Impact of size and delay on neural activity in the rat limbic corticostriatal system

Matthew R. Roesch* and Daniel W. Bryden

Department of Psychology and Program in Neuroscience and Cognitive Science, University of Maryland, College Park, MD, USA

A number of factors influence an animal's economic decisions. Two most commonly studied are the magnitude of and delay to reward. To investigate how these factors are represented in the firing rates of single neurons, we devised a behavioral task that independently manipulated the expected delay to and size of reward. Rats perceived the differently delayed and sized rewards as having different values and were more motivated under short delay and big-reward conditions than under long delay and small reward conditions as measured by percent choice, accuracy, and reaction time. Since the creation of this task, we have recorded from several different brain areas including, orbitofrontal cortex, striatum, amygdala, substantia nigra pars reticulata, and midbrain dopamine neurons. Here, we review and compare those data with a substantial focus on those areas that have been shown to be critical for performance on classic time discounting procedures and provide a potential mechanism by which they might interact when animals are deciding between differently delayed rewards. We found that most brain areas in the cortico-limbic circuit encode both the magnitude and delay to reward delivery in one form or another, but only a few encode them together at the single neuron level.

Keywords: discounting, value, dopamine, orbitofrontal, striatum, amygdala, substantia nigra

INTRODUCTION

Animals prefer an immediate reward over a delayed reward even when the delayed reward is more economically valuable in the long run. In the lab, the neural mechanisms underlying this aspect of decision-making are often studied in tasks that ask animals or humans to choose between a small reward delivered immediately and a large reward delivered after some delay (Herrnstein, 1961; Ainslie, 1974; Thaler, 1981; Kahneman and Tversky, 1984; Rodriguez and Logue, 1988; Lowenstein, 1992; Evenden and Ryan, 1996; Richards et al., 1997; Ho et al., 1999; Cardinal et al., 2001; Mobini et al., 2002; Winstanley et al., 2004b; Kalenscher et al., 2005; Kalenscher and Pennartz, 2008; Ballard and Knutson, 2009; Figner et al., 2010). As the delay to the large reward becomes longer, subjects tend to start discounting the value of the large reward, biasing their choice behavior toward the small, immediate reward (temporal discounting). This choice behavior is considered to be impulsive because over the course of many trials it would be more economical to wait for the larger reward. Impulsive choice is exacerbated in several disorders such as drug addiction, attention-deficit/hyperactivity disorder, and schizophrenia, altering the breakpoint at which subjects abandon the large-delayed reward for the more immediate reward (Ernst et al., 1998; Jentsch and Taylor, 1999; Monterosso et al., 2001; Bechara et al., 2002; Coffey et al., 2003; Heerey et al., 2007; Roesch et al., 2007c; Dalley et al., 2008). Although considerable attention has been paid to the neuroanatomical and pharmacological basis of temporally discounted reward and impulsivity, few have examined the neural correlates involved. Specifically, few have asked how delays impact neural encoding in brain areas known to be involved in reinforcement

learning and decision-making, and how that encoding might relate to less abstract manipulations of value such as magnitude. To address this issue we developed an inter-temporal choice task suitable for behavioral recording studies in rats (Roesch et al., 2006, 2007a,b, 2009, 2010b; Takahashi et al., 2009; Calu et al., 2010; Stalnaker et al., 2010; Bryden et al., 2011).

In this task, rats were trained to nosepoke into a central odor port. After 0.5 s, one of three odors was presented. One odor signaled for the rat to go left (forced-choice), another signaled go right (forced-choice), and the third odor signaled that the rat was free to choose either the left or right well (free-choice) to receive liquid sucrose reward. The two wells were located below the odor port as illustrated in **Figure 1B**. After responding to the well, rats had to wait 0.5 or 1–7 s to receive reward, depending on trial type (**Figure 1A**). The task was designed to allow for equal samples of left and rightward responses (forced-choice) while at the same time having a direct measure of the animal's preference (free-choice). In addition, use of free- and forced-choice trials has allowed us to determine whether the brain processes free-choice differently than forced instrumental responding and whether or not observed neural correlates reflect sensory or motor processing.

At the start of each session, we shifted the rats' response bias to the left or to the right by increasing the delay preceding reward delivery in one of the two fluid wells (1–7 s). During delay blocks, each well yielded one bolus of 10% sucrose solution. After ∼60–80 trials, the response direction associated with the delayed well switched unexpectedly. Thus, the response direction that was associated with a short delay became long, whereas the response direction associated with the long delay in the first block of trials
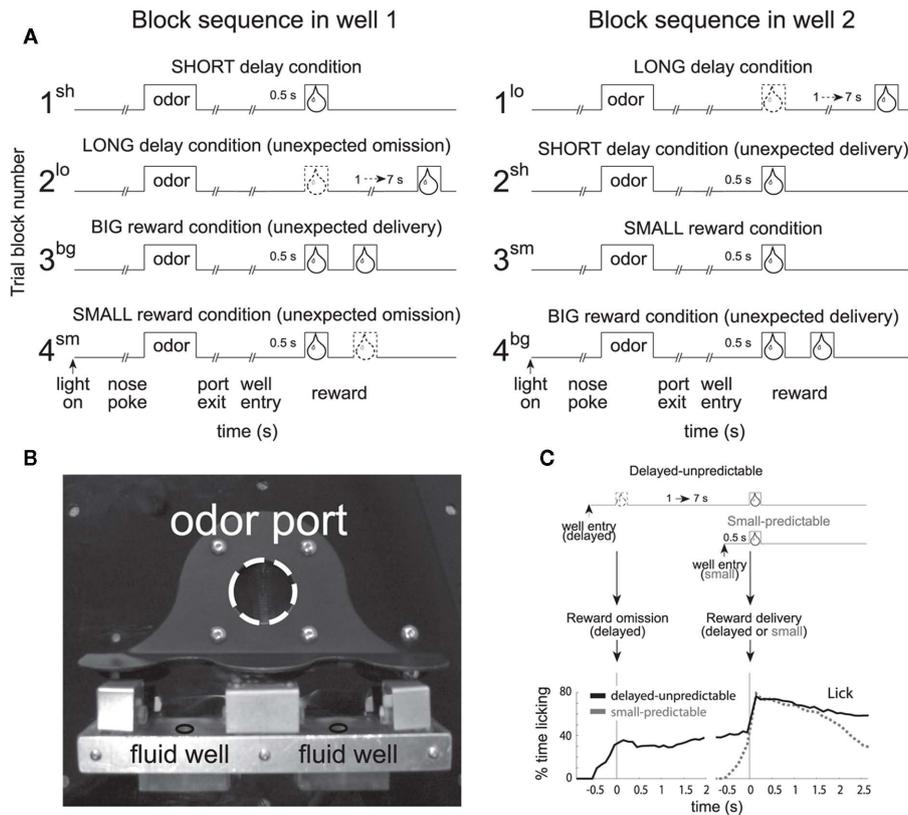
FIGURE 1 | Size and Delay Behavioral Choice Task. (A) Figure shows sequence of events in each trial in 4 blocks in which we manipulated the time to reward or the size of reward. Trials were signaled by illumination of the panel lights inside the box. When these lights were on, nosepoke into the odor port resulted in delivery of the odor cue to a small hemicylinder located behind this opening. One of three different odors was delivered to the port on each trial, in a pseudorandom order. At odor offset, the rat had 3 s to make a response at one of the two fluid wells located below the port. One odor instructed the rat to go to the left to get reward, a second odor instructed the rat to go to the right to get reward, and a third odor indicated that the rat could obtain reward at either well. At the start of each recording session one well was randomly designated as short (a 0.5 s delay before reward) and the other long (a 1 to 7 s delay before reward) (block 1). In the second block of trials these contingencies were switched (block 2). In blocks 3 and 4 delays were held constant (0.5 s) and reward size was manipulated. sh = short; bg = big; lo = long; sm = small; **(B)** Picture of apparatus. **(C)** Percent licking behavior averaged over all recording sessions during trials when a small reward was delayed versus when a small reward was delivered after 0.5 s. Licking is aligned to well entry (left) and reward delivery (right). Adapted from (Roesch et al., 2006, Roesch et al., 2007b, Takahashi et al., 2009).

became short. During delay blocks, the intertrial intervals were normalized so that the length of short and long delay trials were equal, thus there was no overall benefit to choosing the short delay, but as we will describe, rats did so regardless.

These contingencies continued for ∼60–80 trials at which time both delays were set to 0.5 s and the well that was associated with the long delay, now produced a large reward (two to three boli). These contingencies were again switched in the fourth block of trials. Trial block switches were not cued, thus animals had to detect changes in reward contingencies and update behavior from block to block.

It is important to emphasize that reward size and delay were varied independently, unlike common delay discounting tasks. Whereas other studies have investigated the neuronal coding of temporally discounted reward in paradigms that have manipulated size and delay simultaneously, our task allows us to dissociate correlates related to size and delay to better understand how each manipulation of value is coded independently from the other. As we will show below, rats prefer or value short over long delays to

reward and large over small reward as indicated by choice performance. We felt it was necessary to dissociate size correlates from delay correlates because certain disorders and brain manipulations have been shown to impair size and delay processing independently, sometimes in an opposing manner (Roesch et al., 2007c). Although we do not manipulate the size of the reward along with the length of the delay preceding reward delivery in the traditional sense, any effects on choice behavior and neural firing must be dependent on the delay and reflect how time spent waiting for a reward reduces the value of reward. The depreciation of the reward value due to delay has been referred to as the temporally discounted value of the reward (Kalenscher and Pennartz, 2008).

In each of the studies that we will describe below, rats were significantly more accurate and faster on high value reward trials (large reward and short delay) as compared to low value reward trials (small reward and long delay) on forced-choice trials (Roesch et al., 2006, 2007a,b, 2009, 2010b; Takahashi et al., 2009; Calu et al., 2010; Stalnaker et al., 2010; Bryden et al., 2011). On free-choice trials, rats chose high over low value and switched their preference

rapidly after block changes. Thus, rats discounted delayed rewards, choosing it less often and working less hard to achieve it. Preference of immediate reward over delayed reward was not significantly different than preference of the large over small reward.

In addition, behavioral measures have illustrated that delayed rewards were less predictable than more immediate rewards in this task (Takahashi et al., 2009). Even after learning, the rats could not predict the delayed reward with great precision. Licking increased rapidly prior to the small, more immediate reward and showed no change prior to delivery of the delayed small reward (**Figure 1C**; Takahashi et al., 2009). Instead, rats' licking behavior increased around 0.5 s after well entry on delayed trials (**Figure 1C**) corresponding to the time when delivery of immediate reward would have happened in the preceding block of trials (**Figure 1C**). Thus, rats anticipated delivery of immediate reward, even on long delay trials and, although they knew that the delayed reward would eventually arrive, they could not predict exactly when. Similar findings have been described in primates (Kobayashi and Schultz, 2008).

In this article, we review neural correlates related to performance of this task from several brain areas, with a stronger focus on those areas that have been shown to disrupt behavior on standard delay discounting tasks after lesions, inactivation, or other pharmacological manipulations (Cardinal et al., 2004; Floresco et al., 2008).

ORBITOFRONTAL CORTEX

Impulsive choice in humans has long been attributed to damage of orbitofrontal cortex (OFC), but the role that OFC plays in inter-temporal choices remains unclear. OFC lesions decrease and increase discounting functions depending on experimental design and lesion location (Mobini et al., 2002; Winstanley et al., 2004b; Rudebeck et al., 2006; Winstanley, 2007; Churchwell et al., 2009; Sellitto et al., 2010; Zeeb et al., 2010; Mar et al., 2011). From these data it has been clear that OFC is involved in inter-temporal choice suggesting that it must carry information related to the length of delay preceding the delivery of reward.

To investigate how delay and size were encoded in OFC, we recorded from single neurons while rats performed the task described above (Roesch et al., 2006). Consistent with previous work, lateral OFC neurons fired in anticipation of delayed reward. As illustrated in **Figure 2A**, activity of many single neurons continuously fired until the delayed reward was delivered, resulting in higher levels of activity for rewards that were delayed (**Figure 2A**; bottom; gray).

Surprisingly, the majority of OFC neurons in our study did not show this pattern of activity (Roesch et al., 2006). Most OFC neurons did not maintain firing across the delay as illustrated by the single cell example in **Figure 2B**. Under short delay conditions, this neuron fired in anticipation of and during the delivery of immediate reward (top; black). When the reward was delayed (gray), activity declined until the delayed reward was delivered, and thus, did not bridge the gap between the response and reward as in the previous example (**Figure 2A**). Interestingly, activity seemed to reflect the expectation of reward by continuing to fire when the reward would have been delivered on previous trials (i.e., 0.5 s after the response). This old expectancy signal slowly dissipated with learning (**Figure 2B**).
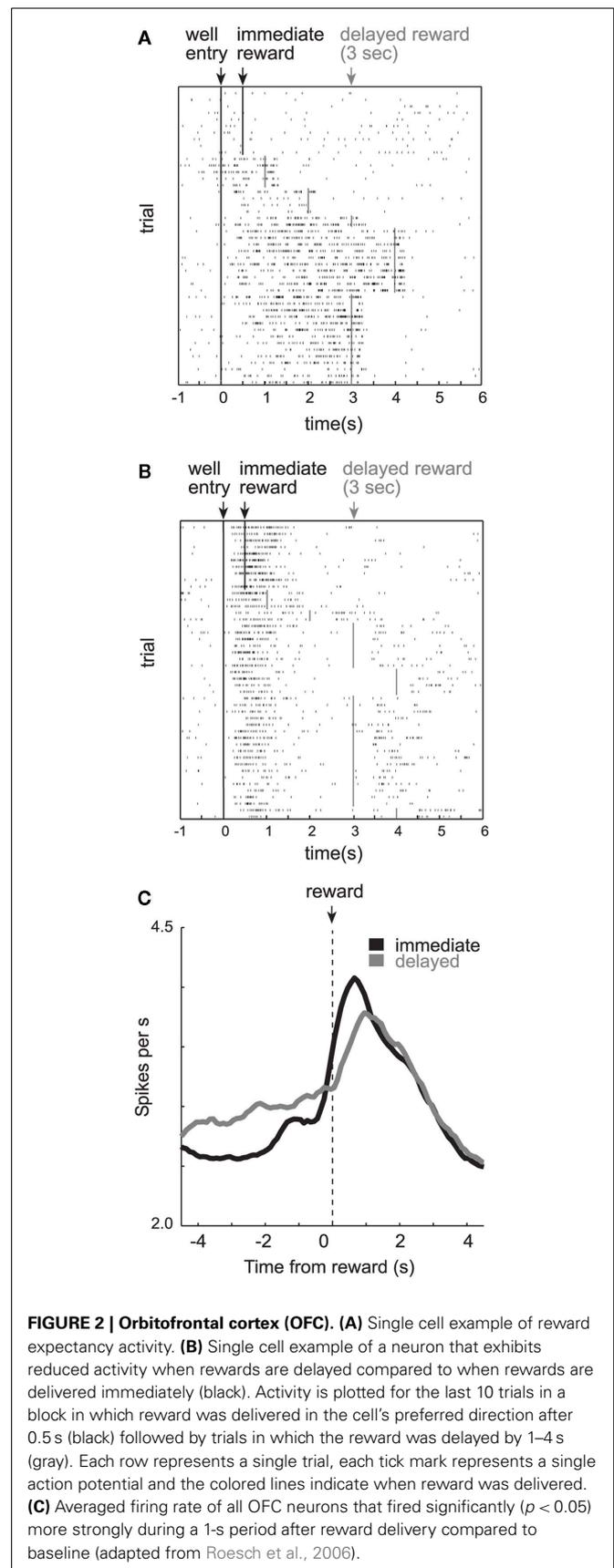

FIGURE 2 | Orbitofrontal cortex (OFC). **(A)** Single cell example of reward expectancy activity. **(B)** Single cell example of a neuron that exhibits reduced activity when rewards are delayed compared to when rewards are delivered immediately (black). Activity is plotted for the last 10 trials in a block in which reward was delivered in the cell's preferred direction after 0.5 s (black) followed by trials in which the reward was delayed by 1–4 s (gray). Each row represents a single trial, each tick mark represents a single action potential and the colored lines indicate when reward was delivered. **(C)** Averaged firing rate of all OFC neurons that fired significantly ($p < 0.05$) more strongly during a 1-s period after reward delivery compared to baseline (adapted from Roesch et al., 2006).

Thus, it appears that although many OFC neurons maintained representations of the reward across the delay, most did not. Overall activity across the population of reward-responsive neurons was stronger during delivery of immediate reward as compared to delayed reward (**Figure 2C**). These changes in firing likely had a profound impact on inter-temporal choice. Indeed, firing of OFC neurons was correlated with the tendency for the rat to choose the short delay on future free-choice trials (Roesch et al., 2006, 2007a).

We suspect that these two types of signals play very different roles during performance of standard delay discounting paradigms (Roesch et al., 2006, 2007a). Reward expectancy signals that maintain a representation of the delayed delivery of reward (**Figure 2A**) might be critical for facilitating the formation of associative representations in other brain regions during learning. For example, it has been shown that input from OFC is important for rapid changes in cue–outcome encoding in basolateral amygdala (Saddoris et al., 2005) and prediction error signaling in DA neurons in ventral tegmental area (VTA). Loss of cue–outcome encoding in downstream areas after OFC lesions may be due to the loss of expectancy signals generated in OFC (Schoenbaum and Roesch, 2005). If the purpose of expectancy signals in OFC is to maintain a representation of the reward when it is delayed so that downstream areas can develop cue or response–outcome associations, then animals with OFC lesions would be less likely to choose those cues or responses when they result in the delayed reward. This interpretation is consistent with reports that lesions of OFC can cause more impulsive responding (Rudebeck et al., 2006).

The majority of OFC neurons fired more strongly for immediate reward (**Figure 2B**). These neurons likely represent when an immediate reward is or is about to be delivered. When the reward is delayed, this expectation of immediate reward is violated and a negative prediction error is generated in downstream areas. Negative prediction error signals would subsequently weaken associations between cues and responses that predict the now delayed reward. These changes would drive behavior away from responses that predict the delayed reward. Elimination of this signal could make animals less likely to abandon the delayed reward as has been shown in previous studies (Winstanley et al., 2004b).

To add to this complexity, a recent study suggests that different regions of OFC might serve opposing functions related to inter-temporal choice (Mar et al., 2011). In this study, Mar and colleagues showed that lesions of medial OFC make rats discount slower, encouraging responding to the larger delayed reward, whereas lateral OFC lesions make rats discount faster, decreasing preference for the larger delayed reward. How does this relate to our data? It suggests that neurons that bridge the gap during the delay preceding reward delivery might be more prominent in lateral OFC. This hypothesis is consistent with human imaging studies showing a positive correlation between OFC activation and preference for delayed reward (McClure et al., 2004, 2007; Hariri et al., 2006; Boettiger et al., 2007; Mar et al., 2011). These data also suggest that neurons that exhibit reduced activity for delayed reward, firing more strongly for immediate reward, might be more prominent in medial OFC. This hypothesis is consistent with human imaging studies showing that activation of medial OFC is positively correlated with preference of more immediate reward (McClure et al., 2004, 2007; Hariri et al., 2006; Mar et al., 2011). Future studies will have to examine whether this theory is true and/or if other signals might be involved in generating the opposing symptoms observed after medial and lateral OFC lesions.
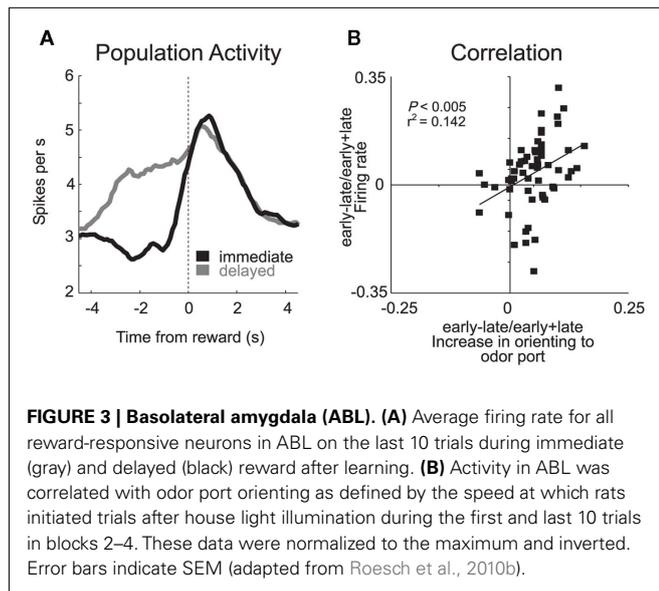
Notably, a number of other prefrontal cortical areas are thought to be involved in processing delayed reward. Most of this work has come from humans and in studies examining neural activity in monkeys. For example, Kim et al. (2008) found that single neurons in monkey prefrontal cortex (PFC) were modulated by both expected size of and delay to reward in a task in which monkeys chose between targets that predicted both magnitude and delay. Human studies have backed these findings and have further suggested that PFC, unlike OFC, might be more critical in evaluating rewards that are more extensively delayed (e.g., months to years; McClure et al., 2004, 2007; Tanaka et al., 2004; Kable and Glimcher, 2007; Ballard and Knutson, 2009; Figner et al., 2010).

BASOLATERAL AMYGDALA

Much of the evidence we have described for the general role of OFC in anticipating future events and consequences can also be found in studies of amygdalar function, in particular, the ABL (Jones and Mishkin, 1972; Kesner and Williams, 1995; Hatfield et al., 1996; Malkova et al., 1997; Bechara et al., 1999; Parkinson et al., 2001; Cousens and Otto, 2003; Winstanley et al., 2004d). This is perhaps not surprising given the strong reciprocal connections between OFC and ABL and the role that ABL is proposed to play in associative learning. ABL also appears to play a critical role during inter-temporal choice. Rats with ABL lesions are more impulsive when rewards are delayed, abandoning the delayed reward more quickly than controls (Winstanley et al., 2004b; Cardinal, 2006; Churchwell et al., 2009; Ghods-Sharifi et al., 2009).

As in many studies, activity patterns observed in ABL during performance of our task were similar to those observed in OFC; neurons represented predicted outcomes at the time of cue presentation and in anticipation of reward (Roesch et al., 2010b). The two areas differed in that signals related to anticipated reward and delivery did not appear to be as reduced in ABL as they were in OFC when rewards were delayed. This is evident by comparing population histograms from both areas (OFC: **Figure 2C** versus ABL: **Figure 3A**). The counts of neurons that fired significantly more strongly for immediate reward did not outnumber the number of neurons that fired more strongly for delayed reward in ABL as they did in OFC.

Another difference between ABL and OFC was that neurons in ABL also fired more strongly when reward was delivered unexpectedly. For example, many ABL neurons fired strongly when the big-reward was delivered at the start of blocks 3 and 4 (**Figure 1A**). Although the mainstream view holds that amygdala is important for acquiring and storing associative information (LeDoux, 2000; Murray, 2007), these data and others like it have recently suggested that amygdala may also support other functions related to associative learning such as detecting the need for increased attention when reward expectations are violated (Gallagher et al., 1990; Holland and Gallagher, 1993b, 1999; Breiter et al., 2001; Yacubian et al., 2006; Belova et al., 2007; Tye et al., 2010). Consistent with this hypothesis, we have shown that activity during unexpected reward delivery and omission was

FIGURE 3 | Basolateral amygdala (ABL). (A) Average firing rate for all reward-responsive neurons in ABL on the last 10 trials during immediate (gray) and delayed (black) reward after learning. **(B)** Activity in ABL was correlated with odor port orienting as defined by the speed at which rats initiated trials after house light illumination during the first and last 10 trials in blocks 2–4. These data were normalized to the maximum and inverted. Error bars indicate SEM (adapted from Roesch et al., 2010b).

correlated with changes in attention that occur at the start of trial blocks (**Figure 3B**) and that inactivation of ABL makes rats less likely to detect changes in reward contingencies (Roesch et al., 2010b).

Unfortunately, it is still unclear what sustained activity during the delay represents in ABL. Sustained activity in ABL might reflect unexpected omission of reward, signaling to the rat to attend more thoroughly to that location so that new learning might occur. It might also serve to help maintain learned associations and/or to learn new associations when delays are introduced between responses and reward delivery. Consistent with this hypothesis, ABL lesions have been shown to reduce the selectivity of neural firing in OFC and ventral striatum (VS; Schoenbaum et al., 2003; Ambroggi et al., 2008). If ABL's role is to help maintain expectancies or attention across the gap between responding and delivery of delayed reward, then loss of this signal would increase impulsive choice as has been shown by other labs (Winstanley et al., 2004b).

Finally, it is worth noting that other parts of the amygdala might be critical for inter-temporal decision-making. Most prominent is the central nucleus of amygdala (CeA), which is critical for changes in attention or variations in event processing that occur during learning when rewards downshift from high to low value (Holland and Gallagher, 1993a,c, 2006; Holland and Kenmuir, 2005; Bucci and Macleod, 2007). We have recently shown that downshifts in value, including when rewards are unexpectedly delayed, increase firing in CeA during learning (Calu et al., 2010). Changes in firing were correlated with behavioral measures of attention observed when reward contingencies were violated, which were lost after CeA inactivation (Calu et al., 2010). Surprisingly, inactivation of CeA did not impact temporal choice in our task (Calu et al., 2010). This might reflect control of behavior via detection of unexpected reward delivery which happens concurrently with unexpected reward omission during each block switch. To the best of our knowledge, it is unknown how CeA lesions would impact performance on the standard small-immediate versus large-delayed

reward temporal discounting task, but we suspect that rats would be less impulsive.
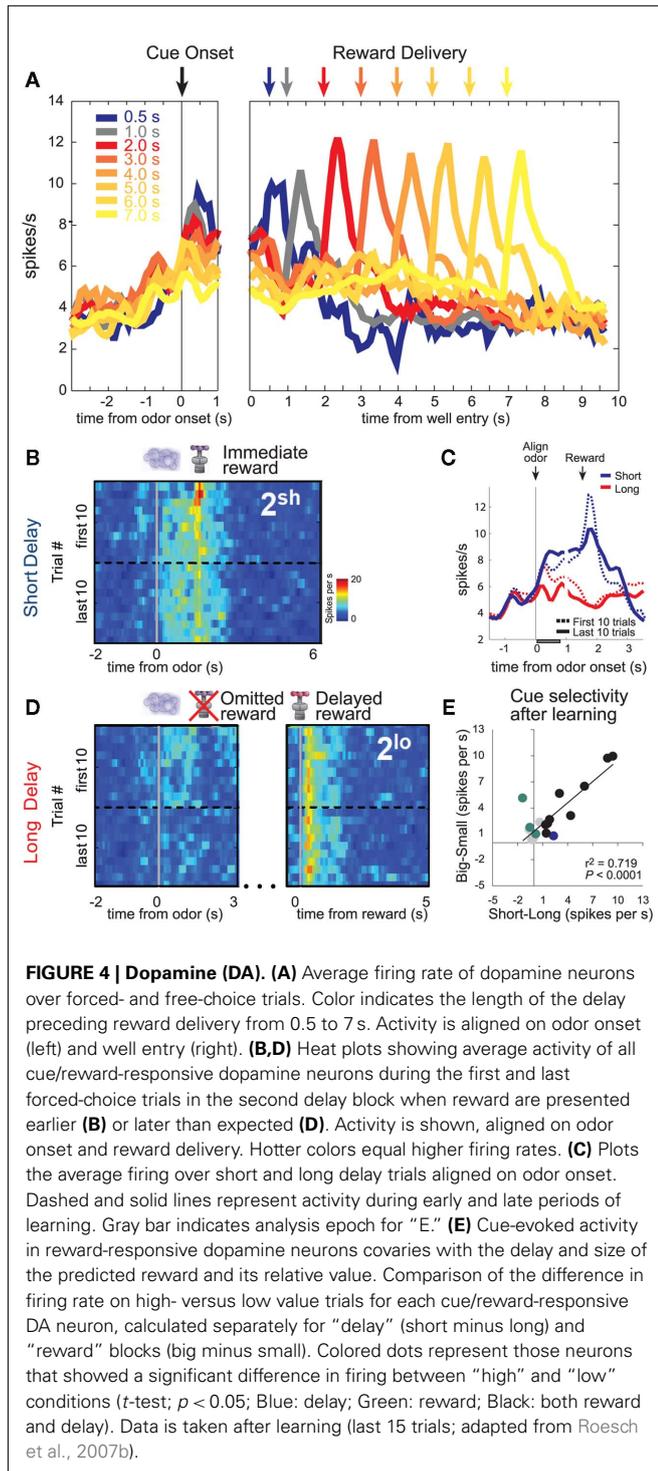
DOPAMINE NEURONS IN VENTRAL TEGMENTAL AREA

Manipulations of DA can either increase or decrease how much animals discount delayed reward (Cardinal et al., 2000, 2004; Wade et al., 2000; Kheramin et al., 2004; Roesch et al., 2007c); however, few studies have examined how DA neurons respond when rewards are unexpectedly delayed or delivered after long delay (Fiorillo et al., 2008; Kobayashi and Schultz, 2008; Schultz, 2010). As in previous work, unexpected manipulation of reward size in our task impacted firing of DA neurons in VTA. Activity was high or low depending on whether reward was unexpectedly larger (positive prediction error) or smaller (negative prediction error), respectively, and activity was high or low depending on whether the odor predicted large or small reward, respectively. Thus, consistent with previous work, the activity of DA neurons appeared to signal errors in reward prediction during the presentation of unconditioned and conditioned stimuli (Mirenowicz and Schultz, 1994; Montague et al., 1996; Hollerman and Schultz, 1998a,b; Waelti et al., 2001; Fiorillo et al., 2003; Tobler et al., 2003; Nakahara et al., 2004; Bayer and Glimcher, 2005; Pan et al., 2005; Morris et al., 2006).

DA neurons also signaled errors in reward prediction when rewards were delayed (Roesch et al., 2007b). Delivery of an unexpected immediate reward elicited a strong DA response (**Figure 4B**; immediate reward; red blotch; first 10 trials), which was subsequently replaced by firing to cues that predicted the short delay after learning (**Figure 4B**; last 10 trials). That is, activity was stronger at the end of the block (dashed blue line) than during the first several trials of that same block (solid blue line) just after odor presentation (**Figure 4C**). Overall, population activity was the strongest during cues that predicted the immediate reward (**Figures 4A–C**; Roesch et al., 2007b). Moreover, neurons that tended to fire more strongly for immediate reward also fired more strongly for cues that predicted large reward (**Figure 4E**).

When rewards were unexpectedly delayed, DA neurons were inhibited at the time when the reward should have arrived on short delay trials (**Figure 4D**; omitted reward; first 10 trials). Again, this negative prediction error signal transferred to cues that predicted the delayed reward after learning (**Figure 4D**; last 10 trials). That is, cue-related activity was still strong at the start of the block before the rat realized that the cue no longer signaled short delay. During odor sampling activity was the weakest when that cue signaled the longest delay (**Figure 4A**; 7 s; cue onset).

Finally, consistent with delayed rewards being unpredictable (**Figure 1C**), rewards delivered after long delay elicited strong firing (**Figure 4A**; 7 s; cue onset and **Figure 4D**; delayed reward). Activity after 2 s did not increase with each successive delay increase. This is likely due to rats updating their expectations as the delay period grew second by second. All of these findings are consistent with the notion that activity in midbrain DA signals errors in reward prediction.

Importantly, our results are consistent with work in humans and primates. Human fMRI studies demonstrate that VTA's efferents are active when participants are making decisions related to more immediate reward (McClure et al., 2004, 2007; Tanaka

FIGURE 4 | Dopamine (DA). (A) Average firing rate of dopamine neurons over forced- and free-choice trials. Color indicates the length of the delay preceding reward delivery from 0.5 to 7 s. Activity is aligned on odor onset (left) and well entry (right). **(B,D)** Heat plots showing average activity of all cue/reward-responsive dopamine neurons during the first and last forced-choice trials in the second delay block when reward are presented earlier **(B)** or later than expected **(D)**. Activity is shown, aligned on odor onset and reward delivery. Hotter colors equal higher firing rates. **(C)** Plots the average firing over short and long delay trials aligned on odor onset. Dashed and solid lines represent activity during early and late periods of learning. Gray bar indicates analysis epoch for "E." **(E)** Cue-evoked activity in reward-responsive dopamine neurons covaries with the delay and size of the predicted reward and its relative value. Comparison of the difference in firing rate on high- versus low value trials for each cue/reward-responsive DA neuron, calculated separately for "delay" (short minus long) and "reward" (big minus small) blocks. Colored dots represent those neurons that showed a significant difference in firing between "high" and "low" conditions (t-test; p < 0.05; Blue: delay; Green: reward; Black: both reward and delay). Data is taken after learning (last 15 trials; adapted from Roesch et al., 2007b).

et al., 2004; Kable and Glimcher, 2007; Ballard and Knutson, 2009). Direct recordings from primate DA neurons during performance of a simple pavlovian task are also consistent with our results (Fiorillo et al., 2008; Kobayashi and Schultz, 2008). As in our study, activity during delivery of delayed reward was positively correlated with the delay preceding it reflecting the uncertainty or unpredictability of the delayed reward. This might reflect the

possibility that longer delays are harder to time (Church and Gibbon, 1982; Kobayashi and Schultz, 2008). Consistent with this interpretation, monkeys could not accurately predict the delivery of the delayed reward as measured by anticipatory licking (Kobayashi and Schultz, 2008). Also consistent with our work, activity during sampling of cues that predicted reward was discounted by the expected delay. Specifically, the activity of DA neurons resembled the hyperbolic function typical of animal temporal discounting studies, reflecting stronger discounting of delayed reward when delays were relatively short.

Thus, across species, it is clear that signals related to prediction errors are modulated by cues that predict delayed reward. Such modulation must act on downstream neurons in cortex and basal ganglia to promote and suppress behavior during inter-temporal choice. Prominent in the current literature is the idea that DA transmission ultimately impacts behavioral output by influencing basal ganglia output structures such as SNr via modulation of D1 and D2 type receptors in dorsal striatum (DS; Bromberg-Martin et al., 2010; Hong and Hikosaka, 2011). Indeed, we and others have recently shown that DS and SNr neurons incorporate anticipated delay into their response selective firing during and prior to the decision to move (Stalnaker et al., 2010; Bryden et al., 2011; Cai et al., 2011).

We propose that bursting of DA neurons to rewards that are delivered earlier than expected and the cues that come to predict them would activate the D1 mediated direct pathway, directing behavior toward the more immediate reward (Bromberg-Martin et al., 2010). Low levels of dopamine, as observed when rewards are unexpectedly delayed would activate the D2 mediated indirect pathway so that movement toward the well that elicited the delayed reward would be suppressed (Frank, 2005; Bromberg-Martin et al., 2010). Consistent with this hypothesis, it has been shown that high and low DA receptor activation promotes potentiation of the direct and indirect pathway, respectively (Shen et al., 2008), and that striatal D1 receptor blockade selectively impairs movements to rewarded targets, whereas D2 receptor blockade selectively suppresses movements to non-rewarded locations (Nakamura and Hikosaka, 2006).

VENTRAL STRIATUM

Post-training lesions of VS, in particular nucleus accumbens core, induces impulsive choice of small-immediate reward over large-delayed reward (Cousins et al., 1996; Cardinal et al., 2001, 2004; Bezzina et al., 2007; Floresco et al., 2008; Kalenscher and Pennartz, 2008). Although there are several theories about the function of VS, one prominent theory suggests that VS serves as a limbic-motor interface, integrating value information with motor output (Mogenson et al., 1980). Consistent with this notion, several labs have shown that VS incorporates expected value information into its neural firing (Bowman et al., 1996; Hassani et al., 2001; Carelli, 2002; Cromwell and Schultz, 2003; Setlow et al., 2003; Janak et al., 2004; Tanaka et al., 2004; Nicola, 2007; Khamassi et al., 2008; Ito and Doya, 2009; van der Meer and Redish, 2009). Until recently, it was unknown whether VS incorporated expected delay information into this value calculation, possibly serving as a potential source by which representations of delayed reward might impact inter-temporal choice.

We have recently shown that single neurons in VS signal the value of the chosen action during performance of our choice task (Roesch et al., 2009). The majority of cue-responsive neurons in VS fired significantly more strongly when rats anticipated high value reward in one of the two movement directions. This is illustrated in **Figures 5A–D**, which plots the average firing rate of all cue-responsive neurons in VS for responses made in each cell's preferred and non-preferred movement fields. Activity was stronger prior to a response in the cell's preferred direction (left column)

FIGURE 5 | Ventral striatum (VS). Population activity of odor-responsive neurons reflected motivational value and response direction on forced-choice trials. **(A–D)** Curves representing normalized population firing rate during performance of forced-choice trials for the odor-responsive neurons as a function of time under the eight task conditions (high value = black; low value = gray). Data are aligned on odor port exit. Preferred and non-preferred directions are represented in left and right columns, respectively. For each neuron, the direction that yielded the maximal response was designated as preferred. Correlations in the preferred **(E)** and non-preferred **(F)** direction between value indices (short − long/short + long and big − small/big + small) computed for firing rate (during odor sampling) and reaction time (speed at which rats exited the odor port after sampling the odor; adapted from Roesch et al., 2009).

when the expected outcome was either a short delay (**Figure 5A**; black) or a large reward (**Figure 5C**; black) compared to a long delay (**Figure 5A**; gray) or a small reward (**Figure 5C**; gray), respectively. This activity most likely reflects common changes in motivation because neural firing during this period was correlated with the motivational level of the rat, which was high under short delay and large reward conditions (**Figure 5E,F**; Roesch et al., 2009). This result is consistent with previous work showing that activity in VS is modulated during inter-temporal choice for immediate rewards (McClure et al., 2004, 2007; Kable and Glimcher, 2007; Ballard and Knutson, 2009) suggesting that VS is involved in decisions regarding discounted reward (but see Day et al., 2011). We suspect that increased activation of neurons that signal movement during short delay trials might cause animals to choose the more immediate reward over the delayed reward through some sort of winner take all mechanism (Pennartz et al., 1994; Redgrave et al., 1999; Nicola, 2007; Taha et al., 2007).

Others suggest that temporally discounted value signals in VS have less to do with the actual choice – which appears to be more reliably encoded in DS – and more to do with encoding the sum of the temporally discounted values of the available options, that is, the overall goodness of the situation. Unlike our task, monkeys, on each trial, were presented with two options simultaneously. Each option varied in magnitude and delay, and the location of the better reward varied randomly. Color and number of cues signaled size and delay, respectively. These contingencies did not change over the course of the experiment.

Not only was activity in VS modulated by the value of the delayed reward in this study, neurons in VS were more likely to encode the sum of the temporally discounted value of the two targets than the differences between them or the choice that the monkey was about to make (Cai et al., 2011). Our results are similar to these in that activity in VS was modulated by size and delay, however we clearly show that activity in VS signaled the value and the direction of the chosen option. This difference likely reflects differences in the task design. In our task, rats form response biases to one direction over the other during the course of the block and rats constantly had to modify their behavior when contingencies changed, thus response–outcome contingencies were very important in our task. Further, we could not access whether or not activity in VS represented the overall value of the two options because the overall value of the reward did not change from block to block. This was an important and interesting feature of the monkey task and it is highly likely that monkeys paid close attention to the overall value associated with each trial before deciding which option to ultimately choose.

Several studies, including ours, have also shown that VS neurons fire in anticipation of the reward (Roesch et al., 2009). This is apparent in **Figure 5A**, which illustrates that activity was higher after the response in the cell's preferred direction on long delay (gray) compared to short delay trials (black). Interestingly, the difference in firing between short and long delay trials after the behavioral response was also correlated with reaction time, however the direction of this correlation was the opposite of that prior to the movement. That is, slower responding on long delay trials, after the choice, resulted in stronger firing rates during the delay preceding reward delivery. If activity in VS during decision-making

reflects motivation, as we have suggested, then activity during this period may reflect the exertion of increased will to remain in the well to receive reward. As described above for OFC and ABL, this expectancy signal might be critical for maintaining responding when rewards become delayed. Loss of this signal would reduce the rat's capacity to maintain motivation during reward delays as described in other contexts (Cousins et al., 1996; Cardinal et al., 2001, 2004; Bezzina et al., 2007; Floresco et al., 2008; Kalenscher and Pennartz, 2008).

Our data suggest two conflicting roles for VS in delay discounting. We speculate that different training procedures might change the relative contributions of these two functions. For example, if animals were highly trained to reverse behaviors based on discounted reward, as in the recording setting used here, they might be less reliant on VS to maintain the value of the discounted reward. In this situation, the primary effect of VS manipulations might be to reduce impulsive choice of the more immediate reward. On the other hand, maintaining reward information across the delay might be more critical early on during learning, when rats are learning contingencies between responses and their outcomes. Increasing delays between the instrumental response and reinforcer impairs learning in normal animals and is exacerbated after VS lesions (Cardinal et al., 2004).

Besides directly driving behavior, as proposed above, other theories suggest that VS might also be involved in providing expectancy information to downstream areas as part of the "Critic" in the actor–critic model (Joel et al., 2002; O'Doherty et al., 2004). In this model the Critic stores and learns values of states which in turn are used to compute prediction errors necessary for learning and adaptive behavior. Neural instantiations of this model suggests that it is VS that signals the predicted value of the upcoming decision, which in turn impacts prediction error encoding by dopamine neurons. Subsequently, DA prediction errors modify behavior via connections with the DS (Actor) and update predicted value signals in VS. Thus, signaling of immediate and delayed reward by VS would have a profound impact on reinforcement learning in this circuit as we will discuss below.

INTEGRATION OF SIZE AND DELAY ENCODING

Do brain areas integrate size and delay information, providing a context-free representation of value (Montague and Berns, 2002; Kringelbach, 2005; Padoa-Schioppa, 2011)? If this hypothesis is so, then neural activity that encodes the delay to reward should also be influenced by changes in reward magnitude, either at a single-unit or population level. We found that when delay and reward size were manipulated across different blocks of trials, OFC, ABL, and DS maintained dissociable representations of the value of differently delayed and sized rewards. Even in VS, where the population neurons fired more strongly to short delay and large reward conditions and activity was correlated with motivational strength, there was only a slight insignificant tendency for single neurons to represent both size and delay components. Although many neurons did encode reward size and delay length at the single cell level, many neurons encoded one but not the other. This apparent trend toward common encoding likely reflects the integration of value into motor signals at the level of VS which is further downstream than areas such as OFC and ABL.
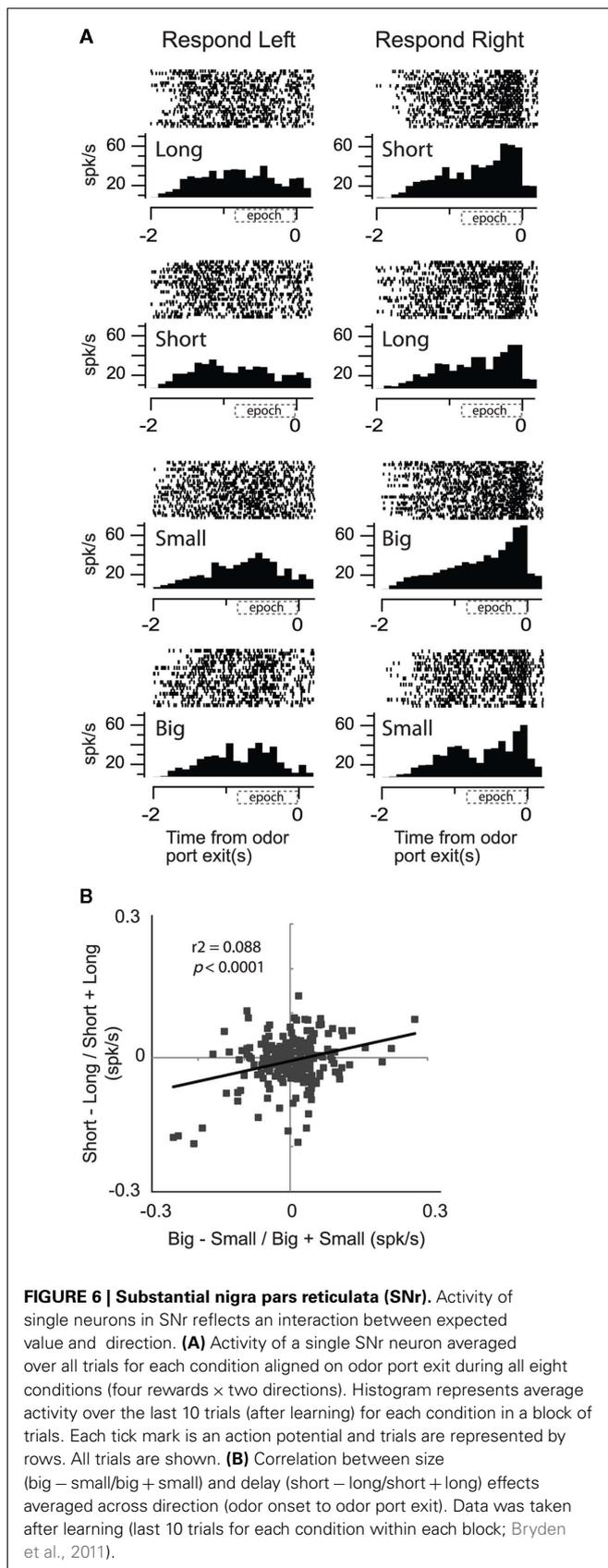
Consistent with this hypothesis, when we recorded from neurons more closely tied to the output of the basal ganglia, we found that activity in SNr showed a significant positive correlation between reward size and delay (Bryden et al., 2011). This is illustrated in the single cell example in **Figure 6A**. Activity was higher for short delay and large reward conditions for movements made into the right well. Unlike OFC, ABL, and VS, those SNr neurons that fired more strongly for cues that predicted short delay (over long delay) significantly tended to fire more strongly for cues that predicted the large reward (over small reward; **Figure 6B**) similar to what we described for DA neurons in VTA (**Figure 4E**).

Although these results are consistent with the notion that activity in SNr reflects a common output, even in SNr, correlations between delay and size were relatively weak; leaving open the interpretation that SNr might also maintain independent representations of expected size and delay. These data suggest that we have to move very close to the motor system before delay and size are represented as a common signal, and it is not clear whether such representations exist in many regions upstream. According to our data, the majority of brain areas involved in the circuit critical for learning and decision-making based on expected outcomes and violations of those expectations encode delayed reward independently from reward size.

The fact that we were able to dissociate the effects of reward size and delay on single-unit activity in these areas indicates that encoding of discounted reward might involve different neural processes than those that signal expected reward value. This dissociation is perhaps not surprising considering recent behavioral data that supports the view that learning about sensory and temporal features of stimuli involve different underlying systems (Delamater and Oakeshott, 2007) and that studies that report abnormal delay discounting functions often report no observable change in behaviors guided by reward size.

With that said, other studies have shown that neural activity related to size and delay are correlated in several frontal areas in primate cortex (Roesch and Olson, 2005a,b; Kim et al., 2008). For example, in primates, OFC neurons that fired more strongly for shorter delays tended to fire more strongly for larger rewards. Our ability to detect independent encoding might reflect a species difference and/or a number of other task parameters; however we would like to think that differences might emerge from different levels of training. With extended training, OFC neurons may become optimized to provide generic value representations. This would have interesting implications as it would suggest that OFC and possibly other brain areas might refrain from putting delay and size on a common value scale until they have been integrated for an extended time. This might be why single neurons in primate frontal cortex and striatum have been shown to be modulated by both size and delay (Kim et al., 2008; Cai et al., 2011).

Another possibility is that common value signals observed in primates reflect the fact that, over time, short delay trials sometimes led to more reward. That is, since short delay trials took less time to complete, more reward could be obtained over the course of the recording session. Unlike the rat work, delays were not normalized in some of these studies (Roesch and Olson, 2005a,b), thus raising the possibility that brain areas might commonly encode size and delay only when shorter delays are genuinely more valuable,

FIGURE 6 | Substantial nigra pars reticulata (SNr). Activity of single neurons in SNr reflects an interaction between expected value and direction. **(A)** Activity of a single SNr neuron averaged over all trials for each condition aligned on odor port exit during all eight conditions (four rewards × two directions). Histogram represents average activity over the last 10 trials (after learning) for each condition in a block of trials. Each tick mark is an action potential and trials are represented by rows. All trials are shown. **(B)** Correlation between size (big − small/big + small) and delay (short − long/short + long) effects averaged across direction (odor onset to odor port exit). Data was taken after learning (last 10 trials for each condition within each block; Bryden et al., 2011).

not just subjectively preferred. The possibility that these variables might be encoded separately in primates is also consistent with recent work showing that risk is sometimes encoded separately from reward size in primate OFC (Kennerley and Wallis, 2009; Kennerley et al., 2009, O'Neill and Schultz, 2010; Schultz, 2010; Wallis and Kennerley, 2010).

A final possibility is that we did not vary delay and magnitude simultaneously. True discounting studies manipulate size and delay at the same time to demonstrate the antagonistic effects of reward magnitude and delay. Certainly, Lee and colleagues have found more integrative encoding of value in the brain than we have using this procedure (Kim et al., 2008; Cai et al., 2011). This would suggest that when size and delay are manipulated simultaneously the brain encodes them together, but when they are split apart, they are represented independently. More work is necessary to determine if this theory holds up. Still, there are other differences between tasks that might impact how the brain encodes these two variables. In our task rats are constantly forming and updating response–outcome associations as they learn to bias behavior in one direction when rewards change in size or delay. Independent representations of size and delay might help the brain cope with these changing circumstances.

That fact that size and delay are not strongly correlated in most brain areas that we have tested does not mean that the rat or that other brain areas might treat them similarly. Remarkably, out of all the brain areas that we have recorded from in this task only the firing of DA neurons in VTA showed a strong clear cut relationship between manipulations of delay and size (**Figure 4E**). DA neurons fired more strongly to cues that predicted a short delay and large reward and were inhibited by cues that predicted a small reward and a long delay (**Figure 4E**). These were the same neurons in which activity reflected prediction errors during unexpected reward delivery and omission when reward was made larger or smaller than expected and when reward was delivered earlier or later than expected (**Figure 4**). The fact that the activity of DA neurons represents cues that predict expected size and delay similarly does not fit well with the finding that other areas do not, considering that it is dopaminergic input that is thought to train up associations in these areas. Why and how delay information remains represented separately from value is an intriguing question and requires further investigation.

CONCLUSION

Here we speculate on the circuit that drives discounting behavior based on the neural correlates related to size and delay as described above. It is important to remember that much of this is based on neural correlates and we are currently trying to work out the circuit using lesion and inactivation techniques combined with single-unit recordings.

According to our data, when an immediate reward is delivered unexpectedly, DA neurons burst, consistent with a signal that detects errors in reward prediction (Roesch et al., 2007b). ABL neurons also respond to unexpected immediate reward but several trials later, consistent with signals that detect the need for increased attention or event processing during learning (Roesch et al., 2010b). As the rat learns to anticipate the reward, expectancy

signals in OFC, ABL, and VS develop. We suspect that development of expectancy signals first occurs in OFC as a consequence of error detection by DA neurons, and that OFC is critical for the development of expectancy signals in ABL and VS. Although all three areas fire in anticipation of reward, they might be carrying unique signals related to reward outcome values, attention, and motivation, respectively. As expectancy signals increase, prediction error signaling at the time of reward delivery decrease and firing of DA neurons start to fire to cues that predict the immediate reward (**Figure 4**). Cue-evoked responses that develop in DA neurons subsequently stamp in associations in OFC, VS, and DS. Interactions between ABL and these areas might be particularly important in this process; lesions of ABL impairs cue development in OFC and VS (Schoenbaum et al., 2003; Schoenbaum and Roesch, 2005; Ambroggi et al., 2008). It is still unclear whether ABL–DA interactions are necessary for cue selectivity to develop in themselves and in downstream areas (Roesch et al., 2010a). After learning, OFC and VS guide decision-making via reward specific outcome values and affective/motivational associations, respectively and with DS guiding behavior by signaling action–value and stimulus–response associations (Stalnaker et al., 2010). Positive prediction errors likely impact striatal output via D1 mediated direct pathways to SNr promoting movement via disinhibition of downstream motor areas (Bromberg-Martin et al., 2010).

So what happens when rewards are delayed? After learning, there are strong expectancy signals in OFC for the immediate reward. Expectancy activity in OFC for the immediate reward persists even when reward is no longer available at that time (e.g., **Figure 2B**). Thus, when the immediate reward is not delivered, a strong negative prediction error is generated by DA neurons (**Figure 4B**). Inhibition of DA should reduce associability with reward in downstream areas, thus inhibiting responses to cues signaling the location of the delayed reward. Further, attenuated expectancy signals generated in OFC would reduce expectancy signals reliant on it, as shown for ABL and possibly for VS, that might aid in maintaining responding for the now delayed reward (Saddoris et al., 2005). Reduction of these signals might further decrease associability with the delayed reward. Subsequently, DA neurons start to inhibit firing to cues that predict the delayed reward, weakening associations in downstream areas such as OFC, ABL, and striatum. Decreased DA transmission in striatum would impact the D2 mediated indirect pathway inducing

increases in SNr firing that increases suppression of movement by inhibiting downstream motor structures (Bromberg-Martin et al., 2010).

It is important to note that these are not the only brain areas involved in temporal discounting and inter-temporal choice. Serotonin clearly plays a role but exactly what role it plays is still a little murky. Serotonin depletion, sometimes, but not always, steepens the discounting of delayed rewards making animals more impulsive (Wogar et al., 1993; Harrison et al., 1997; Bizot et al., 1999; Evenden and Ryan, 1999; Mobini et al., 2000a,b; Cardinal et al., 2004; Winstanley et al., 2004a,c; Denk et al., 2005; Cardinal, 2006), and increased extracellular serotonin concentrations promotes selection of large-delayed reward over smaller immediate reward (Bizot et al., 1988, 1999). Furthermore, recent data demonstrates that serotonin efflux in rat dorsal raphe nucleus increase when animals have to wait for reward and single dorsal raphe neurons fire in anticipation of delayed reward (Miyazaki et al., 2011a,b).

Work in humans has also clearly defined a role for PFC and other cortical areas in inter-temporal choice especially when decisions have to made for rewards that will arrive in the distant future (e.g., months to years; McClure et al., 2004, 2007; Tanaka et al., 2004; Kable and Glimcher, 2007; Ballard and Knutson, 2009; Figner et al., 2010). These systems likely interact on several levels to control behavior when expected rewards are delayed.

In conclusion, it is clear that discounting behavior is complicated and impacts a number of systems. From the results described above, when an individual chooses between an immediate versus delayed reward, the decision ultimately depends on previous experience with the delayed reward and the impact that a delayed reward has on neural processes related to reward expectation, prediction error encoding, attention, motivation, and the development of associations between stimuli, responses, and outcome values. To elucidate the underlying cause of the many disorders that impact impulsivity, we must address which of these processes are impaired and further test the circuit involved in inter-temporal choice.

REFERENCES

Ainslie, G. W. (1974). Impulse control in pigeons. *J. Exp. Anal. Behav.* 21, 485–489.

Ambroggi, F., Ishikawa, A., Fields, H. L., and Nicola, S. M. (2008). Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron* 59, 648–661.

Ballard, K., and Knutson, B. (2009). Dissociable neural representations of future reward magnitude and delay during temporal discounting. *Neuroimage* 45, 143–150.

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

Bechara, A., Damasio, H., Damasio, A. R., and Lee, G. P. (1999). Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J. Neurosci.* 19, 5473–5481.

Bechara, A., Dolan, S., and Hindes, A. (2002). Decision-making and addiction (part II): myopia for the future or hypersensitivity to reward? *Neuropsychologia* 40, 1690–1705.

Belova, M. A., Paton, J. J., Morrison, S. E., and Salzman, C. D. (2007). Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron* 55, 970–984.

Bezzina, G., Cheung, T. H., Asgari, K., Hampson, C. L., Body, S., Bradshaw, C. M., Szabadi, E., Deakin, J. F., and Anderson, I. M. (2007). Effects of quinolinic acid-induced lesions of the nucleus accumbens core on inter-temporal choice: a quantitative analysis. *Psychopharmacology (Berl.)* 195, 71–84.

Bizot, J., Le Bihan, C., Puech, A. J., Hamon, M., and Thiebot, M. (1999). Serotonin and tolerance to delay of reward in rats. *Psychopharmacology (Berl.)* 146, 400–412.

Bizot, J. C., Thiebot, M. H., Le Bihan, C., Soubrie, P., and Simon, P. (1988). Effects of imipramine-like drugs and serotonin uptake blockers on delay of reward in rats. Possible implication in the behavioral mechanism of action of antidepressants. *J. Pharmacol. Exp. Ther.* 246, 1144–1151.

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 158(Val/Val) genotype. *J. Neurosci.* 27, 14383–14391.

Bowman, E. M., Aigner, T. G., and Richmond, B. J. (1996). Neural signals in the monkey ventral striatum related to motivation for juice and cocaine rewards. *J. Neurophysiol.* 75, 1061–1073.

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.

Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834.

Bryden, D. W., Johnson, E. E., Diao, X., and Roesch, M. R. (2011). Impact of expected value on neural activity in rat substantia nigra pars reticulata. *Eur. J. Neurosci.* 33, 2308–2317.

Bucci, D. J., and Macleod, J. E. (2007). Changes in neural activity associated with a surprising change in the predictive validity of a conditioned stimulus. *Eur. J. Neurosci.* 26, 2669–2676.

Cai, X., Kim, S., and Lee, D. (2011). Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron* 69, 170–182.

Calu, D. J., Roesch, M. R., Haney, R. Z., Holland, P. C., and Schoenbaum, G. (2010). Neural correlates of variations in event processing during learning in central nucleus of amygdala. *Neuron* 68, 991–1001.

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

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.

Cardinal, R. N., Robbins, T. W., and Everitt, B. J. (2000). The effects of d-amphetamine, chlordiazepoxide, alpha-flupentixol and behavioural manipulations on choice of signalled and unsignalled delayed reinforcement in rats. *Psychopharmacology (Berl.)* 152, 362–375.

Cardinal, R. N., Winstanley, C. A., Robbins, T. W., and Everitt, B. J. (2004).

Limbic corticostriatal systems and delayed reinforcement. *Ann. N. Y. Acad. Sci.* 1021, 33–50.

Carelli, R. M. (2002). Nucleus accumbens cell firing during goal-directed behaviors for cocaine vs 'natural' reinforcement. *Physiol. Behav.* 76, 379–387.

Church, R. M., and Gibbon, J. (1982). Temporal generalization. *J. Exp. Psychol. Anim. Behav. Process.* 8, 165–186.

Churchwell, J. C., Morris, A. M., Heurtelou, N. M., and Kesner, R. P. (2009). Interactions between the prefrontal cortex and amygdala during delay discounting and reversal. *Behav. Neurosci.* 123, 1185–1196.

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.

Cousens, G. A., and Otto, T. (2003). Neural substrates of olfactory discrimination learning with auditory secondary reinforcement. I. Contributions of the basolateral amygdaloid complex and orbitofrontal cortex. *Int. Physiol. Behav. Sci.* 38, 272–294.

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.

Cromwell, H. C., and Schultz, W. (2003). Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J. Neurophysiol.* 89, 2823–2838.

Dalley, J. W., Mar, A. C., Economidou, D., and Robbins, T. W. (2008). Neurobehavioral mechanisms of impulsivity: fronto-striatal systems and functional neurochemistry. *Pharmacol. Biochem. Behav.* 90, 250–260.

Day, J. J., Jones, J. L., and Carelli, R. M. (2011). Nucleus accumbens neurons encode predicted and ongoing reward costs in rats. *Eur. J. Neurosci.* 33, 308–321.

Delamater, A. R., and Oakeshott, S. (2007). Learning about multiple attributes of reward in Pavlovian conditioning. *Ann. N. Y. Acad. Sci.* 1104, 1–20.

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 (Berl.)* 179, 587–596.

Ernst, M., Zametkin, A. J., Matochik, J. A., Jons, P. H., and Cohen, R. M. (1998). DOPA decarboxylase

activity in attention deficit hyperactivity disorder adults. A [fluorine-18]fluorodopa positron emission tomographic study. *J. Neurosci.* 18, 5901–5907.

Evenden, J. L., and Ryan, C. N. (1996). The pharmacology of impulsive behaviour in rats: the effects of drugs on response choice with varying delays of reinforcement. *Psychopharmacology (Berl.)* 128, 161–170.

Evenden, J. L., and Ryan, C. N. (1999). The pharmacology of impulsive behaviour in rats VI: the effects of ethanol and selective serotonergic drugs on response choice with varying delays of reinforcement. *Psychopharmacology (Berl.)* 146, 413–421.

Figner, B., Knoch, D., Johnson, E. J., Krosch, A. R., Lisanby, S. H., Fehr, E., and Weber, E. U. (2010). Lateral prefrontal cortex and self-control in intertemporal choice. *Nat. Neurosci.* 13, 538–539.

Fiorillo, C. D., Newsome, W. T., and Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. *Nat. Neurosci.* 11, 966–973.

Fiorillo, C. D., Tobler, P. N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1856–1902.

Floresco, S. B., St Onge, J. R., Ghods-Sharifi, S., and Winstanley, C. A. (2008). Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. *Cogn. Affect. Behav. Neurosci.* 8, 375–389.

Frank, M. J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J. Cogn. Neurosci.* 17, 51–72.

Gallagher, M., Graham, P. W., and Holland, P. C. (1990). The amygdala central nucleus and appetitive Pavlovian conditioning: lesions impair one class of conditioned behavior. *J. Neurosci.* 10, 1906–1911.

Ghods-Sharifi, S., St Onge, J. R., and Floresco, S. B. (2009). Fundamental contribution by the basolateral amygdala to different forms of decision making. *J. Neurosci.* 29, 5251–5259.

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, 13213–13217.

Harrison, A. A., Everitt, B. J., and Robbins, T. W. (1997). Central 5-HT depletion enhances impulsive responding without affecting

the accuracy of attentional performance: interactions with dopaminergic mechanisms. *Psychopharmacology (Berl.)* 133, 329–342.

Hassani, O. K., Cromwell, H. C., and Schultz, W. (2001). Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J. Neurophysiol.* 85, 2477–2489.

Hatfield, T., Han, J. S., Conley, M., Gallagher, M., and Holland, P. (1996). Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *J. Neurosci.* 16, 5256–5265.

Heerey, E. A., Robinson, B. M., McMahon, R. P., and Gold, J. M. (2007). Delay discounting in schizophrenia. *Cogn. Neuropsychiatry* 12, 213–221.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *J. Exp. Anal. Behav.* 4, 267–272.

Ho, M. Y., Mobini, S., Chiang, T. J., Bradshaw, C. M., and Szabadi, E. (1999). Theory and method in the quantitative analysis of "impulsive choice" behaviour: implications for psychopharmacology. *Psychopharmacology (Berl.)* 146, 362–372.

Holland, P. C., and Gallagher, M. (1993a). Amygdala central nucleus lesions disrupt increments, but not decrements, in conditioned stimulus processing. *Behav. Neurosci.* 107, 246–253.

Holland, P. C., and Gallagher, M. (1993b). Effects of amygdala central nucleus lesions on blocking and unblocking. *Behav. Neurosci.* 107, 235–245.

Holland, P. C., and Gallagher, M. (1993c). Effects of amygdala central nucleus lesions on blocking and unblocking. *Behav. Neurosci.* 107, 235–245.

Holland, P. C., and Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. *Trends Cogn. Sci. (Regul. Ed.)* 3, 65–73.

Holland, P. C., and Gallagher, M. (2006). Different roles for amygdala central nucleus and substantia innominata in the surprise-induced enhancement of learning. *J. Neurosci.* 26, 3791–3797.

Holland, P. C., and Kenmuir, C. (2005). Variations in unconditioned stimulus processing in unblocking. *J. Exp. Psychol. Anim. Behav. Process.* 31, 155–171.

Hollerman, J. R., and Schultz, W. (1998a). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.

Hollerman, J. R., and Schultz, W. (1998b). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309.

Hong, S., and Hikosaka, O. (2011). Dopamine-mediated learning and switching in cortico-striatal circuit explain behavioral changes in reinforcement learning. *Front. Behav. Neurosci.* 5:15. doi:10.3389/fnbeh.2011.00015

Ito, M., and Doya, K. (2009). Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J. Neurosci.* 29, 9861–9874.

Janak, P. H., Chen, M. T., and Caulder, T. (2004). Dynamics of neural coding in the accumbens during extinction and reinstatement of rewarded behavior. *Behav. Brain Res.* 154, 125–135.

Jentsch, J. D., and Taylor, J. R. (1999). Impulsivity resulting from frontostriatal dysfunction in drug abuse: implications for the control of behavior by reward-related stimuli. *Psychopharmacology (Berl.)* 146, 373–390.

Joel, D., Niv, Y., and Ruppin, E. (2002). Actor-critic models of the basal ganglia: new anatomical and computational perspectives. *Neural. Netw.* 15, 535–547.

Jones, B., and Mishkin, M. (1972). Limbic lesions and the problem of stimulus-reinforcement associations. *Exp. Neurol.* 36, 362–377.

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

Kahneman, D., and Tverskey, A. (1984). Choices, values, and frames. *Am. Psychol.* 39, 341–350.

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

Kalenscher, T., Windmann, S., Diekamp, B., Rose, J., Gunturkun, O., and Colombo, M. (2005). Single units in the pigeon brain integrate reward amount and time-to-reward in an impulsive choice task. *Curr. Biol.* 15, 594–602.

Kennerley, S. W., Dahmubed, A. F., Lara, A. H., and Wallis, J. D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables. *J. Cogn. Neurosci.* 21, 1162–1178.

Kennerley, S. W., and Wallis, J. D. (2009). Evaluating choices by single neurons in the frontal lobe: outcome

value encoded across multiple decision variables. *Eur. J. Neurosci.* 29, 2061–2073.

Kesner, R. P., and Williams, J. M. (1995). Memory for magnitude of reinforcement: dissociation between amygdala and hippocampus. *Neurobiol. Learn. Mem.* 64, 237–244.

Khamassi, M., Mulder, A. B., Tabuchi, E., Douchamps, V., and Wiener, S. I. (2008). Anticipatory reward signals in ventral striatal neurons of behaving rats. *Eur. J. Neurosci.* 28, 1849–1866.

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

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

Kobayashi, S., and Schultz, W. (2008). Influence of reward delays on responses of dopamine neurons. *J. Neurosci.* 28, 7837–7846.

Kringelbach, M. L. (2005). The human orbitofrontal cortex: linking reward to hedonic experience. *Nat. Rev. Neurosci.* 6, 691–702.

LeDoux, J. E. (2000). "The amygdala and emotion: a view through fear," in *The Amygdala: A Functional Analysis,* ed. J. P. Aggleton (New York: Oxford University Press), 289–310.

Lowenstein, G. E. J. (1992). *Choice Over Time.* New York: Russel Sage Foundation.

Malkova, L., Gaffan, D., and Murray, E. A. (1997). Excitotoxic lesions of the amygdala fail to produce impairment in visual learning for auditory secondary reinforcement but interfere with reinforcer devaluation effects in rhesus monkeys. *J. Neurosci.* 17, 6011–6020.

Mar, A. C., Walker, A. L., Theobald, D. E., Eagle, D. M., and Robbins, T. W. (2011). Dissociable effects of lesions to orbitofrontal cortex subregions on impulsive choice in the rat. *J. Neurosci.* 31, 6398–6404.

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.

Mirenowicz, J., and Schultz, W. (1994). Importance of unpredictability for reward responses in primate

dopamine neurons. *J. Neurophysiol.* 72, 1024–1027.

Miyazaki, K., Miyazaki, K. W., and Doya, K. (2011a). Activation of dorsal raphe serotonin neurons underlies waiting for delayed rewards. *J. Neurosci.* 31, 469–479.

Miyazaki, K. W., Miyazaki, K., and Doya, K. (2011b). Activation of the central serotonergic system in response to delayed but not omitted rewards. *Eur. J. Neurosci.* 33, 153–160.

Mobini, S., Body, S., Ho, M. Y., Bradshaw, C. M., Szabadi, E., Deakin, J. F., and Anderson, I. M. (2002). Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology (Berl.)* 160, 290–298.

Mobini, S., Chiang, T. J., Al-Ruwaitea, A. S., Ho, M. Y., Bradshaw, C. M., and Szabadi, E. (2000a). Effect of central 5-hydroxytryptamine depletion on inter-temporal choice: a quantitative analysis. *Psychopharmacology (Berl.)* 149, 313–318.

Mobini, S., Chiang, T. J., Ho, M. Y., Bradshaw, C. M., and Szabadi, E. (2000b). Effects of central 5-hydroxytryptamine depletion on sensitivity to delayed and probabilistic reinforcement. *Psychopharmacology (Berl.)* 152, 390–397.

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.

Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.

Monterosso, J., Ehrman, R., Napier, K. L., O'Brien, C. P., and Childress, A. R. (2001). Three decision-making tasks in cocaine-dependent patients: do they measure the same construct? *Addiction* 96, 1825–1837.

Morris, G., Nevet, A., Arkadir, D., Vaadia, E., and Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9, 1057–1063.

Murray, E. A. (2007). The amygdala, reward and emotion. *Trends Cogn. Sci. (Regul. Ed.)* 11, 489–497.

Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y., and Hikosaka, O. (2004). Dopamine neurons can represent context-dependent prediction error. *Neuron* 41, 269–280.

Nakamura, K., and Hikosaka, O. (2006). Role of dopamine in the primate caudate nucleus in reward

modulation of saccades. *J. Neurosci.* 26, 5360–5369.

Nicola, S. M. (2007). The nucleus accumbens as part of a basal ganglia action selection circuit. *Psychopharmacology (Berl.)* 191, 521–550.

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

O'Neill, M., and Schultz, W. (2010). Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron* 68, 789–800.

Padoa-Schioppa, C. (2011). Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* 34, 333–359.

Pan, W. X., Schmidt, R., Wickens, J. R., and Hyland, B. I. (2005). Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J. Neurosci.* 25, 6235–6242.

Parkinson, J. A., Crofts, H. S., McGuigan, M., Tomic, D. L., Everitt, B. J., and Roberts, A. C. (2001). The role of the primate amygdala in conditioned reinforcement. *J. Neurosci.* 21, 7770–7780.

Pennartz, C. M., Groenewegen, H. J., and Lopes da Silva, F. H. (1994). The nucleus accumbens as a complex of functionally distinct neuronal ensembles: an integration of behavioural, electrophysiological and anatomical data. *Prog. Neurobiol.* 42, 719–761.

Redgrave, P., Prescott, T. J., and Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem? *Neuroscience* 89, 1009–1023.

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.

Rodriguez, M. L., and Logue, A. W. (1988). Adjusting delay to reinforcement: comparing choice in pigeons and humans. *J. Exp. Psychol. Anim. Behav. Process.* 14, 105–117.

Roesch, M. R., Calu, D. J., Burke, K. A., and Schoenbaum, G. (2007a). Should I stay or should I go? Transformation of time-discounted rewards in orbitofrontal cortex and associated brain circuits. *Ann. N. Y. Acad. Sci.* 1104, 21–34.

Roesch, M. R., Calu, D. J., and Schoenbaum, G. (2007b). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–1624.

Roesch, M. R., Takahashi, Y., Gugsa, N., Bissonette, G. B., and Schoenbaum, G. (2007c). Previous cocaine exposure makes rats hypersensitive to both delay and reward magnitude. *J. Neurosci.* 27, 245–250.

Roesch, M. R., Calu, D. J., Esber, G. R., and Schoenbaum, G. (2010a). All that glitters... dissociating attention and outcome expectancy from prediction errors signals. *J. Neurophysiol.* 104, 587–595.

Roesch, M. R., Calu, D. J., Esber, G. R., and Schoenbaum, G. (2010b). Neural correlates of variations in event processing during learning in basolateral amygdala. *J. Neurosci.* 30, 2464–2471.

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., Singh, T., Brown, P. L., Mullins, S. E., and Schoenbaum, G. (2009). Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *J. Neurosci.* 29, 13365–13376.

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.

Rudebeck, P. H., Walton, M. E., Smyth, A. N., Bannerman, D. M., and Rushworth, M. F. (2006). Separate neural pathways process different decision costs. *Nat. Neurosci.* 9, 1161–1168.

Saddoris, M. P., Gallagher, M., and Schoenbaum, G. (2005). Rapid associative encoding in basolateral amygdala depends on connections with orbitofrontal cortex. *Neuron* 46, 321–331.

Schoenbaum, G., and Roesch, M. (2005). Orbitofrontal cortex, associative learning, and expectancies. *Neuron* 47, 633–636.

Schoenbaum, G., Setlow, B., Saddoris, M. P., and Gallagher, M. (2003). Encoding predicted outcome and acquired value in orbitofrontal cortex during cue sampling depends upon input from basolateral amygdala. *Neuron* 39, 855–867.

Schultz, W. (2010). Subjective neuronal coding of reward: temporal value discounting and risk. *Eur. J. Neurosci.* 31, 2124–2135.

Sellitto, M., Ciaramelli, E., and di Pellegrino, G. (2010). Myopic discounting of future rewards after medial orbitofrontal damage in humans. *J. Neurosci.* 30, 16429–16436.

Setlow, B., Schoenbaum, G., and Gallagher, M. (2003). Neural encoding in ventral striatum during olfactory discrimination learning. *Neuron* 38, 625–636.

Shen, W., Flajolet, M., Greengard, P., and Surmeier, D. J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* 321, 848–851.

Stalnaker, T. A., Calhoon, G. G., Ogawa, M., Roesch, M. R., and Schoenbaum, G. (2010). Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum. *Front. Integr. Neurosci.* 4:12. doi:10.3389/fnint.2010.00012

Taha, S. A., Nicola, S. M., and Fields, H. L. (2007). Cue-evoked encoding of movement planning and execution in the rat nucleus accumbens. *J. Physiol. (Lond.)* 584, 801–818.

Takahashi, Y. K., Roesch, M. R., Stalnaker, T. A., Haney, R. Z., Calu, D. J., Taylor, A. R., Burke, K. A., and Schoenbaum, G. (2009). The orbitofrontal cortex and ventral tegmental area are necessary for learning from unexpected outcomes. *Neuron* 62, 269–280.

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. (1981). Some empirical evidence on dynamic inconsistency. *Econ. Lett.* 8, 201–207.

Tobler, P. N., Dickinson, A., and Schultz, W. (2003). Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J. Neurosci.* 23, 10402–10410.

Tye, K. M., Cone, J. J., Schairer, W. W., and Janak, P. H. (2010). Amygdala neural encoding of the absence of reward during extinction. *J. Neurosci.* 30, 116–125.

van der Meer, M. A., and Redish, A. D. (2009). Covert expectation of reward in rat ventral striatum at decision points. *Front. Integr. Neurosci.* 3:1. doi:10.3389/neuro.07.001.2009

Wade, T. R., de Wit, H., and Richards, J. B. (2000). Effects of dopaminergic drugs on delayed reward as a measure of impulsive behavior in rats. *Psychopharmacology (Berl.)* 150, 90–101.

Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412, 43–48.

Wallis, J. D., and Kennerley, S. W. (2010). Heterogeneous reward signals in prefrontal cortex. *Curr. Opin. Neurobiol.* 20, 191–198.

Winstanley, C. A. (2007). The orbitofrontal cortex, impulsivity, and addiction: probing orbitofrontal dysfunction at the neural, neurochemical, and molecular level. *Ann. N. Y. Acad. Sci.* 1121, 639–655.

Winstanley, C. A., Dalley, J. W., Theobald, D. E., and Robbins, T. W. (2004a). Fractionating impulsivity: contrasting effects of central 5-HT depletion on different measures of impulsive behavior. *Neuropsychopharmacology* 29, 1331–1343.

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

Winstanley, C. A., Theobald, D. E., Dalley, J. W., Glennon, J. C., and Robbins, T. W. (2004c). 5-HT2A and 5-HT2C receptor antagonists have opposing effects on a measure of impulsivity: interactions with global 5-HT depletion. *Psychopharmacology (Berl.)* 176, 376–385.

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

Wogar, M. A., Bradshaw, C. M., and Szabadi, E. (1993). Effect of lesions of the ascending 5-hydroxytryptaminergic pathways on choice between delayed reinforcers. *Psychopharmacology (Berl.)* 111, 239–243.

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.

Zeeb, F. D., Floresco, S. B., and Winstanley, C. A. (2010). Contributions of the orbitofrontal cortex to impulsive choice: interactions with basal levels of impulsivity, dopamine signalling, and reward-related cues. *Psychopharmacology (Berl.)* 211, 87–98.