J. Neurobiol.* 51, 115–128.

Getz, L. L. (1978). Speculation on social structure and population cycles of microtine rodents. *Biologist* 60, 134–147.

Getz, L. L., and Carter, C. S. (1996). Prairie-vole partnerships. *Am. Sci.* 84, 56–62.

Getz, L. L., Carter, C. S., and Gavish, L. (1981). The mating system of the prairie vole, *Microtus ochrogaster*: field and laboratory evidence for pair-bonding. *Behav. Ecol. Sociobiol.* 8, 189–194.

Getz, L. L., and Hofmann, J. E. (1986). Social organization in free-living prairie voles, *Microtus ochrogaster*. *Behav. Ecol. Sociobiol.* 18, 275–282.

Gingrich, B., Liu, Y., Cascio, C., Wang, Z., and Insel, T. R. (2000). Dopamine D2 receptors in the nucleus accumbens are important for social attachment in female prairie voles (*Microtus ochrogaster*). *Behav. Neurosci.* 114, 173–183.

Gobrogge, K. L., Liu, Y., Jia, X., and Wang, Z. (2007). Anterior hypothalamic neural activation and neurochemical associations with aggression in pair-bonded male prairie voles. *J. Comp. Neurol.* 502, 1109–1122.

Gray, G. D., and Dewsbury, D. A. (1973). A quantitative description of copulatory behavior in prairie voles (*Microtus ochrogaster*). *Brain Behav. Evol.* 8, 426–452.

Hall, E. R. (1981). The mammals of North America. New York, John Wiley.

Hammock, E. A., and Young, L. J. (2005). Microsatellite instability generates diversity in brain and sociobehavioral traits. *Science* 308, 1630–1634.

Harbaugh, W. T., Mayr, U., and Burghart, D. R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316, 1622–1625.

Hoffmann, R. S., and Koeppl, J. W. (1985). Zoogeography. In Biology of New World Microtus, Vol. 8, R. H. Tamarin, ed. (Special Publication, American Society of Mammalogists), pp. 84–113.

Ikemoto, S. (2007). Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. *Brain Res. Rev.* 56, 27–78.

Insel, T. R., and Hulihan, T. J. (1995). A gender-specific mechanism for pair bonding: oxytocin and partner preference formation in monogamous voles. *Behav. Neurosci.* 109, 782–789.

Insel, T. R., and Shapiro, L. E. (1992). Oxytocin receptor distribution reflects social organization in monogamous and polygamous voles. *Proc. Natl. Acad. Sci. U.S.A.* 89, 5981–5985.

Insel, T. R., and Young, L. J. (2001). The neurobiology of attachment. *Nat. Rev. Neurosci.* 2, 129–136.

Jansson, A., Goldstein, M., Tinner, B., Zoli, M., Meador-Woodruff, J. H., Lew, J. Y., Levey, A. I., Watson, S., Agnati, L. F., and Fuxe, K. (1999). On the distribution patterns of D1, D2, tyrosine hydroxylase and dopamine transporter immunoreactivities in the ventral striatum of the rat. *Neuroscience* 89, 473–489.

Kelley, A. E. (2004). Memory and addiction: shared neural circuitry and molecular mechanisms. *Neuron* 44, 161–179.

Kleiman, D. G. (1977). Monogamy in mammals. *Q. Rev. Biol.* 52, 39–69.

Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., and Fehr, E. (2005). Oxytocin increases trust in humans. *Nature* 435, 673–676.

Krueger, F., McCabe, K., Moll, J., Kriegeskorte, N., Zahn, R., Strenziok, M., Heinecke, A., and Grafman, J. (2007). Neural correlates of trust. *Proc. Natl. Acad. Sci. U.S.A.* 104, 20084–20089.

Lee, D. (2008). Game theory and neural basis of social decision making. *Nat. Neurosci.* 11, 404–409.

Li, M., and Fleming, A. S. (2003). Differential involvement of nucleus accumbens shell and core subregions in maternal memory in postpartum female rats. *Behav. Neurosci.* 117, 426–445.

Lim, M. M., Liu, Y., Ryabinin, A. E., Bai, Y., Wang, Z., and Young, L. J. (2007). CRF receptors in the nucleus accumbens modulate partner preference in prairie voles. *Horm. Behav.* 51, 508–515.

Lim, M. M., Murphy, A. Z., and Young, L. J. (2004a). Ventral striatopallidal oxytocin and vasopressin V1a receptors in the monogamous prairie vole (*Microtus ochrogaster*). *J. Comp. Neurol.* 468, 555–570.

Lim, M. M., Wang, Z., Olazabal, D. E., Ren, X., Terwilliger, E. F., and Young, L. J. (2004b). Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. *Nature* 429, 754–757.

Lim, M. M., and Young, L. J. (2004). Vasopressin-dependent neural circuits underlying pair bond formation in the monogamous prairie vole. *Neuroscience* 125, 35–45.

Liu, Y., Curtis, J. T., and Wang, Z. (2001). Vasopressin in the lateral septum regulates pair bond formation in male prairie voles (*Microtus ochrogaster*). *Behav. Neurosci.* 115, 910–919.

Liu, Y., and Wang, Z. X. (2003). Nucleus accumbens oxytocin and dopamine interact to regulate pair bond formation in female prairie voles. *Neuroscience* 121, 537–544.

Loewenstein, G., Rick, S., and Cohen, J. D. (2008). Neuroeconomics. *Annu. Rev. Psychol.* 59, 647–672.

Lynch, W. J., and Taylor, J. R. (2005). Persistent changes in motivation to self-administer cocaine following modulation of cyclic AMP-dependent protein kinase A (PKA) activity in the nucleus accumbens. *Eur. J. Neurosci.* 22, 1214–1220.

McGuire, B., Getz, L. L., Hofmann, J. E., Pizzuto, T., and Frase, B. (1993). Natal dispersal and philopatry in prairie voles (*Microtus ochrogaster*) in relation to population density, season, and natal social environment. *Behav. Ecol. Sociobiol.* 32, 293–302.

McGuire, B., and Novak, M. (1984). A comparison of maternal behaviour in the meadow vole (*Microtus pennsylvanicus*), prairie vole (*M. ochrogaster*) and pine vole (*M. pinetorum*). *Anim. Behav.* 32, 1132–1141.

McGuire, B., Russell, K. D., Mahoney, T., and Novak, M. (1992). The effects of mate removal on pregnancy success in prairie voles (*Microtus ochrogaster*)

and meadow voles (*Microtus pennsylvanicus*). *Biol. Reprod.* 47, 37–42.

Missale, C., Nash, S. R., Robinson, S. W., Jaber, M., and Caron, M. G. (1998). Dopamine receptors: from structure to function. *Physiol. Rev.* 78, 189–225.

Neve, K. A., Seamans, J. K., and Tranthan-Davidson, H. (2004). Dopamine receptor signaling. *J. Recept. Signal. Transduct. Res.* 24, 165–205.

Numan, M., Numan, M. J., Pliakou, N., Stolzenberg, D. S., Mullins, O. J., Murphy, J. M., and Smith, C. D. (2005). The effects of D1 or D2 dopamine receptor antagonism in the medial preoptic area, ventral pallidum, or nucleus accumbens on the maternal retrieval response and other aspects of maternal behavior in rats. *Behav. Neurosci.* 119, 1588–1604.

Oliveras, D., and Novak, M. (1986). A comparison of paternal behaviour in the meadow vole *Microtus pennsylvanicus*, the pine vole *M. pinetorum* and the prairie vole *M. ochrogaster*. *Anim. Behav.* 34, 519–526.

Pecina, S., Smith, K. S., and Berridge, K. C. (2006). Hedonic hot spots in the brain. *Neuroscientist* 12, 500–511.

Pfaus, J. G., Damsma, G., Wenkstern, D., and Fibiger, H. C. (1995). Sexual activity increases dopamine transmission in the nucleus accumbens and striatum of female rats. *Brain Res.* 693, 21–30.

Pfeiffer, T., Rutte, C., Killingback, T., Taborsky, M., and Bonhoeffer, S. (2005). Evolution of cooperation by generalized reciprocity. *Proc. Biol. Sci.* 272, 1115–1120.

Phillips, P. E., Stuber, G. D., Heien, M. L., Wightman, R. M., and Carelli, R. M. (2003). Subsecond dopamine release promotes cocaine seeking. *Nature* 422, 614–618.

Phillips, P. E., Walton, M. E., and Jhou, T. C. (2007). Calculating utility: preclinical evidence for cost-benefit analysis by mesolimbic dopamine. *Psychopharmacology (Berl.)* 191, 483–495.

Richfield, E. K., Penney, J. B., and Young, A. B. (1989). Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. *Neuroscience* 30, 767–777.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., and Kilts, C. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.

Robinson, D. L., Heien, M. L., and Wightman, R. M. (2002). Frequency of dopamine concentration transients increases in dorsal and ventral striatum of male rats during introduction of conspecifics. *J. Neurosci.* 22, 10477–10486.

Roitman, M. F., Wheeler, R. A., and Carelli, R. M. (2005). Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* 45, 587–597.

Roitman, M. F., Wheeler, R. A., Wightman, R. M., and Carelli, R. M. (2008). Real-time chemical responses in the nucleus accumbens differentiate rewarding and aversive stimuli. *Nat. Neurosci.* 11, 1376–1377.

Rutte, C., and Taborsky, M. (2007). Generalized Reciprocity in Rats. *PLoS Biol.* 5, e196.

Salamone, J. D., and Correa, M. (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav. Brain Res.* 137, 3–25.

Sanfey, A. G. (2007). Social decision-making: insights from game theory and neuroscience. *Science* 318, 598–602.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science* 300, 1755–1758.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron* 36, 241–263.

Self, D. W., Genova, L. M., Hope, B. T., Barnhart, W. J., Spencer, J. J., and Nestler, E. J. (1998). Involvement of cAMP-dependent protein kinase in the nucleus accumbens in cocaine self-administration and relapse of cocaine-seeking behavior. *J. Neurosci.* 18, 1848–1859.

Skuse, D. H., and Gallagher, L. (2009). Dopaminergic-neuropeptide interactions in the social brain. *Trends Cogn. Sci.* 13, 27–35.

Tankersley, D., Stowe, C. J., and Huettel, S. A. (2007). Altruism is associated with an increased neural response to agency. *Nat. Neurosci.* 10, 150–151.

Thomas, J. A., and Birney, E. C. (1979). Parental care and mating system of the prairie vole, *Microtus ochrogaster*. *Behav. Ecol. Sociobiol.* 5, 171–186.

Thomas, S. A., and Wolff, J. O. (2004). Pair bonding and "the widow effect" in female prairie voles. *Behav. Processes* 67, 47–54.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.

Wang, Z., Hulihan, T. J., and Insel, T. R. (1997). Sexual and social experience is associated with different patterns of behavior and neural activation in male prairie voles. *Brain Res.* 767, 321–332.

Wang, Z., Yu, G., Cascio, C., Liu, Y., Gingrich, B., and Insel, T. R. (1999). Dopamine D2 receptor-mediated regulation of partner preferences in female prairie voles (*Microtus ochrogaster*): a mechanism for pair bonding? *Behav. Neurosci.* 113, 602–611.

Wang, Z. X., and Aragona, B. J. (2004). Neurochemical regulation of pair bonding in male prairie voles. *Physiol. Behav.* 83, 319–328.

Wang, Z. X., and Novak, M. (1992). Influence of the social environment on parental behavior and pup development of meadow voles (*Microtus pennsylvanicus*) and prairie voles (*M. ochrogaster*). *J. Comp. Psychol.* 106, 163–171.

Wheeler, R. A., Twining, R. C., Jones, J. L., Slater, J. M., Grigson, P. S., and Carelli, R. M. (2008). Behavioral and electrophysiological indices of negative affect predict cocaine self-administration. *Neuron* 57, 774–785.

Williams, J. R., Catania, K. C., and Carter, C. S. (1992). Development of partner preferences in female prairie voles (*Microtus ochrogaster*): the role of social and sexual experience. *Horm. Behav.* 26, 339–349.

Winslow, J. T., Hastings, N., Carter, C. S., Harbaugh, C. R., and Insel, T. R. (1993). A role for central vasopressin in pair bonding in monogamous prairie voles. *Nature* 365, 545–548.

Wise, R. A. (2004). Dopamine, learning and motivation. *Nat. Rev. Neurosci.* 5, 483–494.

Witt, D. M., Carter, C. S., and Insel, T. R. (1991). Oxytocin receptor binding in female prairie voles: endogenous and exogenous oestradiol stimulation. *J. Neuroendocrinol.* 3, 155–161.

Witt, D. M., Carter, C. S., and Walton, D. M. (1990). Central and peripheral effects of oxytocin administration in prairie voles (*Microtus ochrogaster*). *Pharmacol. Biochem. Behav.* 37, 63–69.

Young, L. J., Lim, M. M., Gingrich, B., and Insel, T. R. (2001). Cellular mechanisms of social attachment. *Horm. Behav.* 40, 133–138.

Young, L. J., and Wang, Z. (2004). The neurobiology of pair bonding. *Nat. Neurosci.* 7, 1048–1054.

Zak, P. J. (2004). Neuroeconomics. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 359, 1737–1748.

Zak, P. J., Kurzban, R., and Matzner, W. T. (2004). The neurobiology of trust. *Ann. N. Y. Acad. Sci.* 1032, 224–227.

Zald, D. H., Boileau, I., El-Dearedy, W., Gunn, R., McGlone, F., Dichter, G. S., and Dagher, A. (2004). Dopamine transmission in the human striatum during monetary reward tasks. *J. Neurosci.* 24, 4105–4112.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 23 April 2009; paper pending published: 15 May 2009; accepted: 23 July 2009; published online: 11 August 2009.
Citation: Aragona BJ and Wang Z (2009) Dopamine regulation of social choice in a monogamous rodent species. Front. Behav. Neurosci. ***3****:15. doi: 10.3389/neuro.08.015.2009*

frontiers in
BEHAVIORAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 20 October 2009
doi: 10.3389/neuro.08.036.2009

Unforgettable ultimatums? Expectation violations promote enhanced social memory following economic bargaining

Luke J. Chang and Alan G. Sanfey*

Department of Psychology, University of Arizona, Tucson, AZ, USA

Edited by:
Daeyeol Lee,
Yale University School of Medicine, USA

Reviewed by:
Scott A. Huettel, Duke University, USA
Timothy Vickery, Yale University, USA

****Correspondence:***
Alan G. Sanfey, Department of Psychology, University of Arizona, 1503 E. University Blvd, Tucson, AZ 85721, USA.
e-mail: asanfey@u.arizona.edu

Recent work in the field of neuroeconomics has examined how people make decisions in interactive settings. However, less is currently known about how these social decisions influence subsequent memory for these interactions. We investigated this question by using functional magnetic resonance imaging to scan participants as they viewed photographs of people they had either recently played an Ultimatum Game with in the role of Responder, or that they had never seen before. Based on previous work that has investigated "cheater detection", we were interested in whether participants demonstrated a relative enhanced memory for partners that made either fair or unfair proposals. We found no evidence, either behaviorally or neurally, supporting enhanced memory based on the amount of money offered by the Proposer. However, we did find that participants' initial expectations about the offers they would experience in the game influenced their memory. Participants demonstrated relatively enhanced subjective memory for partners that made proposals that were contradictory to their initial expectations. In addition, we observed two distinct brain systems that were associated with partners that either offered more or less than the participants' expectations. Viewing pictures of partners that exceeded initial expectations was associated with the bilateral anterior insula, anterior cingulate cortex/premotor area, striatum, and bilateral posterior hippocampi, while viewing partners that offered less than initial expectations was associated with bilateral temporal-parietal junction, right STS, bilateral posterior insula, and precuneus. These results suggest that memory for social interaction may not be guided by a specific cheater detection system, but rather a more general expectation violation system.

Keywords: neuroeconomics, memory, social, expectation, ultimatum game, cheater detection, neural, decision-making

INTRODUCTION

Despite its relative youth, neuroeconomics as a field has made significant progress in describing the neural mechanisms that underlie decision-making (Glimcher et al., 2009). One approach within this domain has focused on circumstances in which the participant must consider the desires and intentions of another agent in reaching his or her eventual decision (Sanfey, 2007). These interactive situations have examined decisions made in a social environment, such as whether to trust or not trust another player or how to negotiate the division of a sum of money with another. The simplicity of these tasks and their ease of quantification provide not only a useful framework for developing mathematical models of optimal behavior within a social interaction (von Neumann and Morgenstern, 1944; Rabin, 1993; Fehr and Schmidt, 1999; Battigalli and Dufwenberg, 2007), but also a controlled environment within which to understand how social interaction interacts with more general cognitive processes such as memory.

One commonly used task in this domain is the Ultimatum Game (Guth et al., 1982). In this simplified bargaining scenario, one player known as the Proposer is endowed with a sum of money and told that their task is to make a proposal to the other player, the Responder, as to how this money should be divided between the two. The Proposer can make any offer he or she wants, from keeping all of the money for themselves to giving all of it away, and any division in between. Once the offer is made, the Responder must decide to either accept or reject the proposal. If the offer is accepted, then the money is simply divided as suggested. However, if the offer is rejected, then neither player receives any money. Both players are fully aware of the rules of the game, and once the Responder makes a decision the game is over.

Many studies across a multitude of disciplines and utilizing a variety of methods have examined social decision-making using the Ultimatum Game, and the behavioral results are generally strikingly similar (Camerer, 2003). Contrary to classical predictions, which suggest that Responders should accept any non-zero offer and as a consequence Proposers should make the lowest offer possible, the modal offer to Responders is typically a little less than half of the total pot, and this amount is almost always accepted. Offers of around 30% of the pot are accepted only about half of the time, and acceptance rates diminish as offers get lower.

One suggested mechanism as to why responders turn down what is in effect 'free' money when rejecting low offers is that people severely dislike inequity (Fehr and Schmidt, 1999), and consequently feel anger in response to unfair offers (Pillutla and Murnighan, 1996; Xiao and Houser, 2005). There is compelling physiological evidence supporting this argument. Unfair offers

are associated with increased autonomic tone (van 't Wout et al., 2006) and increased activity in the anterior insula (Sanfey et al., 2003). In fact, greater insula activity in response to an unfair offer results in an increased likelihood of rejection of that offer (Sanfey et al., 2003). Other studies have found that when neural systems involved in emotion regulation are disrupted in various ways, from using tryptophan depletion (Crockett et al., 2008) to lesions of the ventromedial prefrontal cortex (Koenigs and Tranel, 2007), the result is increased rejection rates of unfair offers.

Though the response to fair and unfair offers provides an interesting window into how the competing motivations of maintaining one's reputation and maximizing one's financial gain interact in decision-making, other relevant questions can be answered using these type of tasks. Of perhaps equal importance to examining the processes that underlie performance in this task is to ask what happens, both behaviorally and neurally, when we re-encounter a player who has made a fair or unfair offer to us in the past. How do our perceptions of others shift when these people have previously treated us either well or poorly? In this initial attempt to investigate this question, we focus on memory for players with whom we have recently interacted, and specifically examine whether the way in which another player has treated us has an impact on how we in turn remember them.

Several theoretical proposals have been made as to whether we are more attuned to remembering those who have treated us either fairly or unfairly in the past. In their highly influential theory of social exchange, Cosmides and Tooby (1992) argue that humans have evolved specific cognitive abilities to promote reciprocal altruism, a construct that has been associated with positive evolutionary fitness (Trivers, 1971). Of particular importance to their theory is the ability to detect, remember, and punish "cheaters" – individuals who benefit themselves by violating a social contract (Cosmides and Tooby, 1992). However, despite the intuitive appeal of this theory, the primary evidence presented in favor of the selective detection of cheaters is that experimental participants demonstrate improved conditional reasoning when asked to detect violations of a social contract, when compared to non-social contract violations (Cosmides, 1989; Gigerenzer and Hug, 1992). Evidence supporting these theoretical claims in the domain of memory is more mixed.

Several studies have directly examined explicit memory for cheaters. There is some evidence that after 1 week participants had better memories for pictures of people with behaviors associated with cheating (e.g. "E.A. is a bishop who was caught embezzling money from his own church.") as compared to pictures of those that were associated with trustworthy behaviors (e.g. "J.H. is a vendor at baseball games who, after finding a wallet containing $250, located the owner using the driver's license.") (Mealy et al., 1996; Chiappe et al., 2004). However, more recent studies that have attempted to address some of the methodological limitations of these experiments have failed to replicate this finding, with no differences between cheaters and trustworthy pictures emerging (Barclay and Lalumiere, 2006; Mehl and Buchner, 2008). There is even some preliminary evidence for increased confidence, though not accuracy, in remembering altruists (Barclay and Lalumiere, 2006), and also that people may have better source memory than recognition memory for cheaters, meaning that people were better at remembering that an individual was a cheater than actually correctly identifying that they had seen the person before (Buchner et al., 2009).

One explanation of these mixed findings is that the memory manipulations used were not particularly socially relevant for the participants. As outlined above, these paradigms typically involve participants reading a vignette describing either a cheating or trustworthy act by a pictured person, and then subsequently performing a recognition memory test on the set of photographs. There are surprisingly few studies that have attempted to have participants first actually engage in meaningful social interactions with other people, and then test their memory for these partners. In one study with a variant of this methodology, participants were asked to imagine playing a constant strategy (i.e. cooperate or defect) in a Prisoner's Dilemma game, and were then shown pictures and the strategies of their partners (Oda, 1997). After being tested 1 week later, the experimenters found that participants remembered defectors better than cooperators and that this effect interacted with gender. However, there was no clear explanation of the interaction with gender, nor was it clear that participants were actually engaged in the game as they were forced to stick with the same strategy.

Within neuroeconomics, there is clear evidence that people use information about a partner's history to inform decisions in future social interactions, such as to avoid trusting a cheater in a subsequent interaction or to punish them if given the opportunity. People are more likely to invest trust in partners perceived to be initially trustworthy as opposed to untrustworthy (Delgado et al., 2005; van 't Wout and Sanfey, 2008), and also seem able to then disregard this prior information when these partners actually abuse their trust. There is also evidence supporting the notion that people are willing to punish cheaters, even at the risk of incurring a financial cost to themselves in Ultimatum (Guth et al., 1982) and Public Goods Games (Fehr and Gachter, 2002; de Quervain et al., 2004). While these findings suggest that people can learn both who to trust and who not to trust and will punish cheaters given the opportunity, there is as yet no conclusive evidence directly supporting better explicit memory for either cheaters or cooperators.

One study (Singer et al., 2004) attempted to investigate this question both behaviorally and neurally by scanning participants using functional magnetic resonance imaging (fMRI) as they viewed faces which had previously behaved in either cooperative or non-cooperative ways in a modified repeated Prisoner's Dilemma game. Behaviorally, the authors report that cooperators were rated as more likeable and defectors as less likeable than control faces. In addition, participants were more accurate in recalling the behavior of both cooperators and defectors as compared to the null games. However, because there was neither money at stake for the null games nor an equal distribution of trials for each condition, the results of this forced choice memory task should be interpreted cautiously. In terms of neural findings, the authors reported that when asked to make a gender assessment of pictures of cooperators as compared to those who played null games, participants had increased activity in the left ventral putamen and left amygdala. In contrast, when participants viewed pictures of defectors compared to null trials, they showed increased activity in the vmPFC. These preliminary findings suggest that viewing faces of defectors and

cooperators from a socially relevant task may be associated with distinct neural systems. However, it remains an open question as to whether or not there may be selectively better explicit memory for cheaters and what processes might underlie this.

A possible mechanism that could explain the aforementioned pattern of results is the notion of deviation from expectation, that is, when partners play in a way differently than we predict. While it is known that alterations of expectation can affect decision-making in the Ultimatum Game (Sanfey, 2009), up to now there has been relatively little investigation of how expectations, and specifically deviation from expectations, can alter patterns of memory in social decision-making.

Some limited evidence comes from a recent study using the Trust Game, which found that participants did not have selective memory for either cooperators or defectors *per se*, but rather demonstrated enhanced memory for both types of opponents in certain circumstances, these circumstances being that the better–remembered opponent played a relatively infrequent strategy. That is, at different times they remembered both cheaters and defectors better, but only when they comprised merely 20% of the total number of interactions (Barclay, 2008). It is important to note that participants in this experiment knew *a priori* that they were playing with computer partners, so it is not clear if these results could be generalized to games played with real opponents. Nonetheless, this study provides compelling evidence that people may have enhanced memory for partners that behave contrary to social conventions, regardless of their behavior. This suggests therefore that people may not rely on a specific cheater detection system, but rather a more general expectation violation system – a notion within the field of memory that has been known for some time (von Restorff, 1933; Ranganath and Rainer, 2003), often discussed in this literature as a "novelty detection" mechanism. It is therefore possible that a more general novelty detection system can potentially be employed as a cheater detection system. Because interactions with cheaters in the real world are likely to be relatively infrequent, the expectation of cheating behavior should be low and as a result incidences of cheating should be particularly memorable. However, importantly, if we do expect substantial cheating behavior in our environment, this account would predict that partners who treat us well should be preferentially encoded and remembered.

We sought to investigate this question by using fMRI to scan the brains of participants immediately after they played a series of Ultimatum Games with a variety of partners. Firstly, we examined if our participants demonstrated more accurate memories for partners that had treated them either fairly (an equal offer) or unfairly (an unequal offer in the partner's favor). Secondly, we were particularly interested in the neural response to viewing a photograph of a previous partner as compared to a photograph of a previously unseen person, and whether the offer that had been made to the participant mediated this neural activity in our participants. Contrary to most prior behavioral studies of memory for cheaters, players in this study engaged in an actual social decision interaction and we directly assessed their social memory while they were being scanned using a standard recognition task. This study therefore can potentially inform an ongoing debate about whether people actually have enhanced memory for cheaters, and if so, whether the brain is equipped with a system to complete this task.

MATERIALS AND METHODS

PARTICIPANTS

Eighteen participants (mean age = 19.9, female = 56%) were recruited via advertisements posted on the campus of the University of Arizona to participate in this study. All participants were screened for any significant health-related or neuropsychiatric disorders and none were currently taking psychoactive medication. Two participants were excluded from the analysis for technical reasons (corrupted data). All participants gave informed consent according to procedures approved by the University of Arizona's Institutional Review Board.

PROCEDURE

Expectations

Prior to being scanned, we elicited participants' beliefs about the kinds of offers they expected to encounter, with participants being asked the number of people out of 100 that they believed would make a \$0, \$1, \$2, \$3, \$4, \$5, \$6, or \$7 offer. Participants' elicited expectations prior to playing the game were used to create a distribution of the frequency of offers that they expected to encounter. The mode of this distribution was used to represent each participant's initial expectation.

Ultimatum game

Participants then played a standard single-shot Ultimatum Game in the role of Responder with 48 different partners while undergoing fMRI. Twenty-four of these partners were human, 12 were computers, and 12 were non-intentional humans (i.e. humans whose responses were randomly generated). Each offer was preceded by a picture of their partner for that round. Though participants were told that the human-intentional offers would be made by other players, in actual fact all offers were controlled by the experimenter, and all participants saw the same set of offers. This set consisted of equal numbers (12 each) of \$1, \$2, \$3 and \$5 offers, all of which were made from a \$10 pot. For each participant, all pictures were randomly paired to an offer amount, ensuring that there was no potential picture by offer amount interaction. Participants were paid \$20 for participating and an additional \$5, which they believed was based on their performance in the game. Further details of the Ultimatum Game portion of the experiment will be described in greater detail in a separate paper. While participants were not directly queried after the experiment about whether or not they believed that they were interacting with real partners, no participant expressed doubt towards the experimenter at any time during the experiment.

Memory experiment

After completing the Ultimatum Game trials, participants were given an incidental memory test while undergoing fMRI (see **Figure 1** for a trial timeline). Participants had not been forewarned about this task. Participants viewed randomly presented pictures of the 24 human-intentional partners they had previously encountered, as well as 24 new faces they had not seen in the Ultimatum Game task. Participants did not view pictures of the computers or the non-intentional human partners. There were an equal number of male and female faces for both the new and old sets of faces. To begin each trial, a jittered fixation was seen for 6 s on average

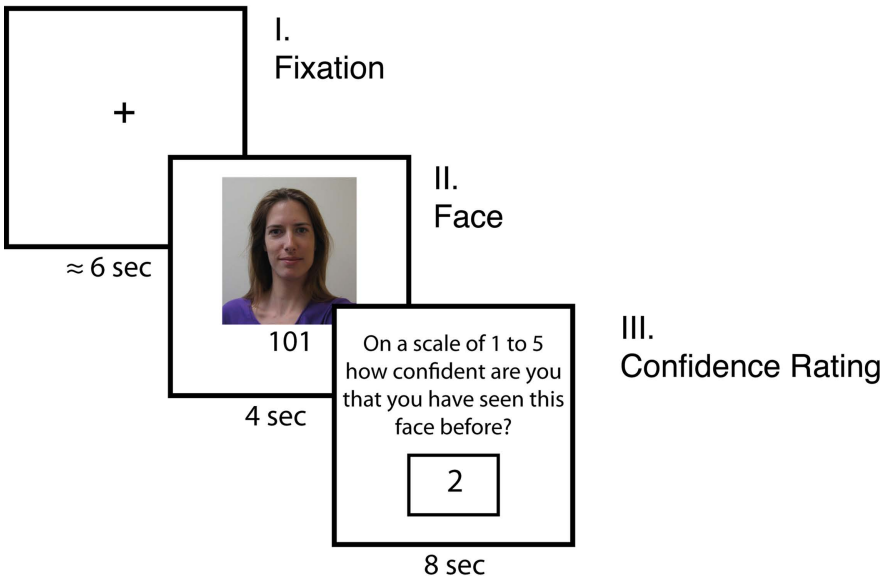


FIGURE 1 | Timeline of memory experiment. Participants first viewed a jittered fixation that was on average 6 s long before seeing each face. Participants saw 24 photographs of previous partners in the Ultimatum Game and 24 new people for 4 s. Participants were then asked to rate their level of confidence that the face was either new or old on a 5 point rating scale.

and then a face was shown for 4 s. After being presented with this photograph, participants were asked to rate their confidence that they had played the Ultimatum Game with this person on a scale of 1 to 5 (1 = "I've definitely never seen this person before; 2 = "I don't think I've seen them before"; 3 = "I'm not sure if I've seen this person before"; 4 = "I think I may have seen them before"; 5 = "I've definitely seen this person before"). Participants were given 8 s to make this judgment, and ratings were entered by scrolling through the possible options with one response button and then selecting the desired rating with the other button. While the rating system was always the same (e.g. 1 = never seen; 5 = definitely seen), the ratings randomly scrolled up or down on each trial to eliminate any possible motor confounds. Therefore, the number of button presses were orthogonal to the actual ratings provided. Stimuli were presented via EPrime software (Psychology Software Tools, Inc, Pttsburgh, PA, USA) using MRI-compatible googles and responses were recorded using a fiber optic button box (Resonance Technologies, Van Nuys, CA, USA).

ANALYSES

All behavioral statistics were computed using the R statistical package (R_Development_Core_Team, 2008). For regressions that included repeated observations, we used the lme4 mixed effects general linear model package (Bates et al., 2008). Participants were treated as a random effect with varying intercepts and slopes. We report the parameter estimates (b), standard error, *t*-values, and *p*-values. Because there is no generally agreed upon method for calculating *p*-values in mixed models, we used two separate methods. First, we calculated the degrees of freedom by subtracting the number of observations minus the number of fixed effects (Kliegl et al., 2007). Second, we generated confidence intervals from the posterior distribution of the parameter estimates using Markov Chain Monte Carlo methods (Baayen et al., 2008). These results were identical unless otherwise noted. For robust regressions we used the rlm function from the MASS package using an MM-estimator (Venables and Ripley, 2002).

D′

To measure participant's ability to discriminate old from new faces, we used D′, a signal detection metric (Wickens, 2002). D′ controls for individual participants' response bias (i.e. their propensity to say yes) and was calculated as the difference between the standardized *z*-score for hits (indicated by a 4 or 5 on the confidence rating for an old face) and the standardized *z*-score for false positives (indicated by a 4 or 5 on the confidence rating for a new face). Because this analysis emphasizes hits and false positives, it ignores differences in levels of subjective confidence, that is, the difference between a 1 and 2, or a 4 or 5 rating. D′ scores were calculated separately for every level of offer amount.

Data acquisition

Each scanning session included a T1-weighted MPRAGE structural scan (TR = 11 ms, TE = 4 ms, matrix = 256 × 256, slice thickness = 1 mm, gap = 0 mm), followed by five functional runs. The first 3 functional runs contained the Ultimatum Game trials and the last two contained the memory trials (240 volumes per run). Functional scans used a 3-shot multiple echo planar imaging (MEPI) GRAPPA sequence using parameters selected to maximize signal in regions associated with high susceptibility artifact, such as orbitofrontal cortex and medial temporal lobe (Stocker et al., 2006; Weiskopf et al., 2006) (TR = 2000 ms, TE = 256 ms, matrix = 96 × 96, FOV = 192 mm, slice thickness = 3.0 mm, 42 axial slices, voxel size 2 × 2 × 3). The MEPI sequence employs parallel imaging and allows for increases in

signal intensity, image resolution, the number of slices that can be acquired in a 2000-ms TR, as well as substantial decreases in geometric distortion (Newbould et al., 2007).

Data preprocessing

Functional imaging data were preprocessed and analyzed using the FSL Software package 4.1.4 (FMRIB, Oxford, UK). The first three volumes of the functional runs were discarded to account for T1 equilibrium effects. Images were corrected for slice scan time using an ascending interleaved procedure. Head motion was corrected using MCFLIRT using a 6-parameter rigid-body transformation. Images were spatially smoothed using a 5 mm full width at half maximum Gaussian kernel. A high pass filter was used to cut off temporal periods longer than 66 s. All images were initially co-registered to the participant's high resolution structural scan and were then co-registered to the MNI 152 person 2-mm template using a 12-parameter affine transformation. All functional analyses are overlaid on the participants' average high resolution structural scan in MNI space.

General imaging analysis methods

A 3-level mixed effects general linear model (GLM) was used to analyze the imaging data. A first-level GLM was defined for each participant's functional run that included a boxcar regressor for each epoch of interest (e.g. face phase), convolved with a canonical double-gamma hemodynamic response function. To account for residual variance we also included the temporal derivatives of each regressor of interest, the six estimated head movement parameters, and any missed trials as covariates of no interest. The resulting GLM was corrected for temporal autocorrelations using FILM prewhitening. A second-level fixed effects model was fit for each participant to account for intra-run variability. For each participant, contrasts were calculated between predictors for different regressors of interest at every voxel in the brain. A one-sample *t*-test was used at the third-level for each contrast using a Bayesian implementation of mixed effects inference (Behrens et al., 2008). We corrected for multiple comparisons with cluster correction utilizing Gaussian random field theory with an initial cutoff of $Z > 2.3$ and a FWE $p < 0.05$.

We report the results of three analyses. The offer amount analysis included individual regressors during the face phase for players who had previously made offers of \$1, \$2, \$3, or \$5, a regressor indicating the duration of the response time during the memory phase, a regressor indicating a distractor face, a regressor for missed trials, the temporal derivatives of each of these predictors and 6 motion parameters (20 predictors total). We report the results for the Unfair (i.e. \$1 and \$2 offers) vs Fair (i.e. \$5) contrast. For the expectation violation analysis we included regressors at the face phase for offers below expectation (i.e. standardized expectation error (SEE) > 0), offers above expectation (i.e. SEE < 0), and offers at expectation (i.e. SEE = 0). SEE is the within-subject *z*-score of the numerical deviation of an offer amount from a participant's initial expectations. In addition, we included a regressor modeling the memory phase for the duration of the response, a regressor indicating a distractor face, a regressor modeling missed trials, and their temporal derivatives and 6 motion parameters (18 predictors total). We report a linear contrast of prediction error (i.e. +1 0 −1) for Positive, Zero, and Negative SEE regressors. Finally, the third analysis was identical to the second analysis except the third level linear contrast was weighted by each participant's standardized initial expectation, effectively utilizing a correlation analysis rather than a one sample *t*-test. This analysis tests the interaction between participant's initial expectation and their SEE. All trials in which the participants indicated that they were unsure (i.e. a rating of 3) were excluded from all analyses (78 trials total for all subjects, or 10.2% of observations).

RESULTS

BEHAVIORAL RESULTS

Ultimatum game

Consistent with previous research (Sanfey et al., 2003; van 't Wout et al., 2006; Harle and Sanfey, 2007), acceptance rates decreased as offers got lower, and participants were significantly more likely to accept fair (\$5) as opposed to unfair (\$1; \$2) offers, illustrated using a mixed effects logit model (Jaeger, 2008), b = 4.24, se = 0.84, odds ratio = 69.21, Wald $Z = 5.07$, $p < 0.05$. The average acceptance rate for all intentional offers was 62.2%, with three participants accepting all offers. Consistent with previous research, participants expected most participants to make fair offers (mean = 4.5, sd = 0.63) (Sanfey, 2009).

Memory

A one-sample *t*-test revealed that participants were on average accurate in their ability to discriminate between old and new faces (mean $D' = 2.04$), $t(15) = 14.68$, $p < 0.001$. Participants were able to correctly identify both old faces (mean correct = 70%, se = 0.04) and new faces (mean correct = 76%, se = 0.04). We used a mixed effects linear model to test whether or not the amount of money offered by the proposer would influence participant's memory for that person, but did not observe a significant effect, $b = 0.03$, se = 0.05, $t = 0.52$, ns. These results indicate that participants were sensitive in their ability to discriminate between new and old faces, but that, on average, this discriminability was not influenced by the amount of money offered by the partner.

However, closer examination of these results indicate that participants demonstrated considerable variability in their ability to remember proposers that made either fair or unfair offers (**Figure 2**). Some participants appeared to demonstrate improved memory for proposers that made unfair offers, while other participants remembered proposers that made fair offers indicated by the random effect slope coefficient). Using robust regression we found that participants' initial expectations predicted the random effects parameter estimates from the previous offer amount analysis, parameter estimate = 0.16, se = 0.04, $t = 4.32$, $p < 0.05$. This analysis indicates that as initial expectations increased, the slope of offer amount on D' decreased. In other words, participants with low initial expectations had positive memory slopes, meaning that they demonstrated augmented memory for proposers that made fair offers, while participants with high initial expectations had negative memory slopes, indicating increased memory for proposers that made unfair offers (**Figure 2A**).

To test this expectation violation hypothesis more explicitly, we used a mixed effects linear model treating subjects as a random intercept. Specifically, we attempted to predict participant's

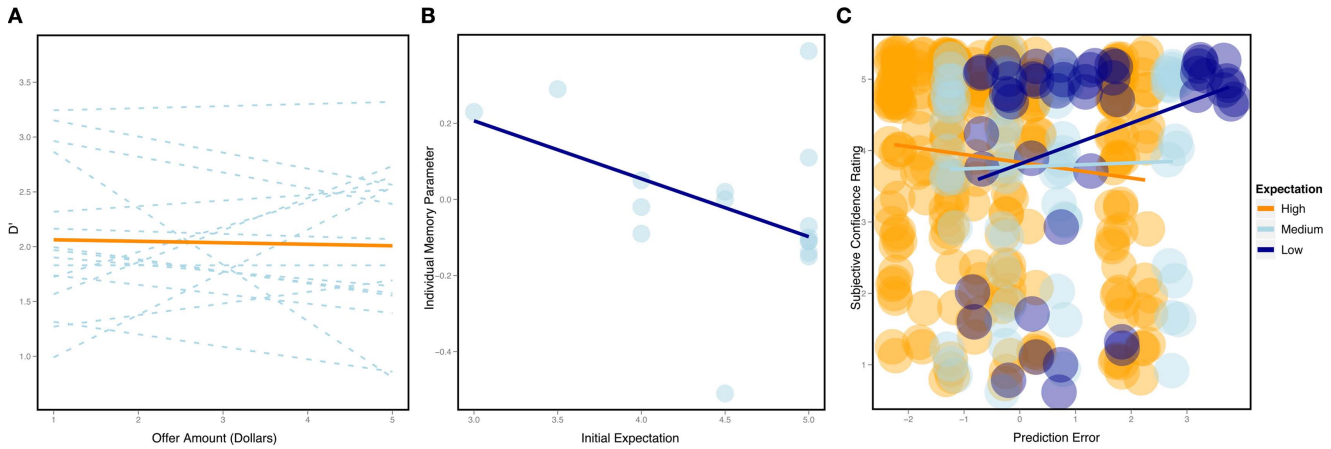


FIGURE 2 | Behavioral results. (A) The effect of offer amount on D′ plotted in orange. The individual regression lines plotted in blue reveal considerable variability in discriminability by offer amount. Positive slopes indicate enhanced memory for partners that made fair offers, while negative slopes indicate enhanced memory for partners that made unfair offers. **(B)** The effect of initial expectation on individual memory parameter [i.e. the slopes from **(A)**]. The individual variability in memory for offer amounts is related to initial expectations. High initial expectations are associated with negative slopes (i.e. enhanced memory for partners that made unfair offers), while low initial expectations are associated with positive slopes (i.e. enhanced memory for partners that made fair offers). **(C)** The effect of expectation error on subjective confidence ratings, split by initial expectations (groupings are only for plotting, not analysis). High initial expectations (orange) demonstrated enhanced memory for offers that were less than their initial expectation. Low initial expectations (dark blue) were associated with enhanced memory for partners that made offers that exceeded their expectations. No relationship was found for medium initial expectations (light blue).

subjective confidence ratings using their centered initial expectation, the centered deviation of the offer amount from their initial expectation, and the interaction between these two variables. We observed an initial expectation by expectation deviation interaction, $b = 0.17$, se $= 0.08$, $t = 2.21$, $p < 0.05$, with no significant main effects. Participants demonstrated enhanced confidence ratings for faces that violated their initial expectations (see **Figure 2**).

IMAGING RESULTS

Offer amount

As noted above, we observed no significant effect of offer amount in predicting participant's ability to discriminate between old and new faces. Similarly, for the corresponding imaging analysis, we did not observe any significant voxels above threshold for this previously fair vs. previously unfair contrast, even at a more liberal $p < 0.001$ (uncorrected) threshold. Therefore, at least on average across participants, there is no particular neural signature for either previously fair or previously unfair partners.

Expectation violation

Our more detailed behavioral analysis indicated that expectation violation, and not offer amount, was associated with enhanced subsequent memory for partners. To explore the neural systems underlying this effect we ran two separate imaging analyses. The first analysis examined the effect of expectation deviation. This contrast was associated with bilateral anterior insula, pre-supplementary motor area (pre-SMA)/anterior cingulate cortex (ACC), the striatum (including the caudate and nucleus accumbens), and bilateral posterior hippocampi/parahippocampi. Negative expectation deviations were associated with bilateral temporal parietal junction (TPJ), right superior temporal sulcus (STS), posterior insula, and precuneus (see **Figure 3**). The second analysis examined the interaction between the initial expectation and the expectation error, by weighting the first analysis by each participant's standardized initial expectation at the third level. No voxels survived our threshold for this analysis. Thus, this set of analyses reveals a network previously associated with expectation violation (i.e. insula, pre-SMA, and NAcc), and memory retrieval (i.e. hippocampi/parahippocampi) when participants view faces of partners who offered more than the participants initially expected, and a network associated with theory of mind processing (i.e. STS/TPJ) and memory (i.e. precuneus) when viewing partners that offered less than the participant initially expected (see **Table 1** for a complete list of regions).

DISCUSSION

This study investigated how economic exchange impacts subsequent memories for social partners. Following a standard Ultimatum Game paradigm, participants were shown photographs of both previously seen and unseen people, and asked to rate their confidence that they had viewed these pictures before. This question is important in understanding the behavioral and neural effects of reappraising a partner with whom one has previously been engaged in social economic interaction. In addition, this research was also interested in investigating the notion of "cheater detection", that is, the idea of relatively enhanced memories for social partners who have treated us badly in the past (Cosmides and Tooby, 1992; Mealy et al., 1996; Singer et al., 2004; Barclay, 2008).

We were primarily interested in whether participants exhibited a relative memory enhancement for partners that made either fair or unfair proposals. A demonstration of the latter (i.e. enhanced memory for unfair proposers) would provide evidence supporting the existence of behavioral cheater-detection effects. However, we

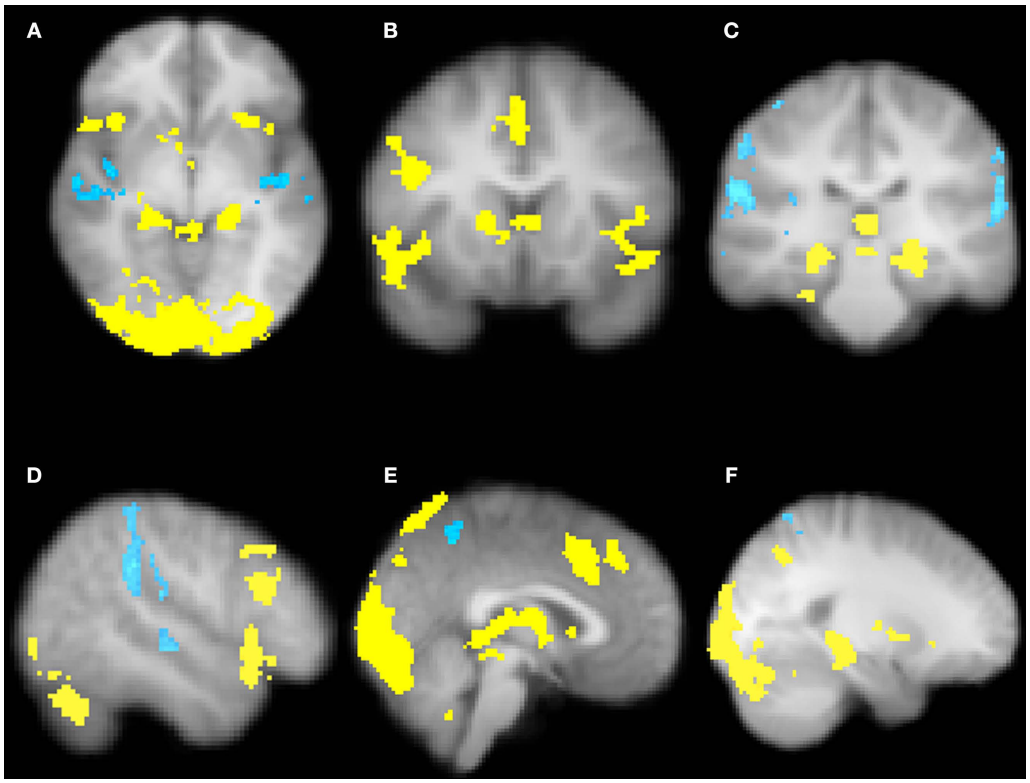


FIGURE 3 | Brain regions associated with expectation violation. These results are a linear contrast of standardized expectation error (SEE) when participants are viewing pictures of their partners during the memory task. Yellow values are associated with partners that offered more than the participant's initial expectation. Blue values are associated with viewing partners that offered less than the participant's initial expectation. **(A)** Axial section shows bilateral insula, ventral striatum, bilateral posterior hippocampi, visual cortex associated with positive SEE, while right superior temporal sulcus and bilateral posterior insula are associated with negative SEE. **(B)** Coronal section shows bilateral anterior insula, ventral striatum, right dorsolateral prefrontal cortex, and pre-supplementary motor area/dorsal anterior cingulate. **(C)** Coronal section shows bilateral posterior hippocampus. **(D)** Right lateral section shows DLPFC, VLPFC, and TPJ and STS. **(E)** Sagittal section shows preSMA/DACC. **(F)** Sagittal section shows posterior hippocampus/parahippocampus. All clusters survive whole brain correction using cluster correction, $Z > 2.3$, $p < 0.05$ and are displayed on the group average T1 image using the radiological convention (left = right).

did not observe a significant effect, either behaviorally or neurally, for an influence of offer fairness on memory. Instead, we found that participants demonstrated considerable variability in their ability to discriminate between partners associated with different levels of fairness (e.g. some participants demonstrated relative enhanced memory for fair partners and some for unfair partners). Importantly, this variability was predicted by their initial expectations about the range of offers they would see in the game. Those who had low initial expectations (operationalized by their modal reported expected offer) were more likely to demonstrate augmented memory for partners that exceeded their expectations, while those who had high initial expectations demonstrated better memory for partners who made offers lower than their initial expectations. We also observed a significant interaction between participants' initial expectations and their expectation error (i.e. the proposal's deviation from initial expectation) in predicting subjective confidence ratings. Participants who had low initial expectations were more likely to remember partners that made offers that were greater than their initial expectations, while participants that had high initial expectations were more likely to remember partners that made offers that were lower than their initial expectations.

This finding is consistent with the results of a recent cheater detection study (Barclay, 2008), in which participants demonstrated enhanced memory in a recognition test for whichever behavior (e.g. cooperate or defect) was more *infrequent* in a previously-played Trust Game. When the majority of partners cooperated, participants remembered defectors better, whereas when most partners defected, participants remembered cooperators better. Our study employed a different approach than that of (Barclay, 2008), namely use of an Ultimatum as opposed to the Trust Game, and additionally we used the participants own expectations as the "baseline", as opposed to examining violations from experienced probabilities, but the two sets of results converge on the same interpretation – that deviations from prior expectations result in greater salience and thus better memory encoding.

This conjecture has been posited for a long time in the memory literature dating to von Restorff (1933). The Von Restorff effect refers to memory enhancement occurring when an item is isolated either by manipulating the context (e.g. item is printed in red in a list of items printed in black) or content (e.g. inserting a nonsense syllable into a list of meaningful words) (Wallace, 1965). This effect is thought to be associated with unexpected change rather than

Table 1 | Brain regions associated with expectation error.

Hemisphere	Region	BA	*Z* value	*X*	*Y*	*Z*
POS > NEG						
L	Angular gyrus	39	3.56	–46	–60	36
L	Anterior insula	48	3.15	–42	12	–4
L	Frontal operculum	48	3.04	–42	18	4
L	Lateral OFC	38	3.23	–38	18	–14
L	Midbrain (substantia nigra)	NA	3.76	–10	–14	–10
L	Occipital cortex	18	4.88	–34	–94	10
L	Occipital cortex (primary visual)	17	4.99	–6	–96	14
L	Posterior hippocampus	27	4.14	–20	–32	–4
L	Superior parietal lobule	7	3.6	–30	–58	42
L	Temporal pole	38	3.36	–50	14	–8
R	ACC	24	3.73	2	20	36
R	Anterior insula	47	3.44	42	16	–8
R	Fusiform gyrus	37	5.03	38	–46	–24
R	Inferior frontal gyrus	48	3.55	54	18	20
R	Middle frontal gyrus	44	3.04	52	14	36
R	Occipital cortex	18	5.11	8	–96	20
R	Parahippcampus	27	4.19	18	–34	–8
R	Posterior hippocampus	20	4.04	24	–26	–10
R	Pre-SMA	32	3.84	4	20	44
R	Cerebellum (right VI)	19	5.08	28	–68	–20
R	Superior frontal gyrus	8	3.31	0	36	50
R	Temporal pole	38	3.43	50	20	–20
NEG > POS						
L	Angular gyrus	39	3.56	–46	–60	36
L	Posterior insula	48	3.01	–38	–12	–2
L	Precuneus	7	3.22	–6	–60	52
L	Superior parietal lobule	5	3.17	–18	–60	66
L	Superior temporal gyrus	22	3.5	–62	–30	12
L	STS	48	3.73	–48	–12	–8
L	Supramarginal gyrus anterior division (TPJ)	40	3.39	–62	–30	40
L	TPJ (parietal operculum cortex)	48	3.41	–58	–38	26
R	Posterior insula	20	3.79	40	–12	–10
R	Precuneus	5	3.08	4	–48	60
R	Superior temporal gyrus	42	3.47	58	–34	16
R	STS	22	3.5	62	–14	–6
R	Supramarginal gyrus posterior division (TPJ)	48	3.77	54	–36	28
R	TPJ (parietal operculum cortex)	48	3.79	62	–28	22

This table reflects the contrast positive expectation error compared to negative expectation error and shows the local maxima of clusters surviving cluster correction Z > 2.3, p < 0.05 in MNI space. Cortical and subcortical regions were identified using the Harvard-Oxford Probabilistic Anatomical Atlas and Mai et al. (2007), while the cerebellar regions were identified using a probabilistic cerebellar atlas (Diedrichsen et al., 2009). Abbreviations: TPJ = temporal-parietal junction, SMA = supplementary motor area, STS = superior temporal sulcus, OFC = orbitofrontal cortex, ACC = anterior cingulate cortex.

actual isolation (Green, 1956). The source of this mechanism has been the focus of considerable research in the memory, attention, and cognitive control literatures and has even served as one of the primary paradigms in studying cognition in preverbal infants (Fantz, 1964). Detecting novel stimuli embedded within more frequent background stimuli has been extensively studied using a paradigm known as the "oddball task". The ability to detect novel stimuli or expectation violations is associated with a distinct event related potential that occurs about 300 ms after the novel stimulus onset (Sutton et al., 1965). While the precise neural origins of this signal are still being worked out (Ranganath and Rainer, 2003), the hippocampus (Knight, 1996; Tulving et al., 1996), ACC (Baudena et al., 1995; Berns et al., 1997), and insula (Linden et al., 1999; Kiehl et al., 2001) have been shown to be reliably involved. Our findings support the existence of this more general system that detects violations of expectations and, importantly, extends these ideas into the domain of social interactive decision-making and neuroeconomics.

In terms of our imaging results, we also found distinct networks consistent with systems previously identified with expectation violation and memory. Viewing faces of partners whose offers exceeded expectations was associated with bilateral posterior hippocampi/parahippocampi, bilateral anterior insula, pre-SMA/ACC, and striatum. These regions have previously been associated with expectation violation, social cognition, and memory. Considerable research has demonstrated that posterior hippocampal regions are critical in successful recognition memory (Eldridge et al., 2000; Yonelinas et al., 2005). In addition, patients with hippocampal damage have been demonstrated to have a selective impairment in generating the characteristic P300 and autonomic skin response following unexpected events while the processing of expected events remained preserved (Knight, 1996). Our observed activation in the striatum, which included the caudate and nucleus accumbens is consistent with the literature on reward prediction error (Schultz et al., 1997) and repeated play with cooperators (Rilling et al., 2002). Partners that exceed initial expectations are associated with a positive prediction error, which is likely to promote further cooperation with these partners in the future (Rilling et al., 2002; Delgado et al., 2005; King-Casas et al., 2005). We also observed activity in the pre-SMA area/ACC and bilateral anterior insula. This network appears to be functionally coupled (Fox et al., 2005; Margulies et al., 2007; Craig, 2009), and has consistently been associated with detecting violations of expectation in a multitude of contexts including stimulus frequency (Braver et al., 2001), changes in sequences (Berns et al., 1997; Huettel et al., 2002) and multi-modal sensory changes (Downar et al., 2000). Thus, viewing pictures of partners who exceeded participants' expectations resulted in increased activity in regions of the brain that have been consistently associated with detecting violations of expectations in paradigms investigating more basic aspects of novelty detection and also in successful memory retrieval.

In contrast, when viewing partners that had made lower offers than the player had expected, we found activation in bilateral TPJ, right STS, bilateral posterior insula, and precuneus. These regions have been implicated in a variety of processes including memory, expectation violation, social cognition, and pain processing. The TPJ has been shown to be involved in expectation violation (Downar et al., 2000) and plays a key role in generating the brain's P300 novelty response (Knight et al., 1989) and in orienting attention (Corbetta et al., 2000). In addition, the TPJ has received attention for its role in thinking about others' mental states (i.e. theory of mind) (Saxe and Kanwisher, 2003), but it is currently unclear if these two processes can be explained by a more general cognitive process (Mitchell, 2008). Thus, viewing pictures of partners who offered less money than was expected is associated with a region of the brain that has been implicated in both social cognition and novelty detection. We also observed increased activity in the right STS, a region which has been hypothesized to detect and evaluate intentions and actions of other's behavior (Frith and Frith, 1999; Saxe et al., 2004). This region has been associated with updating expectations about an opponent's strategy based on their behavior in a repeated Inspection game (Hampton et al., 2008). In addition, the STS and TPJ have been demonstrated to be involved in social prediction error – specifically in both making a prediction about the value of a social partner's advice and updating this prediction after feedback (Behrens et al., 2008). We also observed activity in the bilateral posterior insula, which has been primarily associated with interoceptive processing, that is processing of the physiological condition of the body (Craig, 2002). This region is reliably associated with processing pain from external stimulation (Koyama et al., 2005; Singer et al., 2004) and also direct cortical stimulation (Ostrowsky et al., 2002) and suggests, at least tentatively, that viewing pictures of participants who offered less than expectations is perhaps associated with processing a negative somato-visceral state. Finally, the precuneus has been demonstrated to be involved in memory, and social cognition (Cavanna and Trimble, 2006). The range of these memory processes is diverse and includes episodic memory retrieval (Shallice et al., 1994; Tulving et al., 1994), recognition memory (Henson et al., 1999; Yonelinas et al., 2005), source memory (Lundstrom et al., 2005), and autobiographical memory retrieval (Addis et al., 2004). The precuneus has also been involved in mentalizing perceived intentionality (den Ouden et al., 2005), and in reasoning about another's beliefs (Saxe and Kanwisher, 2003). These results suggest that viewing a picture of a partner that offered less money than was initially expected is associated with brain regions that have been thought to be involved in processing negative somatic states, mentalizing about another person's beliefs, updating expectations about behavior, and memory.

Interestingly, despite the methodological differences between the present study and that of Singer et al. (2004), both studies yield somewhat similar results. Singer et al. (2004)had participants repeatedly make a gender discrimination on photographs of partners with whom they had previously encountered in a repeated Prisoner's Dilemma Game. In contrast, our study employed a single-shot design using the Ultimatum Game and a recognition task that included an equal number of old and new faces. Our imaging analyses focus on partners that made offers that were either higher or lower than the participant's initial expectation, while Singer et al. (2004) independently compared partners that were cooperators or defectors to partner's associated with null games. Despite these methodological discrepancies both studies identify the anterior insula and different components of the striatum as being linked to partners with positive associations (i.e. cooperator or positive expectation error). While Singer et al. (2004) only found vmPFC associated with defectors, we observed activity in the posterior insula, STS, TPJ. Because our study was explicitly designed to study social memory, we were also able to observe activity in regions that have previously been associated with memory retrieval – most notably the hippocampal/parahippocampal regions and precuneus. Thus, our results extend those of Singer et al. (2004), by providing a different perspective on cheater detection (i.e. expectation violation) as well as methods that are more conducive to studying social memory.

In contrast to our behavioral results, we did not observe activation in the brain for the interaction between initial expectations and expectation error. One possible reason why we failed to observe a significant finding for the imaging interaction is the combination of a stringent statistical threshold and a lack of

statistical power. While our mixed effects procedure can account for unequal variances in the interaction analysis, there is an under representation of cases in which there are low initial expectations in this sample. Thus, while our behavioral analyses utilized a p-value of $p < 0.05$, our imaging analyses were restricted to a more stringent criteria to account for multiple comparisons. Indeed, when the statistical threshold is dropped to a more liberal $p < 0.005$ uncorrected level, we find almost identical results to our expectation error analysis including bilateral anterior insula, SMA, bilateral posterior hippocampus, bilateral caudate, left ventral putamen, and bilateral amygdala.

As noted, we did not observe any evidence for a significant behavioral or imaging finding for enhanced memory for partners based on the amount of money they offered. Nor did we observe evidence of a salience detection system, in which partners who made either extremely fair or unfair offers were better remembered. The literature on cheater detection is rife with conflicting results, with some studies finding enhanced memory for cheaters (Mealy et al., 1996; Oda, 1997; Chiappe et al., 2004), others finding enhanced memory for altruists (Barclay and Lalumiere, 2006), and others, like the present study, finding no significant differences (Barclay and Lalumiere, 2006; Mehl and Buchner, 2008). The more general expectation deviation system outlined here is a potential mechanism that could account for the inconsistent results in this domain. However, at present it is not immediately clear why we identified two distinct expectation violation systems that track with the valence of the deviation. In addition, it is important to note that despite the attractiveness of the expectation violation hypothesis, our imaging results do not necessarily rule out the possibility that some of the regions associated with negative expectation violations may be involved in cheater detection. Addressing these issues could be fruitfully explored further in future research.

Like all studies, there are a number of limitations that should be considered before drawing firm conclusions from the results. First, it is always difficult to interpret null findings. Our lack of significant results for offer amount on memory cannot necessarily be interpreted as an absence of an effect. Neuroimaging studies are inherently underpowered (Mumford and Nichols, 2008) and as such are greatly at risk for making Type II error. Second, it is unclear if participants actually believed they were engaged in a real social interaction. While participants were not explicitly probed about whether they believed that they were playing with a real partner, no participant expressed any doubt, nor did their behavior deviate remarkably from other published studies that utilized actual human partners (Camerer, 2003). In addition, consistent with previous research (Sanfey, 2009), most participants indicated that they expected their partners to make fair offers. Finally, it is somewhat of an open question as to whether a single UG interaction is sufficient to label a partner as a "cheater". It is possible that making such a judgment would require multiple interactions. However, a single interaction would be enough to develop an initial impression and there is considerable evidence demonstrating that participants generate negative emotional responses in response to a single unfair offer (Sanfey et al., 2003; van 't Wout et al., 2006).

In summary, our results support a more general system that detects violations of expectations as opposed to a more specialized system engineered to detect cheaters. We found that participants on average were no better or worse at remembering partners who made either fair or unfair proposals, but rather that individual participants exhibited selectively better memory for partners who made offers which violated their initial expectations. Two dissociable neural systems were found to be underlying this effect. While both systems have been previously associated with expectation violation, social cognition, and memory, these regions tentatively suggest that there is distinct processing for positive and negative expectation violations. Positive expectation violations are associated with a system that may incorporate error detection, conscious awareness of the error, reward processing, and enhanced recognition memory, while negative expectation violations are associated with expectation violation, evaluating intentions, pain processing, and autobiographical episodic memory. By incorporating the strengths of several fields – the tasks of behavioral economics, the methodologies of psychology and the sophisticated techniques of neuroscience – we can uniquely investigate how social exchange operates, not just in terms of the immediate decisions but also how these interactions can reverberate over time.

ACKNOWLEDGMENTS

We would like to thank Mascha van't Wout and Katia Harle for their help with the collection of this data and the two anonymous reviewers for their thoughtful suggestions. This research was supported by NIMH R03MH077058 and NIA R21AG030768.

REFERENCES

Addis, D. R., McIntosh, A. R., Moscovitch, M., Crawley, A. P., and McAndrews, M. P. (2004). Characterizing spatial and temporal features of autobiographical memory retrieval networks: a partial least squares approach. *Neuroimage* 23, 1460–1471.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* 59, 340–412.

Barclay, P. (2008). Enhanced recognition of defectors depends on their rarity. *Cognition* 107, 817–828.

Barclay, P., and Lalumiere, M. L. (2006). Do people differentially remember cheaters? *Hum. Nat.* 17, 98–113.

Bates, D., Maechler, M., and Dai, B. (2008). lme4: Linear Mixed-Effects Models Using S4 Classes. R package version 0.999375-28.

Battigalli, P., and Dufwenberg, M. (2007). Guilt in games. *Am. Econ. Rev.* 97, 170–176.

Baudena, P., Halgren, E., Heit, G., and Clarke, J. M. (1995). Intracerebral potentials to rare target and distractor auditory and visual stimuli. III. Frontal cortex. *Electroencephalogr. Clin. Neurophysiol.* 94, 251–264.

Behrens, T. E., Hunt, L. T., Woolrich, M. W., and Rushworth, M. F. (2008). Associative learning of social value. *Nature* 456, 245–249.

Berns, G. S., Cohen, J. D., and Mintun, M. A. (1997). Brain regions responsive to novelty in the absence of awareness. *Science* 276, 1272–1275.

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., and Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cereb. Cortex* 11, 825–836.

Buchner, A., Bell, R., Mehl, B., and Musch, J. (2009). No enhanced recognition memory, but better source memory for faces of cheaters. *Evol. Hum. Behav.* 30, 212–234.

Camerer, C. F. (2003). Dictator, ultimatum, and trust games. In Behavioral Game Theory, C. F. Camerer, ed. (New York, Russell Sage Foundation), pp. 43–116.

Cavanna, A. E., and Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129(Pt 3), 564–583.

Chiappe, D., Brown, D., Dow, B., Koontz, J., Rodriguez, M., and McCulloch, K. (2004). Cheaters are looked at longer and remembered better than

cooperators in social exchange situations. *Evol. Psychol.* 2, 108–120.

Corbetta, M., Kincade, J. M., Ollinger, J. M., McAvoy, M. P., and Shulman, G. L. (2000). Voluntary orienting is dissociated from target detection in human posterior parietal cortex. *Nat. Neurosci.* 3, 292–297.

Cosmides, L. (1989). The logic of social exchange: has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition* 31, 187–276.

Cosmides, L., and Tooby, J. (1992). Cognitive adaptations for social exchange. In The Adapted Mind: Evolutionary Psychology and the Generation of Culture, J. Barkow, L. Cosmides and J. Tooby, eds (New York, Oxfor University Press). pp. 163–228.

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666.

Craig, A. D. (2009). How do you feel–now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70.

Crockett, M. J., Clark, L., Tabibnia, G., Lieberman, M. D., and Robbins, T. W. (2008). Serotonin modulates behavioral reactions to unfairness. *Science* 320(5884), 1739.

de Quervain, D. J., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., and Fehr, E. (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258.

Delgado, M. R., Frank, R. H., and Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618.

den Ouden, H. E., Frith, U., Frith, C., and Blakemore, S. J. (2005). Thinking about intentions. *Neuroimage* 28, 787–796.

Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E., and Ramnani, N. (2009). A probabilistic MR atlas of the human cerebellum. *Neuroimage* 46, 39–46.

Downar, J., Crawley, A. P., Mikulis, D. J., and Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nat. Neurosci.* 3, 277–283.

Eldridge, L. L., Knowlton, B. J., Furmanski, C. S., Bookheimer, S. Y., and Engel, S. A. (2000). Remembering episodes: a selective role for the hippocampus during retrieval. *Nat. Neurosci.* 3, 1149–1152.

Fantz, R. L. (1964). Visual Experience in Infants: Decreased Attention to Familiar Patterns Relative to Novel Ones. *Science* 146, 668–670.

Fehr, E., and Gachter, S. (2002). Altruistic punishment in humans. *Nature* 415, 137–140.

Fehr, E., and Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. Boston, MIT Press.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9673–9678.

Frith, C. D., and Frith, U. (1999). Interacting minds–a biological basis. *Science* 286, 1692–1695.

Gigerenzer, G., and Hug, K. (1992). Domain-specific reasoning: social contracts, cheating, and perspective change. *Cognition* 43, 127–171.

Glimcher, P. W., Camerer, C., Poldrack, R. A., and Fehr, E. eds. (2009). Neuroeconomics: Decision-making and the brain. New York, Academic Press. 538p.

Green, R. T. (1956). Surprise as a factor in the von restorff effect. *J. Exp. Psychol.* 52, 310–344.

Guth, W., Schmittberger, R., and Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* 3, 367.

Hampton, A. N., Bossaerts, P., and O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6741–6746.

Harle, K. M., and Sanfey, A. G. (2007). Incidental sadness biases social economic decisions in the Ultimatum Game. *Emotion* 7, 876–881.

Henson, R. N., Rugg, M. D., Shallice, T., Josephs, O., and Dolan, R. J. (1999). Recollection and familiarity in recognition memory: an event-related functional magnetic resonance imaging study. *J. Neurosci.* 19, 3962–3972.

Huettel, S. A., Mack, P. B., and McCarthy, G. (2002). Perceiving patterns in random series: dynamic processing of sequence in prefrontal cortex. *Nat. Neurosci.* 5, 485–490.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (trsnformation or not) and towards logit mixed models. *J. Mem. Lang.* 59, 434–446.

Kiehl, K. A., Laurens, K. R., Duty, T. L., Forster, B. B., and Liddle, P. F. (2001). Neural sources involved in auditory target detection and novelty processing: an event-related fMRI study. *Psychophysiology* 38, 133–142.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., and Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.

Kliegl, R., Risse, S., and Laubrock, J. (2007). Preview benefit and parafoveal-on-foveal effects from word $n + 2$. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1250–1255.

Knight, R. T. (1996). Contribution of human hippocampal region to novelty detection. *Nature* 383, 256–259.

Knight, R. T., Scabini, D., Woods, D. L., and Clayworth, C. C. (1989). Contributions of temporal-parietal junction to the human auditory P3. *Brain Res.* 502, 109–116.

Koenigs, M., and Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: evidence from the Ultimatum Game. *J. Neurosci.* 27, 951–956.

Koyama, T., McHaffie, J. G., Laurienti, P. J., and Coghill, R. C. (2005). The subjective experience of pain: where expectations become reality. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12950–12955.

Linden, D. E., Prvulovic, D., Formisano, E., Vollinger, M., Zanella, F. E., Goebel, R., and Dierks, T. (1999). The functional neuroanatomy of target detection: an fMRI study of visual and auditory oddball tasks. *Cereb. Cortex* 9, 815–823.

Lundstrom, B. N., Ingvar, M., and Petersson, K. M. (2005). The role of precuneus and left inferior frontal cortex during source memory episodic retrieval. *Neuroimage* 27, 824–834.

Mai, J. K., Paxinos, G., and Voss, T. (2007). Atlas of the Human Brain, 3rd Edn. New York, Academic Press.

Margulies, D. S., Kelly, A. M., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., and Milham, M. P. (2007). Mapping the functional connectivity of anterior cingulate cortex. *Neuroimage* 37, 579–588.

Mealy, L., Daood, C., and Krage, M. (1996). Enhanced memory for faces of cheaters. *Ethol. Sociobiol.* 17, 119–128.

Mehl, B., and Buchner, A. (2008). No enhanced memory for faces of cheaters. *Evol. Hum. Behav.* 29, 35–41.

Mitchell, J. P. (2008). Activity in right temporo-parietal junction is not selective for theory-of-mind. *Cereb. Cortex* 18, 262–271.

Mumford, J. A., and Nichols, T. E. (2008). Power calculation for group fMRI studies accounting for arbitrary design and temporal autocorrelation. *Neuroimage* 39, 261–268.

Newbould, R. D., Skare, S. T., Jochimsen, T. H., Alley, M. T., Moseley, M. E., Albers, G. W., and Bammer, R. (2007). Perfusion mapping with multiecho multishot parallel imaging EPI. *Magn. Reson. Med.* 58, 70–81.

Oda, R. (1997). Biased face recognition in the prisoner's dilemma game. *Evol. Hum. Behav.* 18, 309–315.

Ostrowsky, K., Magnin, M., Ryvlin, P., Isnard, J., Guenot, M., and Mauguiere, F. (2002). Representation of pain and somatic sensation in the human insula: a study of responses to direct electrical cortical stimulation. *Cereb. Cortex* 12, 376–385.

Pillutla, M. M., and Murnighan, J. K. (1996). Unfairness, anger, and spite: emotional rejection of ultimatum offers. *Organ. Behav. Hum. Decis. Process.* 68, 208.

R_Development_Core_Team (2008). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, ISBN 3-900051-07-0, Available at: http://www.R-project.org.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *Am. Econ. Rev.* 83, 1281–1302.

Ranganath, C., and Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nat. Rev. Neurosci.* 4, 193–202.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., and Kilts, C. (2002). A neural basis for social cooperation. *Neuron* 35, 395–405.

Sanfey, A. G. (2007). Social decision-making: insights from game theory and neuroscience. *Science* 318, 598–602.

Sanfey, A. G. (2009). Expectations and social decision-making: biasing effects of prior knowledge on ultimatum responses. *Mind Soc.* 8, 93–107.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science* 300, 1755–1758.

Saxe, R., and Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage* 19, 1835–1842.

Saxe, R., Xiao, D. K., Kovacs, G., Perrett, D. I., and Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia* 42, 1435–1446.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

Shallice, T., Fletcher, P., Frith, C. D., Grasby, P., Frackowiak, R. S., and Dolan, R. J. (1994). Brain regions associated with acquisition and retrieval of verbal episodic memory. *Nature* 368, 633–635.

Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., and Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron* 41, 653–662.

Stocker, T., Kellermann, T., Schneider, F., Habel, U., Amunts, K., Pieperhoff, P., Zilles, K., and Shah, N. J. (2006). Dependence of amygdala activation

on echo time: results from olfactory fMRI experiments. *Neuroimage* 30, 151–159.

Sutton, S., Braren, M., Zubin, J., and John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science* 150, 1187–1188.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.

Tulving, E., Markowitsch, H. J., Craik, F. E., Habib, R., and Houle, S. (1996). Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cereb. Cortex* 6, 71–79.

Tulving, E., Markowitsch, H. J., Kapur, S., Habib, R., and Houle, S. (1994). Novelty encoding networks in the human brain: positron emission tomography data. *Neuroreport* 5, 2525–2528.

van 't Wout, M., Kahn, R. S., Sanfey, A. G., and Aleman, A. (2006). Affective state and decision-making in the ultimatum game. *Exp. Brain Res.* 169, 564–568.

van 't Wout, M., and Sanfey, A. G. (2008). Friend or foe: the effect of implicit trustworthiness judgments in social decision-making. *Cognition* 108, 796–803.

Venables, W. N., and Ripley, B. D. (2002). Modern Applied Statistics with S, 4th Edn. New York, Springer.

von Neumann, J., and Morgenstern, O. (1944). Theory of Games and Economic Behavior. Princeton, NJ, Princeton University Press.

von Restorff, H. (1933). Über die wirkung von bereichsbildungen im spurenfeld. *Psychol. Res.* 18, 299–342.

Wallace, W. P. (1965). Review of the historical, empirical, and theoretical status of the von restorff phenomenon. *Psychol. Bull.* 63, 410–424.

Weiskopf, N., Hutton, C., Josephs, O., and Deichmann, R. (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-brain analysis at 3 T and 1.5 T. *Neuroimage* 33, 493–504.

Wickens, T. D. (2002). Elementary Signal Detection Theory. New York, Oxford University Press.

Xiao, E., and Houser, D. (2005). Emotion expression in human punishment behavior. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7398–7401.

Yonelinas, A. P., Otten, L. J., Shaw, K. N., and Rugg, M. D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *J. Neurosci.* 25, 3002–3008.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 03 June 2009; paper pending published: 22 June 2009; accepted: 27 September 2009; published online: 20 October 2009.

*Citation: Chang LJ and Sanfey AG (2009) Unforgettable ultimatums? Expectation violations promote enhanced social memory following economic bargaining. Front. Behav. Neurosci. **3**:36. doi: 10.3389/neuro.08.036.2009*

frontiers in
BEHAVIORAL NEUROSCIENCE

HYPOTHESIS AND THEORY ARTICLE
published: 08 September 2009
doi: 10.3389/neuro.08.023.2009

Altruistic learning

***Ben Seymour*[1,2]*, *Wako Yoshida*[1] *and Ray Dolan*[1]**

[1] *Wellcome Trust Centre for Neuroimaging, University College London, London, UK*
[2] *ESRC Centre for Economic Learning and Social Evolution, University College London, London, UK*

Edited by:
Daeyeol Lee, Yale University, USA

Reviewed by:
Michael J. Frank, Brown University, USA
Daeyeol Lee, Yale University, USA

***Correspondence:**
Ben Seymour, Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, 12 Queen Square, London WC1N 3BG, UK.
e-mail: bj.seymour@gmail.com

The origin of altruism remains one of the most enduring puzzles of human behaviour. Indeed, true altruism is often thought either not to exist, or to arise merely as a miscalculation of otherwise selfish behaviour. In this paper, we argue that altruism emerges directly from the way in which distinct human decision-making systems learn about rewards. Using insights provided by neurobiological accounts of human decision-making, we suggest that reinforcement learning in game-theoretic social interactions (habitisation over either individuals or games) and observational learning (either imitative of inference based) lead to altruistic behaviour. This arises not only as a result of computational efficiency in the face of processing complexity, but as a direct consequence of optimal inference in the face of uncertainty. Critically, we argue that the fact that evolutionary pressure acts not over the object of learning ('what' is learned), but over the learning systems themselves ('how' things are learned), enables the evolution of altruism despite the direct threat posed by free-riders.

Keywords: reinforcement learning, altruism, evolution, neuroeconomics, strong reciprocity, theory of mind, free-rider problem

INTRODUCTION

Many social interactions are self-beneficial if we behave positively and pro-cooperatively towards others. Opportunities to benefit from cooperation are widespread, and reflect the extrinsic fact that the natural environment is often best harvested, insofar as rewards can be accrued and threats avoided, by working together. But the decision to cooperate is not always straightforward, as in some situations it leaves us vulnerable to exploitation by others.

Game theory specifies a set of potential social interactions in which outcomes of cooperation and defection systematically differ, allowing both experimentalists and theoreticians to probe an individual's propensity for cooperation in different situations (Camerer, 2003). These outcomes typically vary in the extent to which competitive actions may seem preferable and where a short-sighted temptation to exploit the cooperativeness of others has a capacity to subvert cooperation later. Fortunately, the ability to look beyond the immediate returns of defection towards longer-term cooperation allows humans to escape from otherwise competitive equilibria, and this can be viewed as a hallmark of rational, sophisticated behaviour.

However, humans appear to behave positively towards each other in situations in which there is no capacity to benefit from long-term cooperation: for instance, when they play single games in which they never meet the same opponent again, and when their identities are kept anonymous (Fehr et al., 1993; Berg et al., 1995; Fehr and Fischbacher, 2003). This removes the capacity for both direct reciprocity (tit-for-tat) (Trivers, 1971; Axelrod, 1984), and the ability to earn a cooperative and trustworthy reputation that can be communicated by a third-party (Harbaugh, 1998; Bateson et al., 2006; Ariely and Norton, 2007). Furthermore, they will do this even if it is costly to themselves (Xiao and Houser, 2005; Henrich et al., 2006). From an economic perspective this appears to be genuinely altruistic, being strictly irrational since it incurs a direct personal cost with no conceivable long-term benefit.

Arguments against altruistic interpretations of experimentally observed behaviour include suggestions that individuals do not understand the rules of the game, are prone to misbelieve they (or their kin) will interact with opponents again in the future, or falsely infer they are being secretly observed and accordingly act to preserve their reputation in the eyes of experimenters (Smith, 1976). However, the widespread observation of altruism (both rewarding and punishing) across cultures (Henrich et al., 2001), and within meticulously designed experiments conducted by behavioural economists provide compelling support for its presence as a clear behavioural disposition. Furthermore, in fMRI experiments, altruistic actions correlate with brain activity, suggesting that they derive from some sort of intended or motivated behaviour and are not an expression of mere 'effector noise' (i.e. decision error) (de Quervain et al., 2004).

The very existence of altruism raises the difficult question as to why evolution has allowed otherwise highly sophisticated brains to behave so selflessly. This directs attention towards the decision-making systems that subserve economic and social behaviour (Lee, 2006, 2008; Behrens et al., 2009), and questions whether they are structured in such a way that yields altruism either inadvertently, or necessarily. The broader consequence is that if they do, then this reframes the question regarding the ultimate (evolutionary) causes of altruism towards the evolution of these very decision systems, and away from the phenomenological reality of altruism *per se*.

In this paper, we first review the structure of distinct human decision-making systems by considering a goal-directed (cognitive) system, a habitual system, and an innate (Pavlovian) action system and their interactions. We consider how these systems might operate in social contexts where the key problem is how to make optimal decisions when outcomes depend on the uncertainty associated with other agents and their motives. In the face of such computational complexity, we then consider how optimal actions can

be approximated by habit-based decision-making when outcomes are reliably predicted. In this context – through habits – altruism emerges as a consequence of a net economy of computational cost. We also consider the problem of evaluating the best policy when the payoff matrix is unknown but where individuals have an opportunity to learn from others. Observational learning rests upon inferences that might utilise such conspicuous attributes as their personal wealth. We frame observation as an inverse reinforcement learning problem, and consider value functions (including goals and subgoals) that are inferred from others actions, as well as by simpler strategies such as imitation. Notably, with incomplete information – a consequence of not being around to observe the long-term benefits of pro-cooperative behaviours, altruistic outcomes may be inferred as surrogate goals. In this context, altruism arises through optimal inference with incomplete information.

THE ARCHITECTURE OF DECISION-MAKING

Studies of decision-making in behavioural neuroscience and psychology have tended to concentrate on elemental decision-making problems, such as reward accrual in simple, stochastic, non-social environments. This enterprise has been very successful and has combined ingenious experimental designs with more classical focal brain lesion paradigms to yield insights into the underlying structure of decision-making systems. One key emerging insight is the likelihood that there is no singly monolithic decision-making system in the brain. Indeed, the best evidence suggest there are at least three distinct decision-making systems comprising a goal-directed, habitual, and innate (Pavlovian) system – with behavioural control being an admixture of cooperation or independence (Dickinson and Balleine, 2002; Dayan, 2008).

Goal-directed decision-making systems function by building an internal model of the environment. In the simplest case this may simply involve representing the identity of the expected outcome. In more complicated instances, it involves detailed knowledge of the structure of the environment and one's position within it. Although a goal-directed system may subsume several distinct sub-mechanisms, a wide variety of evidence suggest it localises to prefrontal cortex (Daw et al., 2006; Kim et al., 2006; Valentin et al., 2007), hippocampus (Corbit and Balleine, 2000; Kumaran and Maguire, 2006; Lengyel and Dayan, 2007) and dorsomedial striatum (Balleine and Dickinson, 1998; Corbit et al., 2003; Yin et al., 2005).

Habits, on the other hand, lack specific knowledge of the outcome of their decisions. In the parlance of computer science their values are 'cached', and represent only a scalar quantity which describes how good or bad an action is (Daw et al., 2005). In animal learning, such values are characterised by their insensitivity to devaluation: changes in state (e.g. moving from hunger to satiety) do not alter the value of the action, since there is no access to the new value of the goal (Dickinson and Balleine, 1994; Daw et al., 2005). Habits are acquired through experience, and 'rationalised' on account of their reliability in predicting rewarding outcomes. This efficiency derives entirely from the way in which they learn: rewards reinforce actions that are statistically predictive of their occurrence, with reinforced actions acquiring value through simple associative learning rules (Rescorla and Wagner, 1972; Holman, 1975; Adams and Dickinson, 1981). These are well described by Reinforcement Learning algorithms (such as Q learning and SARSA; Sutton and Barto, 1998), and localise to dorsolateral striatum (O'Doherty et al., 2003; Tricomi et al., 2009) and dopaminergic projections from substantia nigra.

Control over decisions is often dynamic and frequently transfers from goal-directed mechanisms (early in a task) to a habit-based system (late in a task). Indeed, this transfer can be manipulated by selective lesions to the neural substrates that underlie each of these systems (Balleine et al., 2009). In formalising accounts of how these systems interact current views centre on the idea of control being mediated by the respective uncertainties with which each system predicts outcomes, a view that provides a reasonable normative account of experimental findings (Daw et al., 2005). At a broader level, the evolutionary rationale for such a dual system is based on computational cost, since habits are vastly less resource demanding than goal-directed mechanisms.

Lastly, animals including humans have an innate, 'hard-wired', decision system. This is often referred to as a Pavlovian system, characterised by the expression of values and responses acquired through simple state-based associative learning. Unconditioned and conditioned Pavlovian responses represent an evolutionarily acquired behavioural repertoire that reflect basic, reliable knowledge gleaned from an organisms evolutionary history: embodying such knowledge structures that approaching sweet tasting fruit and withdrawing from bitter tasting fruit are inherently useful responses to enact. But whereas, on average, this inbuilt knowledge structure is enormously valuable to a naïve individual, it may also be a curse in the (usually) uncommon situations in which it is incorrect. The competitive (inhibitory) interaction between decisions based on experience (instrumental habit and goal-directed mechanisms) and those based on Pavlovian impulse localises to brain regions such as the amygdala and ventral striatum (Cardinal et al., 2002; Seymour and Dolan, 2008). This interaction reflects the classic tension between apparently emotional irrational and rational cognitive systems whereby the emotional expresses an apparent irrationality by way of some peculiarity of the environment.

DECISION-MAKING IN GAMES

A challenge for decision neuroscience is to understand how basic decision-making systems operate within socially interactive environments. Consider the game in **Table 1**: the repeated Prisoner's

Table 1 | An example payoff matrix of two-player Prisoner's dilemma game in which each player can choose either to 'cooperate' or 'defect'. The Left-side numbers represent the payoffs for the first player and the right-side numbers represent the payoffs for the second. Payoffs are symmetric, and chosen so that the sum of the payoffs is greatest when both choose cooperate and least when both players choose defect. However, each player earns the most if he chooses to defect when the other cooperates. Thus, the unique subgame perfect Nash Equilibrium of this game is for both players to defect.

		Player 2	
		Cooperate	**Defect**
Player 1	**Cooperate**	10/10	0/15
	Defect	15/0	1/1

dilemma. Subjects must choose between one of two actions: cooperate or defect, and their payoff depends on this and the choice of the opponent. Now consider a goal-directed, cognitive decision-making policy in the game, which has the ability to consider multiple future hypothetical scenarios (**Figure 1A**). If you neither know, nor care, what the other player does then the best strategy is to defect on the first round, since the outcome is always better regardless of what the other player does. For the same reason, even if you know what he/she will do, it is still better to defect.

However, it is also clear that in the long run, both players are better off if they cooperate: this mutually prescribes the best exploitation of environmental resources. Clearly, you need some way of both knowing that your opponent is committed to cooperation as well as a means of signalling to him/her your intention to cooperate. That is, you need to know that she is sophisticated enough to realise that cooperation is worthwhile, and you yourself need to be sophisticated enough to realise this. There is nothing truly altruistic about this, since you are both just trying to maximise your own payoff in an environment that contains another intelligent agent.

Thus, the existence of another intelligent agent in the environment makes the problem more complex than simpler decision-making problems that exploit inanimate environments. In the latter, the payoff probability usually depends fully on the observable states (they are 'fully observable Markov decision problems'; Bellman, 1957). That is, although the payoff may be probabilistic (either involving risk or ambiguity or both), your predictions depend in no way on how you came to arrive at that state in the first place. In social interactions, this assumption does not apply because outcomes depend on what the state thinks about you. If you have recently behaved uncooperatively, then this history negatively influences the payoff you expect to receive. That is, the outcome depends on unobservable states in the environment (making the problem 'partially observable'). If you find yourself in a seemingly identical state to a previous occasion, for instance playing opponent *x* in the game *y*, then the expected payoffs are not independent of how you got there, since opponent *x* may have a memory of you.

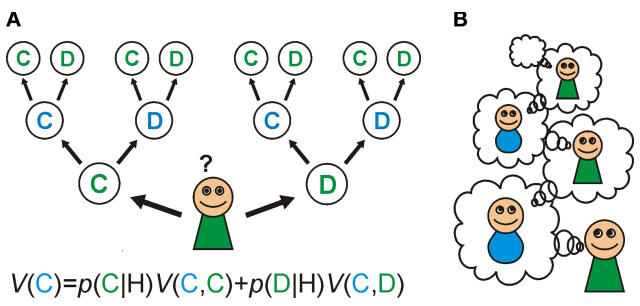


FIGURE 1 | Goal-directed learning of Prisoner's dilemma. (A) In goal-directed learning, players learn the probability of other player's action: cooperation (C) or defection (D) based on the history of their actions (H) as p(C|H) and p(D|H). They estimate the value of their own actions: V(C) and V(D) using the prediction from the learned model and the expected reward from the pair of actions. **(B)** In social games, the model of others leads a recursive process: my model of your action includes a model of your estimate of my actions, and so on. Cooperation in the Prisoner's dilemma depends on these recursive representations; since when I decide to cooperate this time, I must estimate that you are going to cooperate next time as you believe that I am going to cooperate with you.

Consequently, social decision-making benefits greatly from constructing some sort of internal model of the key aspects of the environment. In social games this model needs to capture the intentions of the other player (a component of 'Theory of Mind'). Indeed, your model should also include your opponent's estimate of your intentions: with this model, you can strategically plan to signal to your opponent your intention to cooperate, knowing that it will change their model of you (**Figure 1B**). Accordingly, they should then be more willing to cooperate with you, and you will both be better off in the long run.

It can be seen that this sort of model of others' intentions, and their model of your intentions, captures features of reciprocity, trust, and reputation formation. Indeed maintaining cooperation is in everyone's selfish interest in repeated games when the end of play is not in sight. It does, however, require players to be able to resist the short-term temptation to exploit this mutual reciprocity by the treachery of defection.

Of course, there is no reason why an internal representation of an other-agent's belief model need stop at a knowing the representation of your intentions in their mind. At the next level, it could include your understanding that they know that you know that they know your intentions, and so on. That there are infinite levels of embedded beliefs that make any perfect decision-policy intractable, has inspired models of strategic behaviour that either bound the upper limit of reciprocal beliefs (an example of 'bounded rationality') (Camerer et al., 2004a; Hampton et al., 2008), or estimate the level of reciprocal belief in their opponent directly (Yoshida et al., 2008).

Experimental evidence indicates that in repeated games with the same opponent, people reliably cooperate, as theory predicts. Critically, however, the theory predicts that people shouldn't cooperate towards the end of repeated exchanges, when they play people that they will never meet again and who can't communicate with others that can. The observation that people do cooperate in these situations suggests something is either incorrect about the goal-directed model, or as we suggest, other decision-making systems compete to bias behaviour.

HABITISATION

In simple environments, habits allow you to navigate towards goals and avoid harm with speed and computational efficiency. Habits operate by allowing recently experienced rewards to reinforce actions that are statistically predictive of them. If an outcome is reliably predicted by an action, then the value of that action becomes high. The action set available to an individual at any one time is elicited by the configuration of cues and contexts in the environment, which represents the current 'state'. Importantly, habits don't themselves have access to any specific representation of their outcome, they merely know their value on an ordinal value scale.

Now consider action control in social games. Imagine you are playing a selfish but sophisticated opponent in endless rounds of the Prisoner's dilemma. Early in the game, your model-based system has the ability to consider multiple future rounds of the game, in which mutual cooperation is evaluated as valuable, since you know your opponent also knows this. Accordingly, mutual cooperation is

rewarded as the game dictates. After a few rounds, actions associated with 'cooperate' begin to reliably predict rewarding outcomes, and so the habit learning system, operating concurrently with goal-orientated systems, acquires greater predictive certainty. As this accrues, control is transferred to the habit system, and the computational cost of considering multiple future rounds is relieved. In simple terms, cooperation becomes more 'automatic'.

The critical feature of this type of habit learning is what defines the state by which the habit can be elicited. In animal learning theory, this is termed the 'discriminative stimulus', and is typically experimentally determined by the presence of a cue (Mackintosh, 1983). However, the discriminative stimulus in social games is more complex, and in principle could be determined by the nature of the game being played (Prisoner's dilemma, stag-hunt and so on) or by the identity of the opponent. Below, we consider both possibilities:

Imagine that you ignore the identity of your opponent, and by good fortune play the prisoners dilemma with multiple cooperative opponents: i.e. you exist within a population of sophisticated cooperators (**Figure 2A**). Different types of social interaction will have distinct payoff matrices: some will benefit cooperation, others will not. If you know which game you are playing when you engage in an action, then if your action (e.g. to cooperate) is reliably rewarded it will be accessible to acquisition by a habit learning system that simply encodes that in a given game, cooperation or competition is reliably beneficial.

Indeed even if the payoff matrix is not known, for instance in a novel game in an uncertain environment, a reasonable strategy may be to play by trial and error. This entails exploring different actions and seeing what the outcome is, in which case actions can be reinforced directly by habit systems. Simulation studies demonstrate how readily cooperative equilibria can be reached by simple associative algorithms (such as Q learning) without any model-based control at all (Littman, 1994; Claus and Boutilier, 1998; Hu and Wellman, 2004).

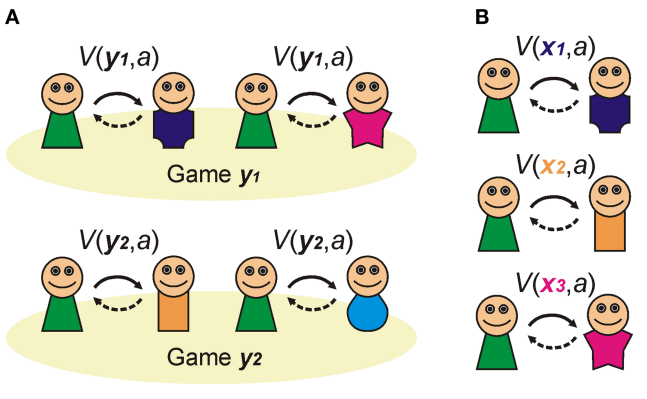


FIGURE 2 | Habitual learning of Prisoner's dilemma. (A) Habit learning in specific games. An agent plays an action *a* when in a particular state that is defined by the game type, e.g. game y1. If the outcome is rewarded, then the action is reinforced, and is more likely to be emitted when the same state is encountered again. **(B)** Habit learning with specific opponents. An agent (in green) plays action *a* (cooperate) when interacting with a particular agent (who defines the state, or discriminative stimulus S_D as x1, for example). If the outcome is rewarding, then the reward reinforces action *a*, such that it's value V(x1,a) increases, and is more likely to be chosen again in the same state in future.

Alternatively, you may choose to ignore the payoff matrix of the game, but concentrate instead on the identity of your opponent (**Figure 2B**). For instance, if you play a specific opponent in a variety of games, and she reliably cooperates with you to your benefit, then you may learn the habitual action to cooperate whenever you play her. In this way, she becomes a positive discriminative stimulus that evokes actions that engage pro-cooperatively with her.

The above mechanisms may acquire control of behaviour if several criteria are satisfied: the state and/or opponent are clearly discernable; the game (i.e. its payoff matrix) is relatively static (or changes slowly) allowing equilibria to be reached; and your internal preferences are stable. However, habit mechanisms are less reliable in the face of perceptual uncertainty, in which case an internal belief model of possible states may be required; if there are sudden changes in the environment that require rapid new learning, or a search for causal antecedents; or if your motivational state changes substantially (cooperation for food becomes less valuable when you are sated). Note that there is no evidence that habit systems 'switch off' in situations in which they behave poorly, rather their influence on control diminishes when their predictions become unreliable (Daw et al., 2005).

Although providing a plausible mechanism for social decision-making it turns out that, to date, evidence for habitised control of social behaviour is largely indirect. First, simple reinforcement learning algorithms do a remarkably good job at predicting behaviour in experiments across a variety of games (Erev and Roth, 1998, 2007). Second, neuroimaging studies show opponent-specific value-related responses accruing according to opponents' cooperativity/competitiveness in games (Singer et al., 2004). Third, neuroimaging studies have also identified dynamic reinforcement learning-like (prediction error) signals during games (King-Casas et al., 2005). Fourth, in single neuron recordings from non-human primates, lateral inter-parietal sulcus neurons in monkeys appear to encode value signals predicted by reinforcement learning in mixed-strategy games (Seo et al., 2009), which adds to previous observations that neurons in dorsolateral prefrontal and anterior cingulate cortex encode quantities related to choice and reinforcement history, respectively (Barraclough et al., 2004; Seo and Lee, 2008).

In reality, humans might be expected to habitise their actions in the context of state information that incorporates both opponent and game type. Although a diversity of subtly different payoff matrices may be common in experiments, it is likely that social interactions in different scenarios represent a relatively discrete set of payoff matrices. When there are small differences between different games, habit systems may generalise across salient features that have characteristic predictive power for beneficial outcomes.

OBSERVATIONAL LEARNING

One especially important social scenario arises when a person interacts with others who are significantly more expert at social interaction. This can occur for a number of reasons: if the payoff matrix that defines the interaction is unknown to us but known to others – either through their experience or private information; because information about other players is known to them but not to us – again through either experience or their own vicariously

acquired knowledge; of if they are more sophisticated – for instance they are more mature or intellectually able. In these situations, you have the choice to engage in interactions and acquire the information directly through your own experience or, better, to observe apparently successful social agents and vicariously acquire knowledge.

As long as success is discernable, as a hallmark of social expertise, then observational learning is likely to yield useful information. The computational problem becomes how to interpret the actions of others, and use observed actions to optimise your own. Computationally, *inverse* reinforcement learning describes this problem of how to reverse engineer observed actions to evaluate their values and goals, and is particularly difficult in situations in which actions do not immediately lead to their benefits. Unfortunately social interactions often display exactly this property: the benefits of cooperation are often long-term, through reputation formation and establishment of trust, and unless an observer has observational access to extended sequences of actions and their ultimate outcomes, the problem becomes even harder.

In general, there are two broad classes of solution. The first is simply to imitate others (Price and Boutilier, 2003). Imitation is the observational twin of habit learning, insofar as the resulting action has no specific representation of the outcome: it simply learns that a particular action is reliably performed in state s. The actions it bears are habit-like, elicited by a discriminative state that represents the environment in which they were learned. Accordingly, the ease of imitation depends on the discernability of the state of the observer. In **Figure 3A**, we illustrate this for a situation in which the state is defined by the game type: as long as it is clear to the subject that they are playing, say Game y = Prisoners Dilemma, then the imitated action will be 'cooperate when playing game y'. The imitated state-action pair could equally well be defined by the identity of the opponent. In this case, the resulting action will be 'cooperate when playing opponent x'. Note that the values of the actions can also be inferred by the frequency with which they are elicited by observation, allowing imitation to encode action values, and not just stimulus-responses.

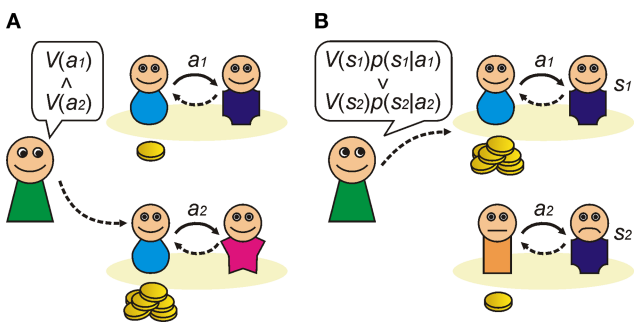


FIGURE 3 | Observation learning of Prisoner's dilemma. Observers learn the strategy from the observation of other players playing a game. **(A)** Imitation learning. An observer estimates the value of action *a* from other players' actions and simply imitates an action which maximises the payoff in a particular environment, which can be defined by game or opponent (or both). Here we show an example in which the environment is 'game = y' and it does not take into account the opponent's type (x) who they are playing with. **(B)** Inverse reinforcement learning. An observer estimates the *players'* value from their actions, for example using subgoals. This means the observer assumes that the players using a model-based learning; i.e. they have a forward model of their opponents. For example, in a repeated Prisoner's dilemma, cooperative actions (a_1) will predict a state of the other players' happiness (s_1) which leads mutual cooperation in the future. The value of action a is calculated as the value of state (e.g. other player's emotional state), $V(s)$, multiplied by the probability of occurrence of the state followed by the action, $p(s|a)$.

The second strategy is more complex, and involves trying to reverse engineer actions so as to evaluate their value or actual outcome (Ng and Russell, 2000). This requires constructing some sort of internal model of the action. For sequential actions, a computationally useful strategy is to represent subgoals – intermediate outcome states that appear to be reliable pre-requisites to eventual success (Abbeel and Ng, 2004). In the case of cooperative games, these subgoals ought to include the welfare of the other cooperators, since this is a powerful determinant of future cooperation. For example, in a repeated Prisoner's dilemma, sophisticated cooperators will themselves predict reward when their opponents cooperate with them, since they have a forward model of future beneficial interactions. Assuming their reward-predicting state is discernable by observations of their emotional state s (their happiness), then this state becomes a statistically reliable subgoal. That is, it follows that the inference that eliciting the state of happiness in another player is a valid predictor of an agent's success (**Figure 3B**).

Although in the case of the agent being observed this is merely an intermediary state in ultimately selfish reciprocal interactions, this information (and its selfishness) is not available to the observer. Even so, it is still valuable knowledge as long as the observer is fortunate enough to use the information in situations in which it actually is beneficial: i.e. in repeated social exchanges. As long as repeated social exchanges outnumber un-repeated exchanges, then observational inference is likely to be a better strategy than ignoring others.

Observational learning in games, and especially putative inverse reinforcement learning, remains relatively under-explored. It is well known that humans use both model-free (imitative) and model-based (inverse-inference) strategies when learning non-social actions through observation (Heyes and Dawson, 1990). Recent imaging evidence shows that people learn values through instruction using similar neural mechanisms involved in personal experience based learning (Behrens et al., 2008), and make inferences about values by pure third-party observation (Klucharev et al., 2009). Furthermore, pro-social feelings towards others (empathic reward), and it's neural representation, have been shown to be modulated by perceived similarity with that person (Mobbs et al., 2009), as one might predict from perspective-taking theories of social observation (Wolpert et al., 2003).

DISCUSSION

We have argued that consideration of the neurobiological mechanisms of learning and decision-making in humans can yield an explanatory account of true altruism. At the heart of this account are the learning systems that allow the brain to optimise reward and efficiency in complex environments. Critically, since evolution is likely to operate primarily over learning and decision mechanisms, and not the content of those systems – how they learn, not what they learn, the ensuing altruistic behaviours are perfectly permissible, despite the fact that they may in some instances become

strictly irrational. This is strengthened by the fact that habit-based and observational learning systems have uses way beyond social decision-making *per se*. The latter, for instance, is elegantly utilised in complex behaviours such as food preparation, tool use, and even language. Hence evolutionary selection for such mechanisms may be driven by a much broader range of decision-making problems than purely social interaction. Accordingly, such learning based accounts may offer both proximate and ultimate explanations for altruism.

The value of the inherent flexibility of learning systems is that it allows them to adapt to a wide range of potentially new and unexpected situations, appropriate for the diversity of the natural environment. But this flexibility carries the cost of inadvertently allowing individually economically disadvantageous actions to emerge, albeit rarely. However, we propose that on average these costs are heavily outweighed by benefits. Part of this supposition incorporates the fact that an innate representation of the caveats of flexible learning in social decision-making (for instance: don't cooperate in one-shot, anonymous exchanges in large groups) is itself cripplingly complex and maladaptive to novelty (it itself becomes a form of impulsivity). In other words, any social decision-making system that attempted to capture the enormous range of possible encounters and interactions, and individually specify optimal policies, would impair rather than augment decision-making under uncertainty. As such, efficient learning based systems are likely to be selected in the course of evolution.

Learning based accounts differ from the conventional approach of studying cooperation in behavioural economics, which often considers static, heuristic decision-policies, such as 'tit-for-tat', 'cooperate and punish', and 'free-ride'. Such models typically succumb to free-riders, including sophisticated (higher-order) free-riders that cooperate but don't enforce or encourage cooperation in others. However, a valuable insight of these models has been the recognition that resistance to free-riders can be provided by acquisition (and defence) of cultural norms of behaviour (Boyd and Richerson, 1988; Boyd et al., 2003; Bowles and Gintis, 2004). Key underlying components of norm-abidance are likely to be observational learning and inference based mechanisms, since these form simple elements of cultural learning. The current paucity of biologically implemented algorithmic models and mechanisms of observational and cultural learning is therefore likely to be an important area of future research. In particular, the relative privacy of culturally acquired information within specific groups is likely to be an important factor in the development of parochialism, which may further allow group-based selection of altruistic behaviour (Bernhard et al., 2006; Choi and Bowles, 2007).

Learning based accounts do not negate innate mechanisms of altruism in the brain. Such mechanisms are thought to underlie many aspects of human impulsivity and irrationality, through their occasionally inflexible competition with instrumental actions (Dayan et al., 2006). If cooperation was so consistently advantageous through human social evolution, that it is quite possible there might be some innate coding. Indeed, the environment in which the social brain evolved is likely to have had a much higher proportion of repeated interactions with the same individuals than our modern environment in which cooperation can occasionally be economically disadvantageous. Innate actions can be thought of as action priors over and above which more sophisticated goal-directed instrumental actions can assume control as experience accrues. Their Achilles heel, however, is the fact that they appear often difficult to overcome (inhibit) completely: they have a residual and significant weight that consistently biases actions in their favour. If such innate coding of cooperation exists in the human brain, then it follows that altruism would be akin to more basic forms of impulsivity.

We note that control by innate systems is characterised by the intrinsic (typically 'emotional') value of a stimulus, as well as by the action it elicits. Accordingly, the states associated with putatively pro-social innate actions could include that following the act of sharing, generosity or generation of equity (Tomasello et al., 2005). In this way, they become intrinsic internal rewards that, phenomenologically, are elicited because they are personally satisfying (and akin to non-social innate behaviours such as novelty-seeking (Wittmann et al., 2008)).

The complexity of different putative accounts of human altruism appeals to neuroscience as an arbitrator (Camerer et al., 2004b). Distinguishing different decision systems purely on anatomical grounds may be difficult, however: brain regions such as the striatum, orbitofrontal cortex, amygdala and hippocampus for instance, appear to be convergence areas for all decision systems. For example the observation of activation of striatum in a study on altruistic punishment (de Quervain et al., 2004), whilst providing a convincing illustration of the fact that such behaviour has a clear proximate basis, says little about the nature of that behaviour in terms of whether it is innate or learned. This underlines the importance for brain imaging techniques that have the ability to distinguish between competing models based on identifying coding of their underlying central parameters (O'Doherty et al., 2007), in situations in which behaviour alone is necessarily ambiguous (Yoshida et al., 2008).

Both habit-based and observation-based accounts of pro-social behaviour make specific experimental predictions. First, if the identities of others can act as discriminative stimuli, then cooperation should carry over between different games with the same individual. Second, if game types can act as discriminative stimuli, then cooperation should carry over between the same game with different individuals. Third, the duration of play should predict the degree of unfolding of cooperation towards the end of repeated games, since extended durations permit stronger habit formation and less susceptibility to anticipatory defection. Fourth, the operation of associative learning mechanisms should be determinable by the use of co-incident cues associated with previous cooperative or uncooperative players, which ought to bias individuals behaviour in future games: in fact evidence already exists for this (Vlaev and Chater, 2006; Chater et al., 2008). Fifth, observational learning can be studied directly by allowing individuals to passively watch interactions between others before engaging in similar games, or different games with the observed opponents. Indeed evidence does exist that previous observation has an influence on future social behaviour, in that people do seem to be biased towards the behaviour of others. What is more difficult to establish is exactly how this information is represented: either as a cached imitated value, or as a model-based representation.

Finally, we note that learning based accounts of altruism are by no means immune to exploitation by selfish and intelligent

learning agents. Any sophisticated model of other agents' behaviour can incorporate the fact that they are habit and observational learners. Consequently, highly sophisticated models of other agents could in theory incorporate representations of their different decision systems: thus knowing that people are habit learners gives predictive insight into what is likely to guide their behaviour in various situations. Whereas determining this might not always be simple to an agent from passive observation, it might be in part revealed by probing: intentionally behaving in a certain way (such as maliciously cultivating pro-social cultures) to manipulate how values are acquired by others, so that they can be exploited later.

REFERENCES

Abbeel, P., and Ng, A. Y. (2004). Apprenticeship learning via inverse reinforcement learning. In Proceedings of the Twenty-first International Conference on Machine Learning, Banff, Alberta, Canada. ACM International Conference Proceeding Series. New York, NY, ACM.

Adams, C. D., and Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol. B Comp. Physiol. Psychol.* 33, 109–121.

Ariely, D., and Norton, M. I. (2007). Psychology and experimental economics: a gap in abstraction. *Curr. Dir. Psychol. Sci.* 16, 336–339.

Axelrod, R. (1984). The Evolution of Cooperation. New York, Basic Books.

Balleine, B. W., and Dickinson, A. (1998). Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37, 407–419.

Balleine, B. W., Liljeholm, M., and Ostlund, S. B. (2009). The integrative function of the basal ganglia in instrumental conditioning. *Behav. Brain Res.* 199, 43–52.

Barraclough, D. J., Conroy, M. L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* 7, 404–410.

Bateson, M., Nettle, D., and Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biol. Lett.* 2, 412–414.

Behrens, T. E., Hunt, L. T., and Rushworth, M. F. (2009). The computation of social behavior. *Science* 324, 1160–1164.

Behrens, T. E. J., Hunt, L. T., Woolrich, M. W., and Rushworth, M. F. S. (2008). Associative learning of social value. *Nature* 456, 245–245.

Bellman, R. (1957). Dynamic Programming. Princeton, NJ, Princeton University Press.

Berg, J. E., Dickhaut, J., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games Econ. Behav.* 10, 122–142.

Bernhard, H., Fischbacher, U., and Fehr, E. (2006). Parochial altruism in humans. *Nature* 442, 912–915.

Bowles, S., and Gintis, H. (2004). The evolution of strong reciprocity: cooperation in heterogeneous populations. *Theor. Popul. Biol.* 65, 17–28.

Boyd, R., Gintis, H., Bowles, S., and Richerson, P. J. (2003). The evolution of altruistic punishment. *Proc. Natl. Acad. Sci. USA* 100, 3531–3535.

Boyd, R., and Richerson, P. J. (1988). The evolution of reciprocity in sizable groups. *J. Theor. Biol.* 132, 337–356.

Camerer, C. F. (2003). Behavioural Game Theory: Experiments in Strategic Interaction. Princeton, NJ, Princeton University Press.

Camerer, C. F., Ho, T. H., and Chong, J. K. (2004a). A cognitive hierarchy model of games. *Q. J. Econ.* 119, 861–898.

Camerer, C. F., Loewenstein, G., and Prelec, D. (2004b). Neuroeconomics: why economics needs brains. *Scand. J. Econ.* 106, 555–579.

Cardinal, R. N., Parkinson, J. A., Hall, J., and Everitt, B. J. (2002). Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.* 26, 321–352.

Chater, N., Vlaev, I., and Grinberg, M. (2008). A new consequence of Simpson's paradox: stable cooperation in one-shot prisoner's dilemma from populations of individualistic learners. *J. Exp. Psychol. Gen.* 137, 403–421.

Choi, J. K., and Bowles, S. (2007). The coevolution of parochial altruism and war. *Science* 318, 636–640.

Claus, C., and Boutilier, C. (1998). The dynamics of reinforcement learning in cooperative multiagent systems. In Proceedings of the Fifteenth National Conference on Artificial Intelligence. Menlo Park, CA, American Association for Artificial Intelligence, pp. 746–752.

Corbit, L. H., and Balleine, B. W. (2000). The role of the hippocampus in instrumental conditioning. *J. Neurosci.* 20, 4233–4239.

Corbit, L. H., Muir, J. L., and Balleine, B. W. (2003). Lesions of mediodorsal thalamus and anterior thalamic nuclei produce dissociable effects on instrumental conditioning in rats. *Eur. J. Neurosci.* 18, 1286–1294.

Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature* 441, 876–879.

Dayan, P. (2008). The role of value systems in decision making. In Better than Conscious? Decision Making, the Human Mind, and Implications for Institutions, C. Engel and W. Singer, eds (Frankfurt, MIT Press), pp. 51–70.

Dayan, P., Niv, Y., Seymour, B., and Daw, D. (2006). The misbehavior of value and the discipline of the will. *Neural Netw.* 19, 1153–1160.

de Quervain, D. J. F., Fischbacher, U., Treyer, V., Schelthammer, M., Schnyder, U., Buck, A., and Fehr, E. (2004). The neural basis of altruistic punishment. *Science* 305, 1254–1258.

Dickinson, A., and Balleine, B. (1994). Motivational control of goal-directed action. *Anim. Learn. Behav.* 22, 1–18.

Dickinson, A., and Balleine, B. W. (2002). The role of learning in the operation of motivational systems. In Stevens' Handbook of Experimental Psychology, 3rd Edn., Vol. 3, Learning, Motivation, and Emotion, H. Pashler and R. Gallistel, eds (New York, John Wiley & Sons).

Erev, I., and Roth, A. E. (1998). Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. *Am. Econ. Rev.* 88, 848–881.

Erev, I., and Roth, A. E. (2007). Multi-agent learning and the descriptive value of simple models. *Artif. Intell.* 171, 423–428.

Fehr, E., and Fischbacher, U. (2003). The nature of human altruism. *Nature* 425, 785–791.

Fehr, E., Kirchsteiger, A., and Riedl, A. (1993). Does fairness prevent market clearing? An experimental investigation. *Q. J. Econ.* 108, 437–459.

Hampton, A. N., Bossaerts, P., and O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. USA* 105, 6741–6746.

Harbaugh, W. T. (1998). The prestige motive for making charitable transfers. *Am. Econ. Rev.* 88, 277–282.

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., and McElreath, R. (2001). In search of Homo economicus: behavioral experiments in 15 small-scale societies. *Am. Econ. Rev.* 91, 73–78.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., and Ziker, J. (2006). Costly punishment across human societies. *Science* 312, 1767–1770.

Heyes, C. M., and Dawson, G. R. (1990). A demonstration of observational-learning in rats using a bidirectional control. *Q. J. Exp. Psychol. B Comp. Physiol. Psychol.* 42, 59–71.

Holman, E. W. (1975). Some conditions for dissociation of consummatory and instrumental behavior in rats. *Learn. Motiv.* 6, 358–366.

Hu, J. L., and Wellman, M. P. (2004). Nash Q-learning for general-sum stochastic games. *J. Mach. Learn. Res.* 4, 1039–1069.

Kim, H., Shimojo, S., and O'Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLoS Biol.* 4, e233. doi: 10.1371/journal.pbio.0040233.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., and Montague, P. R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science* 308, 78–83.

Klucharev, V., Hytonen, K., Rijpkema, M., Smidts, A., and Fernandez, G. (2009). Reinforcement learning signal predicts social conformity. *Neuron* 61, 140–151.

Kumaran, D., and Maguire, E. A. (2006). The dynamics of hippocampal activation during encoding of overlapping sequences. *Neuron* 49, 617–629.

Lee, D. (2006). Neural basis of quasi-rational decision making. *Curr. Opin. Neurobiol.* 16, 191–198.

Lee, D. (2008). Game theory and neural basis of social decision making. *Nat. Neurosci.* 11, 404–409.

Lengyel, M., and Dayan, P. (2007). Hippocampal contributions to control: The third way. *Adv. Neural Inf. Process. Syst.* 20, 889–896.

Littman, M. L. (1994). Markov games as a framework for multi-agent reinforcement learning. In Proceedings of the Eleventh International Conference on Machine Learning. San Francisco, CA, Morgan Kaufmann, pp. 157–163.

Mackintosh, N. J. (1983). Conditioning and associative learning. New York, Oxford University Press.

Mobbs, D., Yu, R. J., Meyer, M., Passamonti, L., Seymour, B.,

Calder, A. J., Schweizer, S., Frith, C. D., and Dalgleish, T. (2009). A key role for similarity in vicarious reward. *Science* 324, 900.

Ng, Y. N., and Russell, S. (2000). Algorithms for inverse reinforcement learning. In Proceedings of the Seventeenth International Conference on Machine Learning. San Francisco, CA, Morgan Kaufmann, pp. 663–670.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., and Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* 38, 329–337.

O'Doherty, J. P., Hampton, A., and Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. *Ann. NY Acad. Sci.* 1104, 35–53.

Price, B., and Boutilier, C. (2003). Accelerating reinforcement learning through implicit imitation. *J. Artif. Intell. Res.* 19, 569–629.

Rescorla, R. A., and Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In Classical Conditioning, Vol. II, A. H. Black and W. F. Prokasy, eds (New York, Appleton-Century-Crofts).

Seo, H., Barraclough, D. J., and Lee, D. (2009). Lateral intraparietal cortex and reinforcement learning during a mixed-strategy game. *J. Neurosci.* 29, 7278–7289.

Seo, H., and Lee, D. (2008). Cortical mechanisms for reinforcement learning in competitive games. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 363, 3845–3857.

Seymour, B., and Dolan, R. (2008). Emotion, decision making, and the amygdala. *Neuron* 58, 662–671.

Singer, T., Kiebel, S. J., Winston, J. S., Dolan, R. J., and Frith, C. D. (2004). Brain responses to the acquired moral status of faces. *Neuron* 41, 653–662.

Smith, A. (1976). The Theory of Moral Sentiments, D. D. Raphael and A. L. Macfie, eds (Oxford, Oxford University Press).

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

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691.

Tricomi, E., Balleine, B. W., and O'Doherty, J. P. (2009). A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.* 29, 2225–2232.

Trivers, R. (1971). The evolution of reciprocal altruism. *Q. Rev. Biol.* 46, 35–57.

Valentin, V. V., Dickinson, A., and O'Doherty, J. P. (2007). Determining the neural substrates of goal-directed learning in the human brain. *J. Neurosci.* 27, 4019–4026.

Vlaev, I., and Chater, N. (2006). Game relativity: how context influences strategic decision making. *J. Exp. Psychol. Learn. Mem. Cogn.* 32, 131–149.

Wittmann, B. C., Daw, N. D., Seymour, B., and Dolan, R. J. (2008). Striatal activity underlies novelty-based choice in humans. *Neuron* 58, 967–973.

Wolpert, D. M., Doya, K., and Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 358, 593–602.

Xiao, E., and Houser, D. (2005). Emotion expression in human punishment behavior. *Proc. Natl. Acad. Sci. USA* 102, 7398–7401.

Yin, H. H., Ostlund, S. B., Knowlton, B. J., and Balleine, B. W. (2005). The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22, 513–523.

Yoshida, W., Dolan, R. J., and Friston, K. J. (2008). Game theory of mind. *PLoS Comput. Biol.* 4(12), e1000254. doi: 10.1371/journal.pcbi.1000254.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 11 June 2009; paper pending published: 13 July 2009; accepted: 14 August 2009; published online: 08 September 2009.

*Citation: Seymour B, Yoshida W and Dolan R (2009) Altruistic learning. Front. Behav. Neurosci. **3**:23. doi: 10.3389/neuro.08.023.2009*

frontiers in
BEHAVIORAL NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 18 May 2009
doi: 10.3389/neuro.08.006.2009

Neural correlates of attitude change following positive and negative advertisements

Junko Kato[1]*†, Hiroko Ide[1†], Ikuo Kabashima[1], Hiroshi Kadota[2], Kouji Takano[2] and Kenji Kansaku[2]*

[1] *Graduate School of Law and Politics, The University of Tokyo, Bunkyo, Tokyo, Japan*
[2] *Cognitive Functions Section, Department of Rehabilitation for Sensory Functions, Research Institute of National Rehabilitation Center for Persons with Disabilities, Tokorozawa, Saitama, Japan*

Edited by:
Daeyeol Lee, Yale University, School of Medicine, USA

Reviewed by:
Ming Hsu, University of Illinois at Urbana, Champaign, USA
Daeyeol Lee, Yale University, School of Medicine, USA

Correspondence:
Dr. Junko Kato, The University of Tokyo, Graduate School of Law and Politics, 7-3-1 Hongo Bunkyo, Tokyo 113-0033, Japan.
e-mail: katoj@j.u-tokyo.ac.jp

Dr. Kenji Kansaku, Cognitive Functions Section, Department of Rehabilitation for Sensory Functions, Research Institute of National Rehabilitation Center for Persons with Disabilities, 4-1 Namiki, Tokorozawa, Saitama 359-8555, Japan.
e-mail: kansaku-kenji@rehab.go.jp

†*Junko Kato and Hiroko Ide contributed equally to this work.*

Understanding changes in attitudes towards others is critical to understanding human behaviour. Neuropolitical studies have found that the activation of emotion-related areas in the brain is linked to resilient political preferences, and neuroeconomic research has analysed the neural correlates of social preferences that favour or oppose consideration of intrinsic rewards. This study aims to identify the neural correlates in the prefrontal cortices of changes in political attitudes toward others that are linked to social cognition. Functional magnetic resonance imaging (fMRI) experiments have presented videos from previous electoral campaigns and television commercials for major cola brands and then used the subjects' self-rated affinity toward political candidates as behavioural indicators. After viewing negative campaign videos, subjects showing stronger fMRI activation in the dorsolateral prefrontal cortex lowered their ratings of the candidate they originally supported more than did those with smaller fMRI signal changes in the same region. Subjects showing stronger activation in the medial prefrontal cortex tended to increase their ratings more than did those with less activation. The same regions were not activated by viewing negative advertisements for cola. Correlations between the self-rated values and the neural signal changes underscore the metric representation of observed decisions (i.e., whether to support or not) in the brain. This indicates that neurometric analysis may contribute to the exploration of the neural correlates of daily social behaviour.

Keywords: attitude changes, fMRI, neuropolitics, neuroeconomics, human

INTRODUCTION

Exploring human nature is one of the primary motivations for investigations in neuroscience. We explored the neural mechanism involved in preference changes toward others in a situation that is relevant to social life, the exposure to positive and negative advertisements.

The (dis)favour in which others are held, and changes thereto, might affect decisions. Attitudes and attitude changes have constituted major concerns of political psychology (Mutz et al., 1996). In this regard, stable preferences towards others, such as those based on partisanship and membership in social groups, as well as preferences that are more susceptible to change constitute important subjects for study. Psychological tests to identify affinity with sociopolitical groups (such as political organizations and ethnic groups) provide reliable indicators of stable preferences. Neuropolitical experiments use stimuli that reinforce or oppose the stable political preferences of subjects. Initial studies have discovered circuitry in regions of the brain that are related to emotion (Kaplan et al., 2007; Knutson et al., 2006; Phelps et al., 2000; Westen et al., 2006). Traditionally, economic theory has assumed that people are concerned exclusively with their own interests, but the recognition of (positive and negative) concern about the welfare of others has already become the conventional wisdom in experimental and behavioural economics (Camerer, 2008; Carpenter, 2008). In neuroeconomics, social preferences are identified as divergence from purely self-interested choices in which subjects consider only their own primary rewards. Neuroeconomics experiments have identified the neural circuitry of social preferences as part of the reward-related regions of the brain (Fehr and Camerer, 2007; Lee, 2008; Loewenstein et al., 2008).

Studies on political decisions have paralleled those on economic ones (Mueller, 2003) in social science, and political scientists have increasingly considered this parallel as relevant (Katznelson and Weingast, 2005). Both political and economic decisions can be regarded as involving self- as well as other-oriented concerns. However, unlike economic self-interested behaviour, political self-interested behaviour is not necessarily defined by utility functions, and concern about others is not explicitly related to considerations of the welfare of others. Whereas economic theory defines utility functions that relate subjective goals (represented as choices) to objective values (such as the amount, probability, and time delay of reward), a variety of social behaviours, including political behaviours, might not necessarily be related to objective measures. Thus, our experiments used self-ratings of affinity towards others that political scientists have used to quantify political preferences. A recent experiment in neuroeconomics has shown that neural activity related to the acquisition of rewards tracks the subjective values of delayed monetary rewards (Kable and Glimcher, 2007). While subjective values of political preferences are not externally quantified, their changes that are focused in our experiments are metrically represented in brain activities.

Our experiments focused on the prefrontal cortices that are linked to cognitive control (Canessa et al., 2005; Lieberman, 2007; Miller and Cohen, 2001) because we hypothesized that this region contributes to changes in preferences caused by relevant stimuli that affect social cognition. Focusing on real implications, we explored the association between preference changes and neural activities by using videos from the 1992 US presidential campaign and the commercials of major cola brands for comparison. Although social scientists agree that negative campaigns affect voters' behaviours, they are divided about their influence on individual psychology and, ultimately, attitude (Lau et al., 2007). To explore the neural mechanisms, our experiment combined the neurometric analysis with two distinct behavioural observations: the binary judgment on which rival candidate (or commodity) is favoured after viewing the videos and the self-scaled affinity towards candidates in post-task questionnaires.

On the one hand, after viewing negative campaign videos, subjects showing stronger functional magnetic resonance imaging (fMRI) activation in the dorsolateral prefrontal cortex lowered their ratings for the candidate they originally supported more than did those with weaker fMRI activation in the same cortical area. On the other hand, subjects showing stronger activation in the medial prefrontal cortex tended to increase their ratings for the candidates attacked in the negative campaign videos more than did those with weaker activation. The same regions were not activated while viewing negative advertisements for cola, which were used for purposes of comparison. These results imply that neural activity after exposure to negative information about previously supported political candidates was linked to cognitive control of socially relevant stimuli. The activation of distinct prefrontal areas indicates that different kinds of cognitive controls were associated with opposite responses to negative information about the previously supported candidates, that is, they were associated with increasing and decreasing political support. fMRI signal changes in the dorsolateral prefrontal cortex showed negative correlations with changes in ratings for the candidates after viewing the negative campaign videos against them and the changes in the medial prefrontal cortex showed positive correlations. The neuronal representation of self-rated affinity towards others might lead to methodological advances in the analysis of those social behaviours that cannot be quantitatively defined by an external measurement explicitly expressed as a utility function.

MATERIALS AND METHODS

SUBJECTS AND BEHAVIOURAL TASK

Forty, healthy volunteers (8 women and 32 men, aged 18–27 years), who were native English speakers or able to understand TV news in English, were pre-assessed to exclude those ineligible for magnetic resonance (MR) scanning. All of the subjects were neurologically normal and strongly right-handed according to the Edinburgh Inventory (Oldfield, 1971). A pre-scanning questionnaire also asked their gender, age, and ideology. We recruited those who were under 30 years old at the time of the experiment (summer 2007) to avoid those who had seen presidential campaign advertisements in 1992. No one had seen any of the campaign ads we presented. All participants gave informed consent for the study, which had been approved by the Institutional Review Board.

We used videos from the 1992 US presidential campaign and ads for two cola brands for comparison. Participants spent three sessions viewing presidential campaign advertisements (Bush vs. Clinton) and three viewing cola advertisements (Coke® vs. Pepsi®) inside an MRI machine (**Figure 1**). The order of the six sessions (both political and commercial ads) was counterbalanced among the participants; about half the participants experienced the six sessions in the reverse order. After each session, they were asked which candidate (or cola brand) they favoured. For both the campaign and cola advertisements, positive advertisements about each candidate (or brand) were shown in the first session, followed by a second session of negative advertisements that attacked the candidate/brand of choice. The third session again showed positive advertisements for both sides, but

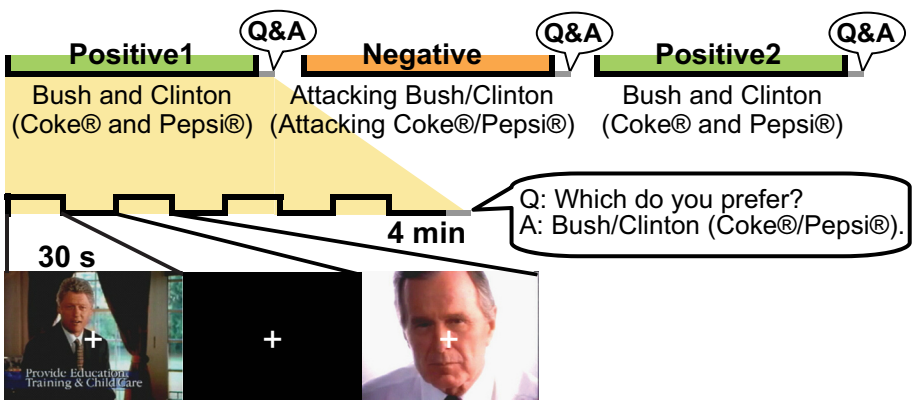


FIGURE 1 | Tasks inside the MR machine. Three sessions of campaign advertisements and three sessions of cola advertisements were shown. After each session, the participants were asked which of the two candidates/brands they favoured. For both the campaign and cola advertisements, the first session consisted of positive advertisements for both candidates/brands. The candidate/brand favoured after the first session was attacked during the second session. The third session again consisted of positive advertisements for both candidates/brands, but the content differed from that of the first session. One advertisement session consisted of four segments of 30-s advertisements with each advertisement followed by a 30-s rest period.

the content differed from the first session. One advertisement session consisted of four segments of 30-s advertisements; each advertisement was followed by a 30-s rest period. For the political campaign advertisements, participants were asked in a post-task questionnaire to rate how (un)favourable they felt towards each candidate after each session. Using an analogy with thermometry for expressing cold- and warm-heartedness, they were instructed to give a rating of 50 [in a range from 0 (least favourable) to 100 (most favourable)] when they were neutral about a candidate. This rating was built on a measure of the so-called "feeling thermometer," used by the Center for Political Studies, University of Michigan, to analyse presidential elections since 1968 (Weisberg and Miller, 1979), and studies of elections and social groups (Cairns et al., 2006).

During the experiment, each participant lay supine on the stretcher of an MR scanner (Exelart; Toshiba, Tokyo, Japan) with his/her head fixed with straps and pads inside the head coil to restrict head motion. The six sessions of advertisement videos were projected onto a screen located at the rear of the scanner and viewed through a mirror attached to the top of the head coil. The sound accompanying the advertisements was delivered through headphones. The participants wore earplugs under the headphones to minimise the MR scanner noise. Before starting the experiment, each participant confirmed that he/she could see the screen and hear the sound clearly. They were instructed to fixate on or view the images around a white cross (a fixation point) at the centre of the screen to minimise artefacts related to eye movements. The stimulus was presented and synchronised with the MR scanner using Presentation® (Neurobehavioral Systems, San Francisco, CA, USA).

IMAGING

During the video sessions, gradient echo T2* weighted echo-planar images with BOLD contrast were acquired at 1.5 T (TR/TE = 3000/40 ms, FA = 85°, slice thickness/gap = 6/2 mm, FOV = 25 × 25 cm^2, matrix size = 64 × 64, 18 slices). Each of the six sessions consisted of 85 scans, the first five of which were discarded to allow for T1 equilibration effects. T1-weighted structural images were also acquired after the video sessions.

IMAGING ANALYSIS

The imaging data were pre-processed and analysed using SPM5 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, University College London, UK). The realignment processing assured that the participant's head movement was less than 2 mm. The realigned images were then normalised to a Montreal Neurological Institute (MNI) EPI template and smoothed with an 8-mm full-width at half maximum Gaussian kernel. A high-pass filter with a cut-off period of 128 s was applied to remove low-frequency noise, and an autoregressive (order one) model was used to correct for short-range serial correlations.

A fixed-effects analysis was conducted for each participant to obtain a contrast image for each advertisement session. The contrast images of the 40 participants were then used for random effects analysis. To evaluate fMRI activation within the prefrontal cortices associated with a change in choice of the favoured candidate, we first conducted a *t*-test between the Changed and Unchanged Groups for each of the two negative and two subsequent positive advertisement sessions. In the comparison, we inclusively masked areas where the mean per cent signal change for the group that had lower activations was more than zero to avoid picking up brain areas that show deactivations predominantly in our analysis. Random effect SPM{t} maps were thresholded at an uncorrected $p < 0.05$, with a cluster- size threshold of 15 voxels. The MNI co-ordinates were converted to Talairach co-ordinates (Talairach and Tournoux, 1988) using the nonlinear transformations suggested by Brett[1]; the corresponding Brodmann areas (BA) were first assumed roughly using Talairach Daemon (Lancaster et al., 2000) and then determined using the atlas of Talairach and Tournoux (Talairach and Tournoux, 1988). We then extracted the time courses of each cluster in the prefrontal area using MarsBaR[2] for the six advertisement sessions. The average per cent signal change for each participant for each session was calculated relative to the average signal obtained during the rest period.

Using the per cent signal change data obtained from each individual, we performed a between-subjects correlation analysis for the behavioural indicators: the aforementioned self-rated preferences for the candidates (feeling thermometer). All clusters that survived the *t*-test with inclusive masking were examined. A behavioural indicator representing the degree of preference change towards a favoured candidate during a campaign advertisement session was calculated by subtracting "the pre-session rating of the favoured candidate" from "the post-session rating of the candidate." An additional indicator that represents a relative preference for a favoured candidate was calculated by subtracting "the pre-session rating of the unfavoured candidate" from "the pre-session rating of the favoured candidate." For the regions of interest, a two-sample *t*-test of the signal changes of the Changed and Unchanged Groups (two-tailed; $p < 0.05$) was also performed for each session.

RESULTS

BEHAVIOURAL ANALYSIS

The choice of favoured candidate and cola brand after each session for the political advertisements and cola advertisements are shown in **Figures 2 and 3**. Of the 34 people who chose Clinton after the first session, 14 changed their choices after seeing negative advertisements attacking Clinton, whereas four of out of six people who had chosen Bush chose Clinton after the negative advertisements. In other words, 18 of the 40 participants who saw negative advertisements attacking the candidate of their choice changed their choices after viewing the negative advertisements (**Figure 2**). Seven of the 20 people who had chosen Pepsi® changed their minds, whereas 4 out of the 20 people who had chosen Coke® changed their choices to Pepsi after viewing the negative advertisements. That is, 11 people changed their choices after viewing negative cola advertisements (**Figure 3**).

We performed Wilcoxon's rank-sum tests for sex, age, and ideology to identify statistically significant relationships between changes in choices and participants' attributes. Only age was significantly related to changes in choice, and only with regard to the second positive political session. This indicates that the younger

[1]http://imaging.mrc-cbu.cam.ac.uk/imaging/MniTalairach

[2]http://marsbar.sourceforge.net

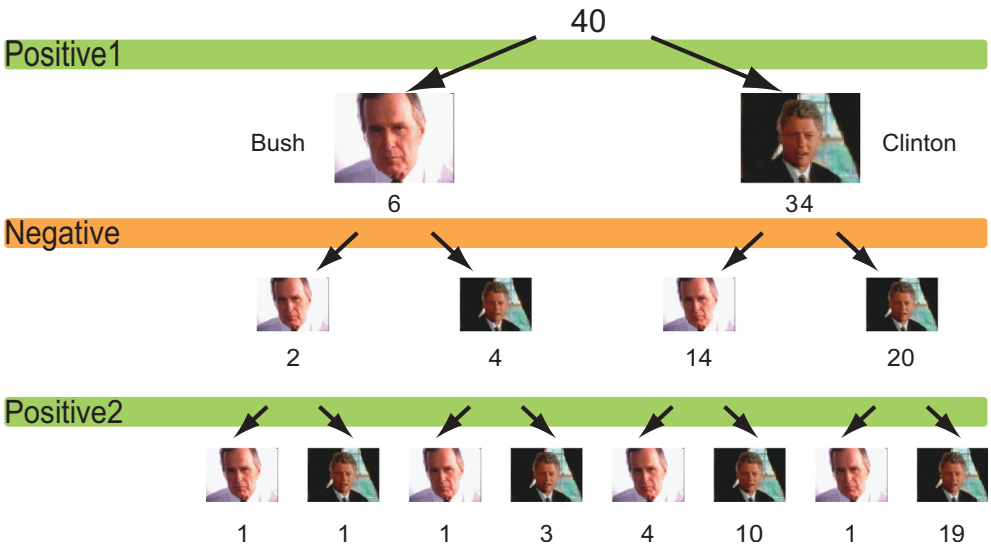


FIGURE 2 | Choice of candidates. The number of participants who chose either Bush or Clinton after each campaign advertisement session.

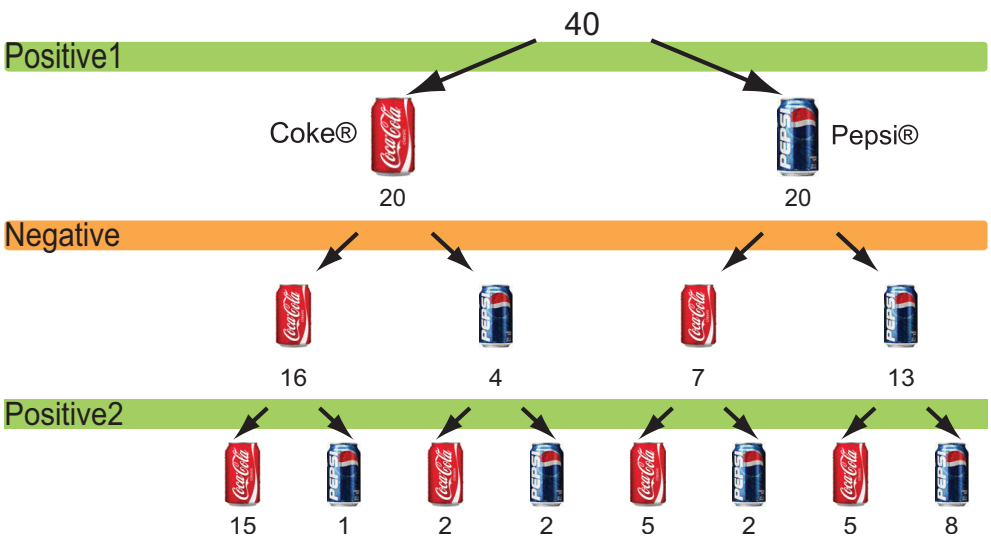


FIGURE 3 | Choice of cola brands. The number of participants who chose either cola brand after each cola advertisement session.

participants changed more often after viewing the second, positive campaign session.

We also conducted correlation analyses to examine if our main behavioural indicator, preference change for the favoured candidate during the negative campaign advertisement session (the post-session rating of the attacked candidate vs. the pre-session rating of the attacked candidate), could be explained by other behavioural indicators. We found that the indicator was negatively correlated (−0.4363; $p = 0.0049$) with the pre-session rating of preference towards the favoured candidate after the first positive advertisement session and negatively correlated (−0.3757; $p = 0.0169$) with the pre-session rating of preference for the non-favoured candidate.

NEURAL ANALYSIS

To identify brain regions in which subjective values, especially preference changes related to changes in choices, are represented metrically, we first compared those whose candidate choice did not change after viewing the advertisements (Unchanged Group) with those whose choice did change (Changed Group). The Talairach co-ordinates of the clusters that remained in this analysis are listed in **Table 1**.

As is the case in much of the literature on social cognition, our investigation focused on the prefrontal cortices. During the negative campaign advertisements, the Unchanged Group had more activations than the Changed Group in the medial prefrontal regions. The degree of preference change towards the favoured (attacked) candidate during the negative advertisement session

Table 1 | Brain areas detected in the group (random effect) analyses.

Brain region (Brodmann area)	*x*	*y*	*z*	*t*-Statistic	*p*-Value
POLITICAL NEGATIVE ADVERTISEMENTS					
Unchanged > Changed					
Left middle temporal gyrus (39)	–42	–52	12	2.81	0.004
Left middle temporal gyrus (21)	–65	–35	–8	2.70	0.005
Left lingual gyrus (18)	–8	–64	3	2.54	0.008
Right superior parietal lobule (7)	28	–52	39	2.53	0.008
Right superior frontal gyrus (8)	6	47	47	2.19	0.017
Right cuneus (19)	22	–76	33	2.12	0.020
Right superior temporal gyrus (39)	46	–52	14	2.04	0.024
Left superior frontal gyrus (8)	–16	39	44	2.03	0.024
Left cerebellum	–6	–41	–5	2.02	0.025
Right superior temporal gyrus (41)	42	–36	9	2.00	0.026
Right superior parietal lobule (7)	36	–67	49	1.97	0.027
Left cuneus (18)	–14	–84	23	1.93	0.030
Changed > Unchanged					
Right superior temporal gyrus (22)	65	–4	6	4.32	0.000
Left cerebellum	–34	–82	–16	3.41	0.001
Right occipital gyrus (19)	40	–78	4	4.10	0.000
Left inferior/middle frontal gyrus (9/6)	–42	16	40	3.99	0.000
Right cerebellum	14	–26	–14	3.55	0.001
Left superior temporal gyrus (38)	–50	14	–21	3.35	0.001
Right inferior/middle frontal gyrus (46/9)	53	30	11	3.30	0.001
Left cuneus (19)	–30	–88	28	3.10	0.002
Left superior parietal lobule (7)	–24	–71	55	2.63	0.006
Left inferior frontal gyrus (46)	–46	30	13	2.51	0.008
Left precentral gyrus (4)	–63	–10	28	2.36	0.012
Right inferior frontal gyrus (47)	42	28	–18	2.35	0.012
Right cuneus (19)	14	–94	27	2.33	0.012
Right fusiform gyrus (37)	53	–59	–16	2.29	0.014
Left superior temporal gyrus (38)	–36	5	–25	2.15	0.019
Left middle temporal gyrus (22)	–53	–41	4	2.14	0.019
Right superior parietal lobule (7)	16	–63	53	2.05	0.023
Right cerebellum	12	–67	–13	2.03	0.024
POLITICAL POSITIVE2 ADVERTISEMENTS					
Unchanged > Changed					
Left precuneus (7)	–16	–77	46	3.90	0.000
Right posterior cingulate (29)	10	–42	8	3.37	0.001
Right lingual gyrus/cuneus (18)	4	–84	–11	2.89	0.003
Left fusiform gyrus (19)	–22	–59	–9	2.47	0.009
Anterior cingulate (32)	0	25	–11	2.46	0.009
Right fusiform gyrus (19)	24	–55	–7	2.45	0.009
Left middle temporal gyrus (21)	–65	–37	–8	2.43	0.010
Right fusiform gyrus (37)	46	–59	–14	2.34	0.012
Right superior temporal gyrus (38)	44	10	–27	2.34	0.012
Left fusiform gyrus (19)	–44	–74	–11	2.30	0.013
Left cerebellum	–34	–34	–24	2.27	0.014
Right middle temporal gyrus (21)	46	–14	–9	2.13	0.020
Right superior/middle temporal gyrus (21/22)	48	–29	1	2.11	0.021
Left hippocampus	–30	–16	–11	2.00	0.026
Right lingual gyrus (18)	26	–74	–8	1.96	0.028
Right supramarginal gyrus (40)	57	–50	19	1.96	0.029

(Continued)

Table 1 | ***Continued***

Brain region (Brodmann area)	*x*	*y*	*z*	*t*-Statistic	*p*-Value
Left middle occipital gyrus (19)	–40	–71	20	1.92	0.031
Left precuneus (19)	–28	–70	35	1.87	0.034
Changed > Unchanged					
Left middle/superior temporal gyrus (21/38)	–57	3	–10	4.02	0.000
Right medial frontal gyrus (10)	16	61	6	3.24	0.001
Left transverse temporal gyrus (41)	–42	–27	12	3.09	0.002
Left inferior frontal/precentral gyrus (9/6)	–32	7	31	2.95	0.003
Right inferior parietal lobule (7)	38	–56	54	2.85	0.004
Left middle temporal gyrus (37)	–48	–56	–1	2.75	0.005
Right superior temporal gyrus (42)	51	–17	5	2.64	0.006
Right cerebellum	34	–60	–27	2.50	0.008
Right inferior frontal gyrus (46)	51	41	11	2.44	0.010
Left precentral gyrus (6)	–32	–14	67	2.42	0.010
Right parahippocampal gyrus (28)	20	–11	–25	2.42	0.010
Left precuneus (7)	–28	–48	43	2.34	0.012
Right middle occipital gyrus (18)	34	–75	11	2.22	0.016
Right superior temporal gyrus (22/42)	67	–31	11	2.16	0.018
Left cerebellum	–26	–44	–31	2.07	0.022
Right inferior frontal gyrus (9)	42	9	29	2.07	0.023
Right precuneus (7)	22	–56	36	2.05	0.023
Right inferior frontal gyrus (11)	24	27	–13	2.04	0.024
Left middle frontal gyrus (11)	–16	27	–11	2.01	0.026
Left precuneus (7)	–20	–57	34	2.00	0.026
Right cingulate gyrus (23)	6	–40	24	1.90	0.032
COLA NEGATIVE ADVERTISEMENTS					
Unchanged > Changed					
Right middle/superior temporal gyrus (22/21/42)	51	–43	4	3.90	0.000
Right middle frontal gyrus (9)	57	17	32	3.43	0.001
Right superior temporal gyrus (38)	44	–1	–15	3.28	0.001
Right superior frontal gyrus (8/9)	12	45	44	3.11	0.002
Left superior temporal gyrus (40/42)	–46	–46	21	3.05	0.002
Right middle frontal gyrus (6)	44	0	39	2.97	0.003
Right middle temporal gyrus (21)	55	–16	–9	2.91	0.003
Left inferior frontal gyrus (46)	–53	32	9	2.72	0.005
Left superior/medial frontal gyrus (9)	–4	58	30	2.68	0.005
Left superior occipital gyrus (19)	–46	–81	19	2.67	0.006
Left parahippocampal gyrus (35)	–18	–33	–7	2.36	0.012
Right lingual gyrus (18)	6	–86	–4	2.27	0.014
Left precentral gyrus (4/6)	–40	–6	44	2.16	0.018
Right parahippocampal gyrus	22	–16	–11	2.14	0.019
Left superior frontal gyrus (8)	–8	38	52	2.12	0.020
Right superior parietal lobule (7)	14	–57	67	2.09	0.021
Right middle frontal gyrus (6)	28	11	62	2.07	0.022
Left precentral gyrus (6)	–57	7	33	1.90	0.032
Changed > Unchanged					
Right cuneus (18)	14	–99	3	3.21	0.001
Right thalamus	16	–19	8	2.91	0.003
Right precentral/inferior frontal gyrus (6/9)	30	7	25	2.68	0.005
Right cerebellum	34	–56	–24	2.67	0.005
Left cuneus (18)	–12	–101	9	2.35	0.012
Left thalamus	–22	–21	10	2.30	0.014

(Continued)

Table 1 | ***Continued***

Brain region (Brodmann area)	*x*	*y*	*z*	*t*-Statistic	*p*-Value
Left cerebellum	–34	–67	–25	2.14	0.019
Posterior cingulate (29)	0	–50	12	2.01	0.025
COLA POSITIVE2 ADVERTISEMENTS					
Unchanged > Changed					
Right middle frontal gyrus (9/8)	36	22	19	2.18	0.018
Left inferior parietal lobule (40)	–65	–34	27	2.15	0.019
Right middle temporal gyrus (21)	51	6	–29	2.02	0.025
Changed > Unchanged					
Left precuneus/cuneus (7/19)	–10	–79	43	5.56	0.000
Right inferior frontal gyrus (9)	46	–1	22	3.96	0.000
Right middle frontal gyrus (8)	38	35	42	3.11	0.002
Right inferior occipital gyrus (19)	44	–76	–5	3.10	0.002
Left middle frontal gyrus (6/8)	–40	22	47	2.77	0.004
Left putamen	–24	13	–7	2.74	0.005
Left inferior frontal gyrus (9)	–50	5	31	2.65	0.006
Left supramarginal gyrus/inferior parietal lobule (40)	–42	–43	30	2.65	0.006
Left postcentral/precentral gyrus (3/4)	–32	–25	45	2.60	0.007
Left inferior frontal gyrus (47)	–32	35	–5	2.52	0.008
Right precuneus (7/19)	26	–70	35	2.48	0.009
Right cingulate gyrus (31)	22	–55	19	2.42	0.010
Right fusiform gyrus (37)	50	–49	–14	2.39	0.011
Left cingulate gyrus (24/31)	–8	–15	45	2.34	0.012
Left superior temporal gyrus (22/42)	–46	–19	–1	2.23	0.016
Right thalamus	18	–25	–4	2.22	0.016
Right inferior frontal gyrus (47)	30	18	–19	2.21	0.016
Right inferior frontal gyrus (45)	46	22	8	2.19	0.017
Left cuneus (18)	–8	–95	12	2.18	0.018
Right precentral gyrus (6)	38	–5	56	2.14	0.019
Right superior temporal gyrus (21)	53	–4	–12	2.10	0.021
Left inferior frontal gyrus (47)	–48	21	–3	2.03	0.024
Left inferior parietal lobule (40)	–48	–36	17	2.00	0.026
Left superior parietal lobule (5)	–22	–44	59	1.98	0.027
Left middle temporal gyrus (21)	–65	–16	–4	1.90	0.032
Right postcentral gyrus (2)	36	–27	40	1.88	0.033

t-test, uncorrected p < 0.05. Inclusively masked with the areas where the mean per cent signal change for the group that had lower activations was greater than zero. Cluster size of more than 15 voxels.

had a significant positive correlation with the per cent signal change in one of these regions (Talairach co-ordinate: –16, 39, 44; BA8; **Figure 4A**).

By contrast, the Changed Group had more activations than the Unchanged Group in the right and left dorsolateral prefrontal regions, and the same degree of preference change towards the favoured (attacked) candidate during the negative advertisement session had significant negative correlations with the per cent signal change in these two regions [(–42, 16, 40; BA9/6) and (53, 30, 11; BA46/9), respectively; **Figures 5A and 6A**]. Among the prefrontal regions that showed significantly different activations between the Unchanged Group and Changed Group (uncorrected $p < 0.05$), only these three regions exhibited significant correlations with our preference-related indicators.

Furthermore, an additional preference rating, the relative preference for the favoured candidate measured before the negative advertisements was significantly positively correlated with the signal change in the left medial prefrontal region (**Figure 7**), but not with that in the two dorsolateral prefrontal regions.

For these three regions, we also compared the signal change in the Unchanged and Changed Groups for all six advertisement sessions to see if significantly different activations also occurred in other sessions. We found that the left medial prefrontal region was activated significantly more in the Unchanged Group during the negative political advertisement session only (**Figure 4B**) This indicates that the activation of this part of the brain was specific to the political task in our experiment and to the negative advertisements among political sessions. Conversely, the

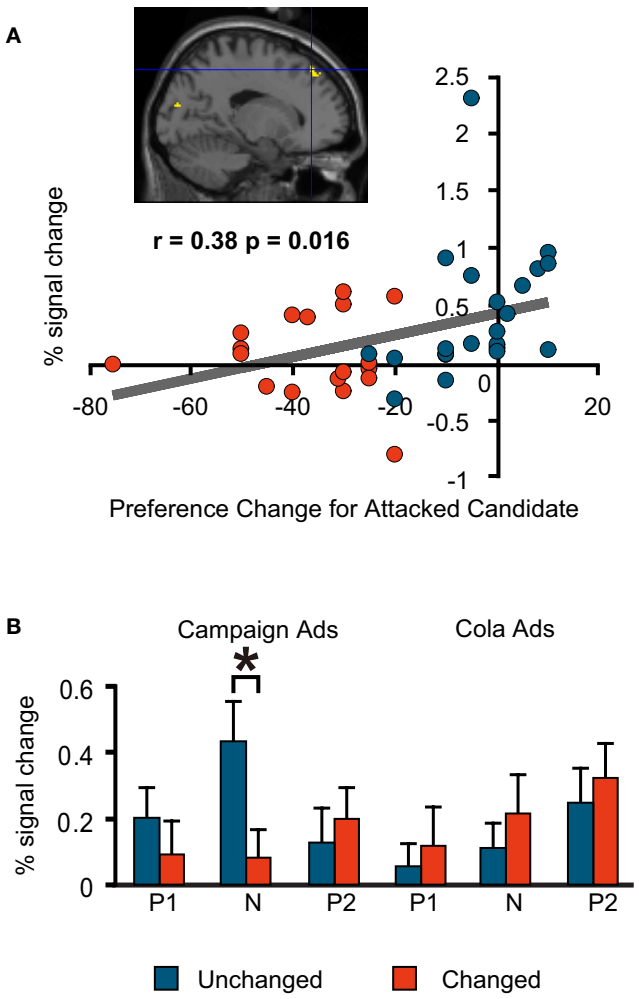


FIGURE 4 | fMRI signal change at the medial prefrontal cortex (−16, 39, 44; BA8). (A) Correlation between the signal change during the negative advertisements and preference change towards a favoured candidate during the negative advertisements (=post-negative-advertisement preference for the attacked candidate − pre-negative-advertisement preference for the attacked candidate). Blue, Unchanged; red, Changed. Note that each plot represents an individual subject. **(B)** Mean comparison between the Unchanged and Changed Groups for each advertisement session. P1, Positive1; N, Negative; P2, Positive2. The asterisk indicates a significant difference at $p < 0.05$ (*t*-test). The error bars are the S.E.M.)

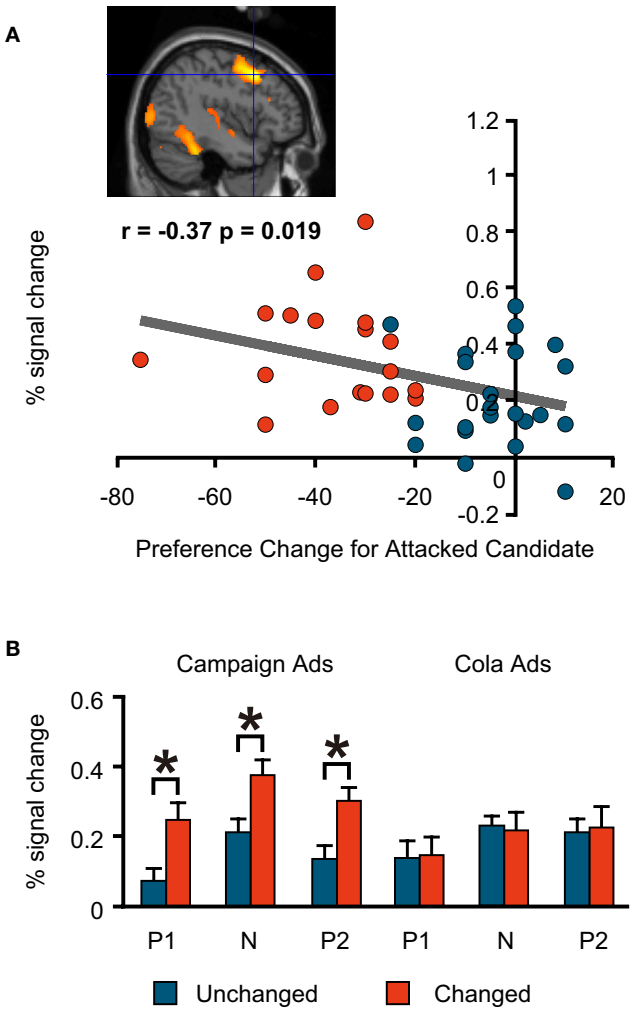


FIGURE 5 | fMRI signal change at the left dorsolateral prefrontal cortex (−42, 16, 40; BA9/6). (A) The correlation between the signal change during the negative advertisements and the preference change towards a favoured candidate during the negative advertisements (=post-negative-advertisement preference for the attacked candidate − pre-negative-advertisement preference for the attacked candidate). **(B)** Mean comparison between the Unchanged and Changed Groups for each advertisement session.

dorsolateral prefrontal regions were activated significantly more in the Changed Group in all political sessions, but not in the cola sessions (**Figures 5B and 6B**); therefore, these areas were politics-specific, but not negative-specific in our tasks. In addition, prefrontal regions that were activated more in the Unchanged Group during the cola negative advertisements were found, but at co-ordinates that were different from those found for the political negative advertisements (for the co-ordinates, see **Table 1**). These findings also indicate that the prefrontal regions had different associations with preference change in response to the negative campaign advertisements and the negative cola advertisements.

During the second sessions of positive advertisements for both the campaign and cola advertisements, the Changed Group showed more activations in the dorsolateral prefrontal cortices (51, 41, 11; BA46), (42, 9, 29; BA9) and (−32, 7, 31; BA9/6) for campaign and (46, −1, 22; BA9) and (−50, 5, 31; BA9) for cola. Activation in the dorsolateral prefrontal cortex (36, 22, 19; BA8/9) was also found in the Unchanged Group, but only during the cola session. As in the case of the negative campaign advertisements, the per cent signal change in the right dorsolateral prefrontal cortex (51, 41, 11; BA46) during the second, positive campaign advertisements had a negative correlation with the preference change for the candidate that had been favoured before the second, positive advertisements (**Figure 8**). In addition to the dorsolateral prefrontal cortices, the reward-related ventromedial prefrontal area (16, 61, 6; BA10) was found to be activated more in the Changed Group during the second, positive campaign advertisements. The area did not show correlated activities with the self-scaled values, but survived a correction for

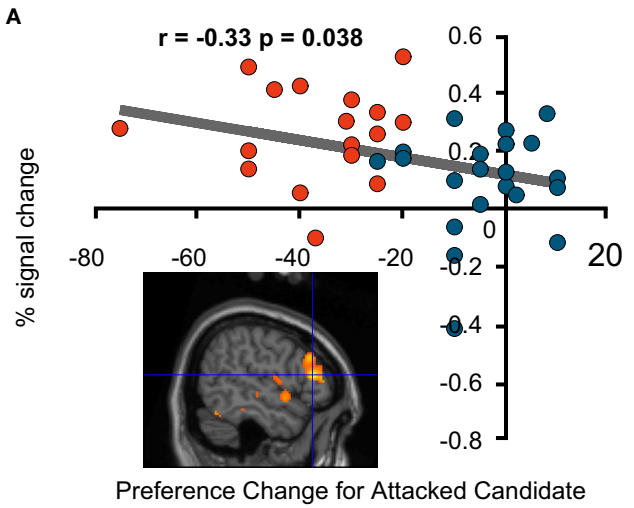


FIGURE 6 | fMRI signal change at the right dorsolateral prefrontal cortex (53, 30, 11; BA46/9). (A) Correlation between the signal change during the negative advertisements and the preference change towards a favoured candidate during the negative advertisements (=post-negative-advertisement preference for the attacked candidate – pre-negative-advertisement preference for the attacked candidate). **(B)** Mean comparison between the Unchanged and Changed Groups for each advertisement session.

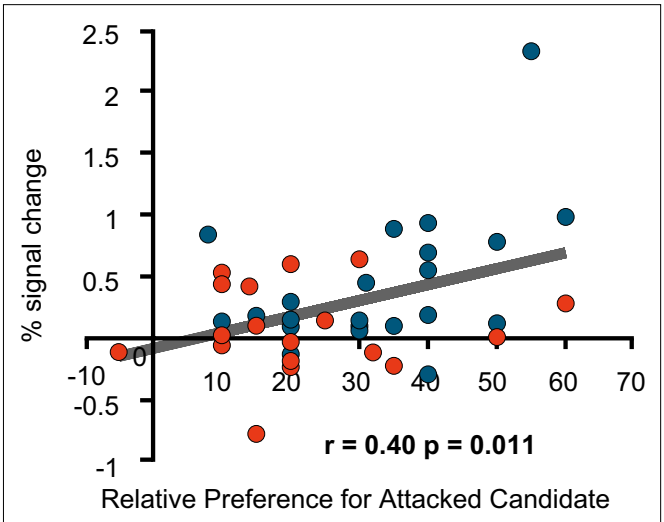


FIGURE 7 | fMRI signal change at the medial prefrontal cortex. The figure shows the correlation between the signal change at (–16, 39, 44; BA8) during the negative advertisements and the relative preference for a favoured candidate before the negative advertisements (=the pre-session rating of the attacked candidate – the pre-session rating of an attacking candidate).

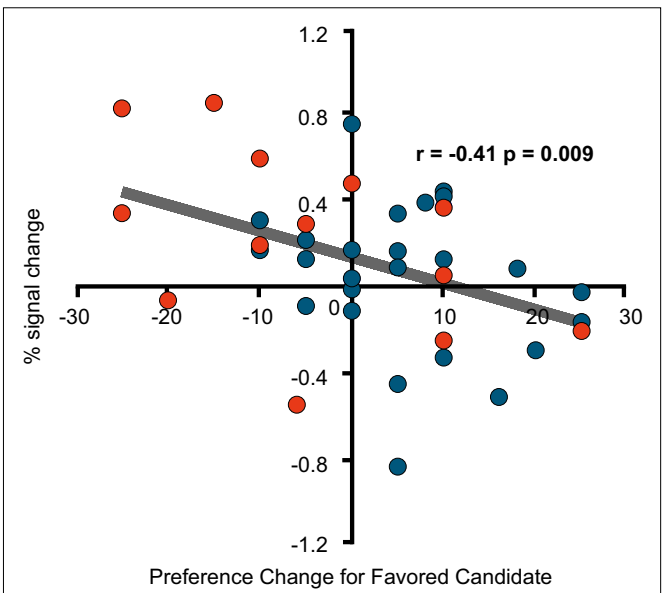


FIGURE 8 | fMRI signal change at the right dorsolateral prefrontal cortex. Correlation between the signal change at (51, 41, 11; BA46) during the second, positive advertisements and the preference change for the favoured candidate during the second, positive advertisements [=the post-session rating of the (originally) favoured candidate – the pre-session rating of the favoured candidate].

multiple comparisons using a false discovery rate when a one-sample *t*-test ($p < 0.05$) was conducted for the Changed Group.

DISCUSSION

We demonstrated that neural activation in several different regions in the prefrontal cortex traced a change in subjective values of affinity towards political candidates. The findings involved implications for understanding preference changes towards others after receiving negative information on them. Neural signal changes in the lateral prefrontal cortex had significant negative correlations with an increase in preference for the supported candidate after viewing the negative campaign videos. Changes in the medial prefrontal cortex had positive correlations. The medial and lateral prefrontal cortices were associated with opposite responses, i.e., continued support or changing sides, respectively, in our experiment. Thus, we confirmed the neural correlates of two critical elements in real politics: a binary choice between two competing alternatives (i.e., maintaining or switching support) and a transformation in preference that predicates the choice.

ACTIVATION IN THE MEDIAL PREFRONTAL CORTEX

The prefrontal cortex, where we found the activated regions, is thought to be associated with social cognitive control (Canessa et al., 2005; Lieberman, 2007; Miller and Cohen, 2001). The medial prefrontal cortex, especially the posterior rostral medial prefrontal cortex, is implicated in conflict monitoring (Amodio and Frith, 2006). Deductive reasoning is thought to be important

when selecting one among conflicting alternatives (Goel, 2007). In published experiments, activation of BA8 was reported when monitoring a conflict among competing alternatives in the face of uncertainty. In predicting a winner between two UFOs with different colours, shapes, and figures, BA8 was commonly recruited in the face of uncertainty that derived from varying winning probabilities, whereas different strategies to cope with uncertainty might recruit different neural circuits (Volz et al., 2004, 2005). The medial BA8 was also recruited in a task that required the individual to judge the validity of an argument and select one from among alternatives that are different in deductive complexity. The activated region (−16, 39, 44: BA8) among the Unchanged Group in our experiment is similar to the one used when subjects were required to judge the validity of logical statements matched in linguistic complexity by using deduction (Monti et al., 2007). The nature of the behaviour framed in previous experiments parallels the behaviour of our experiment in the sense that the subjects were required to make a binary choice based on deduction when facing uncertainty. The neural correlate implies that the Unchanged Group chose to continue to support a favoured candidate based on a deductive examination of conflicting information presented in the negative advertisements. The activation of BA8 was observed in the Unchanged Group during politically negative advertisements, but not during positive ones or in the cola tasks.

ACTIVATION IN THE DORSOLATERAL PREFRONTAL CORTEX

Induction was also adopted as a strategy for making a decision if an explicit pattern in uncertainty did not emerge, that is, the stimulus (i.e., the message in the campaign advertisements) did not necessarily control the level of uncertainty by varying probabilities (Volz et al., 2004, 2005) or logical (in)consistency (Monti et al., 2007). The region activated among the Changed Group was the bilateral dorsolateral prefrontal cortices (BA9/6, BA46/9), which are associated with induction rather than deduction (Goel and Dolan, 2004) and with working memory (Miller and Cohen, 2001). More specifically, the activation in the dorsolateral prefrontal cortex is thought to increase when attention is focussed on goal-relevant stimuli while minimising distraction from cross-modal stimuli (Weissman et al., 2004). The bilateral regions where activation was observed in the published experiments were similar to the brain regions (−42, 16, 40; 53, 30, 11) in our experiment among the Changed Group. After viewing the negative videos, those in the Changed Group were thought to have used induction to make a binary choice.

A recent finding plausibly extends the role of distraction minimisation in a social context when one person's intention is inconsistent with the second person's behaviour (Weissman et al., 2008). BA9 is recruited in tasks that involve the inconsistency of others' intentions. In a previous experiment, subjects were required to view an animation in which a boy followed or did not follow a verbal instruction from a woman standing behind him to touch a blinking dial or a different part of the table. BA9 was recruited when the boy did not follow the woman's instruction (Weissman et al., 2008). In our experiment, a similar region (−42, 16, 40) was recruited when the subjects accepted the negative information regarding the favoured candidate provided by a rival (and consequently changed their minds). In both experiments, subjects found a "social" situation when viewing contradictory behaviours on the part of two characters in two settings, i.e., animated characters and political figures in campaign advertisements. The characters in both videos were regarded as social creatures, so that the inconsistency in their behaviour (i.e., the boy going against the woman's instruction and the rival's attack on the favoured candidate) may be interpreted as ensuing from their contradictory intentions. The neural correlate was also thought to be capable of distinguishing the inconsistency in a social situation from the inconsistency of a single person's behaviour. The posterior superior temporal sulcus, instead of BA9, was recruited in the task that involved viewing an animated character reaching to grasp a blinking dial (correct or expected) or an empty space (incorrect or unexpected; Pelphrey et al., 2004).

Activation of these regions was observed in the political tasks, regardless of the positive and negative advertisements. The results imply that the Changed Group considered the information in the latest political advertisements to be more relevant.

ACTIVATION DURING THE SECOND POSITIVE ADVERTISEMENTS

During the second positive advertisements we found activation in the bilateral posterior dorsolateral prefrontal cortices among the Changed groups in both the cola and political tasks. The bilateral dorsolateral prefrontal cortices are also recruited in minimising cross-modal distraction (Weissman et al., 2004), in coordinating between different tasks (Derrfuss et al., 2004) and deciding to gamble with the feedback of varying probabilities of winning (Satterthwaite et al., 2007). The regions activated in the posterior part of the dorsolateral prefrontal cortices in the published works are similar to those activated in the Changed Group during the second positive advertisement sessions for both the cola and the political figures.

For the right anterior dorsolateral prefrontal cortex, however, the activation was observed among the Changed group during the second positive political advertisement, but among the Unchanged Group during the second positive cola session (**Table 1**). The anterior and posterior dorsolateral prefrontal cortices have been interpreted as being associated with response selection from memory and working memory maintenance, respectively (Duncan and Owen, 2000; Rowe et al., 2000). According to the literature (Wagner et al., 2001), tasks requiring memory selection (i.e., semantic comparisons of three words based on subjective desirability) recruited a similar region in the anterior part of the right dorsolateral prefrontal cortex, where our experiment found a negative correlation between the neural signal change and preference change towards others.

Smaller preference changes after the second positive political advertisements may have resulted from memory selection, whereas larger changes accompanying switching support from one candidate to another may have resulted from memory maintenance. The activation associated with the positive political advertisement may be linked to a form of cognitive control distinct from the one for the negative advertisement, although the social implication of cognitive control has not been specified in its entirety.

COMPLEMENTARY RELATIONSHIP BETWEEN BEHAVIOURAL AND NEUROMETRIC ANALYSIS

In our experimental context, the Unchanged Group examined the negative information deductively, but the Changed Group

used induction and considered the newer information to be more relevant, regardless of the negative or positive implication. The neural correlate, however, hinges on the task used in the experiment. More specifically, activation of the medial prefrontal cortex might be associated with preference changes if deductive reasoning supports the validity of the negative information; the activation of the lateral prefrontal cortex might be associated with rejection of the negative message by induction. These neural processes do not necessarily result in a one-to-one correspondence among observed attitudes, choices, and decisions. In this regard, the behavioural data analysis provides important supplementary evidence to the neurometric analysis. For example, greater activation of the medial prefrontal region (medial BA8) was also observed among those with a relatively high preference for the favoured candidate vs. the candidate who was not fovoured before the negative advertisements, and this group did not necessarily overlap with the Unchanged Group. A preference gap before the negative advertisements had a weak, statistically insignificant correlation with a drop in preference for the attacked (originally favoured) candidate after the negative advertisements (0.1; $p = 0.5392$). This implies that the activation was related to deductive judgement rather than to the presence or absence of preference changes that demarcated the Changed and Unchanged Groups. Similar to the Unchanged Group, those with a more discriminating preference scaling used a deductive approach to forming judgements. The neural circuitry should be considered linked to a specific form of cognitive control rather than to a specific choice of behaviour, represented by switching or maintaining support.

The additional behavioural data analysis provided evidence that those more detached from the objects (i.e., political candidates) are less susceptible to changes after newer stimuli. An increase (+) and decrease (−) in preference for the (originally) favoured candidate during the negative advertisements were negatively correlated with the preference towards the favoured candidate before the negative advertisements. This change in preference for the (originally) favoured candidate was also negatively correlated with the preference towards the non-favoured candidate before the negative advertisements. These results imply that those who generally reported higher preferences for both were more likely to drop the preference for the favoured candidate after the negative advertisements (and were more likely to change their minds). However, those who originally reported relatively low preferences for both candidates tended to report fewer changes during negative advertisements. Here, the stable preference plausibly resulted from a generally low level of affinity, regardless of preference order, rather than a greater attachment to the favoured candidate. To choose one among others while weighing often conflicting information constitutes an important social capability known as "social intelligence." Experimental psychologists believe that those who are more sensitive to information that potentially reveals a lack of trustworthiness in others tend to maintain once formed trust (Ostrom and Walker, 2003; Yamagishi et al., 1999). This implication is consistent with the observation of fewer preference changes among those who were detached from both candidates. This is also consistent with the neural correlates of deductive reasoning among the Unchanged Group in our experiments.

Alternatively, neural correlates often have implications for understanding attitudes, choices, and decisions in our social lives. Our experiment used videos from a past electoral campaign that was not related to the subjects' immediate political experience, so that stable preferences, such as partisanship, should not have been involved. This is distinguished from a neuropolitical experiment in which the stimulus has immediate political relevance for subjects with clear partisan loyalties, and thus it might explain the activation of the region that is linked to cognitive control rather than to emotional regulation. The activation of the region linked to emotion might result from preservation of deeply rooted preferences relating to partisan ideology and belonging to specific social groups (Kaplan et al., 2007; Knutson et al., 2006; Phelps et al., 2000; Westen et al., 2006).

Comparisons of experimental results obtained during the political and cola sessions suggest curious implications with regard to social behaviour. The same regions were not activated when negative commercial advertisements were viewed. Our results suggest the possibility that commercial advertisements might not be equivalent to political advertisements as social stimuli.

ACTIVATION IN REWARD-RELATED REGIONS

In neuroeconomics, the reward-related region of the brain is thought to be associated with social preferences that also influence decision making about one's own rewards. The pleasure derived from achieving social motives was regarded as similar to the pleasure ensuing from one's own reward (Fehr and Camerer, 2007). From among the reward-related regions, our experiments found activation in the ventromedial prefrontal cortex (Fehr and Camerer, 2007; Lee, 2008; Loewenstein et al., 2008) among those who changed their preferences after viewing the second positive campaign video.

The ventromedial cortex activated in our study was adjacent to the region recruited in the prior experiments during a binary judgement in opposing contexts, that is, when given and not given formally irrelevant information (i.e., with and without a framing effect) (Deppe et al., 2005a; McClure et al., 2004). Some studies reported that the ventromedial prefrontal cortex was associated with a reward decision under the influence of a specific brand name of goods (Deppe et al., 2005a, b) or with the modulation of emotional rejection to unfair treatments by others at the expense of one's own reward (Koenigs and Tranel, 2007). In contrast, other studies found that it was associated with the exclusive evaluation of one's own rewards when controlling the brand name of goods (McClure et al., 2004) or when not influenced by behavioural feedback, such as punishment or anticipation (Knutson et al., 2001; O'Doherty et al., 2003). An explanation for the opposite results is that different ventromedial prefrontal cortex regions might have been recruited by social stimuli from those associated exclusively with reward consideration, i.e., without interfering with social motives (Harris et al., 2007). However, the region activated in our experiment (16, 61, 6: BA10) is similar to those activated in the two studies with opposing results (Deppe et al., 2005a; McClure et al., 2004).

An alternative explanation to this contradiction is that the ventromedial prefrontal cortex is commonly recruited by social cognition involving affective judgements (Northoff et al., 2006). Evaluating

the relationship between stimuli and oneself (i.e., self-relatedness) is a critical component in the evaluation of rewards and social cognition. The evaluation of one's rewards, based on subjective values, is essentially "affective," but an intrinsic reward consideration may or may not be related to consideration of the relationships between oneself and others. A prior study (de Greck et al., 2008) identified an association between part of the ventromedial prefrontal cortex and self-related considerations; this association was observed in addition to those involving other reward-related regions, such as the striatum, thought to be closely linked to social interactions in reward decisions (Lee, 2008). The apparently contradictory results reported in published experiments imply that the neural correlates of self-revaluation might be dissociated from reward-related regions. Our results support this possibility.

METRIC REPRESENTATION OF PREFERENCE AND BEHAVIOUR

The metric representation of behaviour is prerequisite for closely relating its changes to neural activities. The hypothesized continuity of social preferences with reward consideration enables one to define social decision by externally quantifiable variables, i.e., rewards. We verified the validity of a self-scaled affinity towards others as a metric representation of the brain activities involved in preference changes towards others. As a rating measure of (un)favourable feelings towards presidential candidates, our experiment adopted a self-reported rating method, which was based on a post-task questionnaire. We verified the use of this method by identifying the neural correlates of cognitive control that were tracked by retrospective scaling reported in the post-task questionnaire. This constitutes a promising representation of attitudes and might enable the application of psychometric–neurometric comparisons to a variety of social behaviours. Our research confirms that the self-rated preference towards others used in our experiment, can be used as subjective values. It thereby proposes an alternative measurement to externally quantifiable variables defined by utility functions in neuroeconomics (Kable and Glimcher, 2007).

Our findings imply that the neurometric analysis requires careful interpretation of the neural data analysis to derive behavioural implications. Although the neurometric analysis helps to find the neural correlates of mental states (Kay et al., 2008), the subtle working of the brain precludes an easy extension of this finding to the neural correlates of behaviour.

ACKNOWLEDGEMENTS

This study was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas (#19046001) to J. Kato, I. Kabashima, and K. Kansaku. We thank M. Chun for fruitful discussions and for co-organising an interdisciplinary meeting, "Mind, Brain and Society: Neurocognitive Approaches to the Social Sciences", between neuroscientists and social scientists at Yale University on April 25, 2008, with J. Kato. We are grateful to H. Komiyama and R. Levin who launched the Todai-Yale Initiative that sponsored the symposium and to M. Asashima and G. Joseph who helped our academic exchange fruitful in a variety of ways. We thank T. Kochiyama, T. Shimotomai, D. Salat, K. Sakai, and F. MacDonald for their help. We also thank D. Lee for his continuous encouragement and the two reviewers for improving this manuscript. All trademarks appearing in this article are the property of their respective owners.

REFERENCES

Amodio, D. M., and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277.

Cairns, E., Kenworthy, J., Campbell, A., and Hewstone, M. (2006). The role of in-group identification, religious group membership and intergroup conflict in moderating in-group and out-group affect. *Br. J. Soc. Psychol.* 45, 701–716.

Camerer, C. F. (2008). Behavioural game theory. In The New Palgrave Dictionary of Economics, Vol. 1, 2nd edn, S. N. Durlauf and L. E. Blume, eds (New York, Palgrave Macmillan), pp. 444–448.

Canessa, N., Gorini, A., Cappa, S. F., Piattelli-Palmarini, M., Danna, M., Fazio, F., and Perani, D. (2005). The effect of social content on deductive reasoning: an fMRI study. *Hum. Brain Mapp.* 26, 30–43.

Carpenter, J. (2008). Social preferences. In The New Palgrave Dictionary of Economics, Vol. 7, 2nd edn, S. N. Durlauf and L. E. Blume, eds (New York, Palgrave Macmillan), pp. 651–655.

de Greck, M., Rotte, M., Paus, R., Moritz, D., Thiemann, R., Proesch, U., Bruer, U., Moerth, S., Tempelmann, C., Bogerts, B., and Northoff, G. (2008). Is our self based on reward? Self-relatedness recruits neural activity in the reward system. *Neuroimage* 39, 2066–2075.

Deppe, M., Schwindt, W., Kramer, J., Kugel, H., Plassmann, H., Kenning, P., and Ringelstein, E. B. (2005a). Evidence for a neural correlate of a framing effect: bias-specific activity in the ventromedial prefrontal cortex during credibility judgments. *Brain Res. Bull.* 67, 413–421.

Deppe, M., Schwindt, W., Kugel, H., Plassmann, H., and Kenning, P. (2005b). Nonlinear responses within the medial prefrontal cortex reveal when specific implicit information influences economic decision making. *J. Neuroimaging* 15, 171–182.

Derrfuss, J., Brass, M., and von Cramon, D. Y. (2004). Cognitive control in the posterior frontolateral cortex: evidence from common activations in task coordination, interference control, and working memory. *Neuroimage* 23, 604–612.

Duncan, J., and Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci.* 23, 475–483.

Fehr, E., and Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends Cogn. Sci.* 11, 419–427.

Goel, V. (2007). Anatomy of deductive reasoning. *Trends Cogn. Sci.* 11, 435–441.

Goel, V., and Dolan, R. J. (2004). Differential involvement of left prefrontal cortex in inductive and deductive reasoning. *Cognition* 93, B109–B121.

Harris, L. T., McClure, S. M., van den Bos, W., Cohen, J. D., and Fiske, S. T. (2007). Regions of the MPFC differentially tuned to social and nonsocial affective evaluation. *Cogn. Affect. Behav. Neurosci.* 7, 309–316.

Kable, J. W., and Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* 10, 1625–1633.

Kaplan, J. T., Freedman, J., and Iacoboni, M. (2007). Us versus them: political attitudes and party affiliation influence neural response to faces of presidential candidates. *Neuropsychologia* 45, 55–64.

Katznelson, I., and Weingast, B. R. (2005). Preferences and Situations: Points of Intersection Between Historical and Rational Choice Institutionalism. New York, Russell Sage Foundation.

Kay, K. N., Naselaris, T., Prenger, R. J., and Gallant, J. L. (2008). Identifying natural images from human brain activity. *Nature* 452, 352–355.

Knutson, B., Adams, C. M., Fong, G. W., and Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J. Neurosci.* 21, RC159.

Knutson, K. M., Wood, J. N., Spampinato, M. V., and Grafman, J. (2006). Politics on the brain: an MRI investigation. *Soc. Neurosci.* 1, 25–40.

Koenigs, M., and Tranel, D. (2007). Irrational economic decision-making after ventromedial prefrontal damage: evidence from the ultimatum game. *J. Neurosci.* 27, 951–956.

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, E. S., Rainey, L., Kochunov, P. V., Nickerson, D., Mikiten, S. A., and Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Hum. Brain Mapp.* 10, 120–131.

Lau, R. R., Sigelman, L., and Rovner, I. B. (2007). The effects of negative political campaigns: a meta-analytic reassessment. *J. Politics* 69, 1176–1209.

Lee, D. (2008). Game theory and neural basis of social decision making. *Nat. Neurosci.* 11, 404–409.

Lieberman, M. D. (2007). Social cognitive neuroscience: a review of core processes. *Annu. Rev. Psychol.* 58, 259–289.

Loewenstein, G., Rick, S., and Cohen, J. D. (2008). Neuroeconomics. *Annu. Rev. Psychol.* 59, 647–672.

McClure, S. M., Li, J., Tomlin, D., Cypert, K. S., Montague, L. M., and Montague, P. R. (2004). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron* 44, 379–387.

Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202.

Monti, M. M., Osherson, D. N., Martinez, M. J., and Parsons, L. M. (2007). Functional neuroanatomy of deductive inference: a language-independent distributed network. *Neuroimage* 37, 1005–1016.

Mueller, D. C. (2003). Public Choice III. Cambridge, Cambridge University Press.

Mutz, D. C., Sniderman, P. M., and Brody, R. A. (1996). Political Persuasion and Attitude Change. Ann Arbor, University of Michigan Press.

Northoff, G., Grimm, S., Boeker, H., Schmidt, C., Bermpohl, F., Heinzel, A., Hell, D., and Boesiger, P. (2006). Affective judgment and beneficial decision making: ventromedial prefrontal activity correlates with performance in the Iowa gambling task. *Hum. Brain Mapp.* 27, 572–587.

O'Doherty, J., Critchley, H., Deichmann, R., and Dolan, R. J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *J. Neurosci.* 23, 7931–7939.

Oldfield, R. C. (1971). Assessment and analysis of handedness: Edinburgh inventory. *Neuropsychologia* 9, 97–113.

Ostrom, E., and Walker, J. (2003). Trust and Reciprocity: Interdisciplinary Lessons from Experimental Research. New York, Russell Sage Foundation.

Pelphrey, K. A., Morris, J. P., and McCarthy, G. (2004). Grasping the intentions of others: the perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *J. Cogn. Neurosci.* 16, 1706–1716.

Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., and Banaji, M. R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. *J. Cogn. Neurosci.* 12, 729–738.

Rowe, J. B., Toni, I., Josephs, O., Frackowiak, R. S., and Passingham, R. E. (2000). The prefrontal cortex: response selection or maintenance within working memory? *Science* 288, 1656–1660.

Satterthwaite, T. D., Green, L., Myerson, J., Parker, J., Ramaratnam, M., and Buckner, R. L. (2007). Dissociable but inter-related systems of cognitive control and reward during decision making: Evidence from pupillometry and event-related fMRI. *Neuroimage* 37, 1017–1031.

Talairach, J., and Tournoux, P. (1988). Co-Planar Stereotaxic Atlas of the Human Brain. New York, Thieme.

Volz, K. G., Schubotz, R. I., and von Cramon, D. Y. (2004). Why am I unsure? Internal and external attributions of uncertainty dissociated by fMRI. *Neuroimage* 21, 848–857.

Volz, K. G., Schubotz, R. I., and von Cramon, D. Y. (2005). Variants of uncertainty in decision-making and their neural correlates. *Brain Res. Bull.* 67, 403–412.

Wagner, A. D., Maril, A., Bjork, R. A., and Schacter, D. L. (2001). Prefrontal contributions to executive control: fMRI evidence for functional distinctions within lateral Prefrontal cortex. *Neuroimage* 14, 1337–1347.

Weisberg, H. F., and Miller, A. H. (1979). Evaluation of the Feeling Thermometer: A Report to the National Election Study Board Based on Data from the 1979 Pilot Survey. Ann Arbor, MI, American National Election Studies.

Weissman, D. H., Perkins, A. S., and Woldorff, M. G. (2008). Cognitive control in social situations: a role for the dorsolateral prefrontal cortex. *Neuroimage* 40, 955–962.

Weissman, D. H., Warner, L. M., and Woldorff, M. G. (2004). The neural mechanisms for minimizing cross-modal distraction. *J. Neurosci.* 24, 10941–10949.

Westen, D., Blagov, P. S., Harenski, K., Kilts, C., and Hamann, S. (2006). Neural bases of motivated reasoning: an fMRI study of emotional constraints on partisan political judgment in the 2004 US Presidential election. *J. Cogn. Neurosci.* 18, 1947–1958.

Yamagishi, T., Kikuchi, M., and Kosugi, M. (1999). Trust, gullibility, and social intelligence. *Asian J. Soc. Psychol.* 2, 145–161.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 February 2009; paper pending published: 23 February 2009; accepted: 06 May 2009; published online: 18 May 2009.

Citation: Kato J, Ide H, Kabashima I, Kadota H, Takano K and Kansaku K (2009) Neural correlates of attitude change following positive and negative advertisements. Front. Behav. Neurosci. (2009) ***3****:6. doi:10.3389/neuro.08.006.2009*