OXFORD

ORIGINAL ARTICLE

# Reward Size Informs Repeat-Switch Decisions and Strongly Modulates the Activity of Neurons in Parietal Cortex

## Jan Kubanek[1,2] and Lawrence H. Snyder[1,2]

[1]Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, MO 63110, USA and [2]Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO 63130, USA

Address correspondence to Jan Kubanek, Department of Anatomy and Neurobiology, Washington University School of Medicine, 660 S Euclid Avenue, St. Louis, MO 63110, USA. Email: jan@eye-hand.wustl.edu

## Abstract

Behavior is guided by previous experience. Good, positive outcomes drive a repetition of a previous behavior or choice, whereas poor or bad outcomes lead to an avoidance. How these basic drives are implemented by the brain has been of primary interest to psychology and neuroscience. We engaged animals in a choice task in which the size of a reward outcome strongly governed the animals' subsequent decision whether to repeat or switch the previous choice. We recorded the discharge activity of neurons implicated in reward-based choice in 2 regions of parietal cortex. We found that the tendency to retain previous choice following a large (small) reward was paralleled by a marked decrease (increase) in the activity of parietal neurons. This neural effect is independent of, and of sign opposite to, value-based modulations reported in parietal cortex previously. This effect shares the same basic properties with signals previously reported in the limbic system that detect the size of the recently obtained reward to mediate proper repeat-switch decisions. We conclude that the size of the obtained reward is a decision variable that guides the decision between retaining a choice or switching, and neurons in parietal cortex strongly respond to this novel decision variable.

Key words: economic choice, exploitation, exploration, law of effect, reinforcement

## Introduction

Operant behavior rests on experienced outcomes (Thorndike 1898, 1911; Skinner 1963; Tversky and Kahneman 1986). According to the law of effect, when an outcome is favorable, animals and humans tend to repeat the previous behavior or choice. When an outcome is bad (insufficient or negative), animals and humans tend to avoid the previous behavior or choice, and generate an alternative behavior or choice (Thorndike 1927; Skinner 1953; Lerman and Vorndran 2002; Seo and Lee 2009; Kubanek, Snyder et al. 2015).

A central interest of neuroeconomics, behavioral psychology, and systems neuroscience has been how these behavioral tendencies are implemented at the neuronal level. To this end, many studies have engaged animals in relatively complex tasks in which the outcomes of several past trials predicted the outcome of the current trial. Based on the outcome history, an animal could, according to a particular stimulus schedule or a particular internal behavioral model (Rescorla and Wagner 1972; Sutton and Barto 1998a; Niv 2009), predict the presence or absence of a reward. Neurons in several regions of the brain increase their activity when a high value is anticipated to result from a particular choice (Platt and Glimcher 1999; Roesch and Olson 2003; Dorris and Glimcher 2004; Sugrue et al. 2004, 2005; Samejima et al. 2005; Lau and Glimcher 2008).

The value of an option is critically based on the outcome experienced from choosing that option. Characterizing the neural signatures of each behavioral outcome would greatly facilitate the study of the neuronal drives that underlie the law of effect.

In this regard, in tasks in which animals gamble for a reward, it has been found that many neurons in frontal structures, including the dorsomedial frontal cortex, dorsolateral prefrontal cortex, the anterior cingulate, supplementary eye fields, and in parietal cortex, including the lateral intraparietal area (LIP), distinguish between rewarded and unrewarded choices (Seo et al. 2007, 2009; Seo and Lee 2009; Abe and Lee 2011; Kennerley et al. 2011; So and Stuphorn 2012; Strait et al. 2014). In multiple regions of the striatum, neurons encode the size of the anticipated reward, and some of these neurons respond to the reward outcome itself (Hassani et al. 2001; Cromwell and Schultz 2003). An effect of reward outcome has also been found in the cingulate motor areas (CMAs) in a task in which a reduced reward size was associated with a switch in choice (Shima and Tanji 1998).

Here we investigate the neuronal effects of a behavioral outcome in a simple choice task in which the size of each outcome had a strong effect on animals' tendency to repeat or switch their previous choice. In this task, we recorded responses of neurons in parietal cortex. Parietal neurons have been shown to encode reward-related variables, such as the size of a reward, the probability of a reward, the desirability associated with a given action, or whether a choice was rewarded or not (Platt and Glimcher 1999; Dorris and Glimcher 2004; Musallam et al. 2004; Sugrue et al. 2004; Kable and Glimcher 2009; Seo et al. 2009). A large reward increases the probability of repeating the previous choice, and so it could be expected that neurons with response fields (RFs) representing that choice may show an elevation in activity (Platt and Glimcher 1999; Dorris and Glimcher 2004; Musallam et al. 2004; Sugrue et al. 2004). Alternatively, a large reward reduces the drive to explore alternatives, which may lead to a decrease in the neuronal activity (Shima and Tanji 1998).

We investigate the effect of the size of obtained reward on parietal firing rate and behavior using the data of Kubanek and Snyder (2015a).

## Materials and Methods

### Subjects

We trained 2 male rhesus monkeys (*Macaca mulatta*, 7 and 8 kg) to make a choice between 2 visual targets based on the reward obtained from each target. The animals made the choice using either a saccade or a reach. In both monkeys, we recorded from the hemisphere that is contralateral to the reaching arm. All procedures conformed to the Guide for the Care and Use of Laboratory Animals and were approved by the Washington University Institutional Animal Care and Use Committee.

The animals sat head-fixed in a custom designed monkey chair (Crist Instrument) in a completely dark room. Visual stimuli were back-projected by a CRT projector onto a custom touch panel positioned 25 cm in front of the animals' eyes. Eye position was monitored by a scleral search coil system (CNC Engineering).

### Task

The animal first fixates on and puts his hand on a purple central target (a square of 1 × by 1 visual degrees). After 120 ms, 2 white targets appear, one in the RF of the recorded neuron and one at the opposite location. At the same time, the central target changes color randomly to either red or blue. After a variable delay of 800–1600 ms, the central target disappears, thus cueing the animal to move. To receive a drop of liquid reward, the animal must make a saccade or a reach (if the central target is red or blue, respectively) to within 6 visual degrees of the chosen target. Trials in which the

animal moved the wrong effector, moved prematurely, or moved inaccurately were aborted and not subsequently analyzed.

Critically, in this task, each target is associated with a reward on each trial. The reward consists of a drop of water, delivered by the opening of a valve for a particular length of time. The associated rewards have a ratio of either 3:1 or 1.5:1. The ratio is held constant in blocks of 7–17 trials (exponentially distributed with a mean of 11) and then changed to either 1:3 or 1:1.5. The time that the reward valve is held open is drawn from a truncated exponential distribution that ranges from 20 to 400 ms. The mean of the exponential distribution differs for each target and depends on the reward ratio for that block. For a reward ratio of 1.5:1 (3:1), the means for the richer and poorer target are 140 and 70 ms (250 and 35 ms), respectively (truncating these distributions between 20 to 400 ms leads to the desired 1.5:1 (3:1) ratios of mean rewards). The distributions of the reward values actually chosen by the monkeys are shown in Figure 2D. This randomization prevented the animals from stereotypically choosing the more valuable option. To help prevent animals from overlearning the specific distributions of reward durations, we further randomized reward delivery for any chosen target by multiplying valve open times by a value between 80% and 120%. This value was changed on average every 70 trials (exponential distribution truncated to between 50 and 100). An auditory cue was presented to the monkeys during the time the valve was open.

### Electrophysiological Recordings

We lowered glass-coated tungsten electrodes (Alpha Omega, impedance 0.5–3 MΩ at 1 kHz) 2.8–10.8 mm below the dura into LIP, and 2.1–11.6 mm below the dura into parietal reach region (PRR). We detected individual action potentials using a dual-window discriminator (BAK Electronics). A custom program ran the task and collected the neural and behavioral data. Anatomical magnetic resonance scans were used to localize the lateral and medial bank of the intraparietal sulcus. We next identified a region midway along the lateral bank containing a high proportion of neurons with transient responses to visual stimulation, strong peri-saccadic responses, and sustained activity on saccade trials that was greater or equal to that on reach trials (LIP); and a second region towards the posterior end of the medial bank containing a high proportion of neurons with transient responses to visual stimulation and sustained activity that was greater on reach than saccade trials (PRR). The requirement of sustained activity was considered as one of the defining properties of areas LIP and PRR. Once a cell was isolated, we characterized its RF. Specifically, we tested responses to targets at one of 8 equally spaced polar angles and 2 radial eccentricities (12 or 18 visual degrees), and defined the RF by the direction and eccentricity that elicited the maximal transient response from a given neuron. We recorded from cells that showed maintained activity during the delay period for either a saccade or a reach (about half of all cells in LIP and in PRR). We recorded as many trials in the main task as possible (an average of 340 valid trials per cell). We recorded from 40 LIP neurons and 31 PRR neurons in monkey A, and from 20 LIP neurons and 34 PRR neurons in monkey B.

### Desirability of the RF Target

We inferred the desirability of each target based on the animals' behavior in this task (Platt and Glimcher 1999; Dorris and Glimcher 2004; Sugrue et al. 2004). To do so, we applied a reinforcement-learning model (Sutton and Barto 1998b; Seo and Lee 2009). In the reinforcement-learning model, on every trial t, the desirability

of the RF target is defined as the difference between the value function assigned to the RF target, $V_t(o)$, and the value function assigned to the opposite target, $V_t(o')$:

$$\text{Desirability} = V_t(o) - V_t(o').$$

The value function of a selected option $o$ on trial $t$, $V_t(o)$, is updated according to a learning rule:

$$V_t(o) = \alpha V_{t-1}(o) + (1-\alpha)r_{t-1},$$

where $V_{t-1}(o)$ is the value function of option $o$ on previous trial, $r_{t-1}$ is the reward received on the previous trial, and $\alpha$ denotes the learning rate. The value function of the unchosen option, $V_t(o')$ is not updated.

The probability of choosing the RF option $o$ is then a logistic function of the desirability:

$$P(o) = \frac{1}{1 + \exp(-\beta(\text{desirability} + E))}$$

Here $\beta$ is the inverse temperature parameter and $E$ is an intercept to account for fixed biases for one target over the other. We used separate intercepts for each effector.

The parameters $\alpha$, $\beta$, and the 2 intercept terms were fitted to behavioral data obtained when recording from each cell using the maximum likelihood procedure, maximizing the log likelihood criterion ($L$):

$$L = \Sigma_t \ln P_t(o(t))$$

where $P_t(o(t))$ is, as given above, the probability of choosing option $o(t)$ on trial $t$ (note that $P_t(o(t)) = 1 - P_t(o'(t))$).

We fitted separate reinforcement-learning model coefficients to account for the behavioral data obtained while recording from each of the parietal neurons. This gave $\alpha = 0.13 \pm 0.086$ and $\beta = 0.021 \pm 0.0063$ (mean ± SD, across all neurons).

This model accounted for the macroscopic choice behavior and to a degree also for the microscopic (trial-by-trial) choice behavior (Kubanek and Snyder 2015a,b), performed similarly as alternative models (Platt and Glimcher 1999; Dorris and Glimcher 2004; Sugrue et al. 2004, 2005), and, critically, the modeled RF target desirability reproduced the previous findings (Platt and Glimcher 1999; Dorris and Glimcher 2004; Sugrue et al. 2004) that parietal neurons increase their firing with increasing RF desirability (Fig. 6).

## A More Complex Model

The randomization of the reward size in each trial was imposed and adjusted over the course of training so that animals could not anticipate the time of a reward ratio transition. Indeed, our analyses of these data confirmed that the animals' behavior did not exhibit an appreciable anticipation of a ratio transition (Kubanek and Snyder 2015a,b). This is apparent in Figure 1B of Kubanek and Snyder 2015b, in which prior to a ratio transition (trial 0), the average choice proportions are constant across trials. An anticipation of a transition would manifest itself as a rising ramp in the average choice proportions prior to the actual transition. No such rising ramp is observed.

Furthermore, the reinforcement-learning model captures the behavioral dynamics prior to and following a transition very well. This is demonstrated in Figure 4B of Kubanek and Snyder 2015a,

in which the monkeys' behavior (solid lines) and the model's behavior (dashed line) are closely matched.

Nonetheless, we tested a more complex model that has explicit knowledge about the transition statistics. Following repeated choices, the model encourages a switch by decreasing the probability of choosing the same option. This probability, $h_s$, matches the transition statistics of the task (i.e., $h_s$ is the hazard rate function corresponding to an exponential distribution with the mean equal to 11 trials and trimmed between 7 and 17 trials). This switch tendency is incorporated into the model as follows:

$$P(o) = \frac{1}{1 + \exp(-\beta(\text{desirability} + E))} - \gamma h_s(o)$$

The additional parameter, $\gamma$, is fitted to the data in the same way as the parameter fits described for the simpler model.

As expected—given the lack of appreciable signs of transition anticipation—this more complex model performed only marginally better than the simpler model (correlation between the model's predictions, $P(o)$, and the animals' choices: $r = 0.56$ for the simpler model, $r = 0.58$ for the more complex model). We therefore used the simpler model.

## Results

We engaged monkeys in a reward-based choice task in which animals chose between 2 targets (Fig. 1). One target offered, in blocks of 7–11 trials, a higher average mean reward than the other target. Critically, the delivered reward varied from trial to trial, even for choices of the same target. In particular, the open times of the valve that delivered liquid reward were drawn from one of 2 exponential distributions, depending on which target was selected (see Materials and Methods).

As expected (Neuringer 1967; Rescorla and Wagner 1972; Leon and Gallistel 1998; Platt and Glimcher 1999; Schultz 2006), the animals' choice behavior in this task was highly sensitive to the size of the reward that was delivered on the preceding trial (Fig. 2A). When animals obtained a large reward (valve open duration >160 ms), the animals almost always repeated their previous choice (probability to switch equal to 0.04). In contrast, when the obtained reward was small (valve open duration <80 ms), the animals often avoided their previous choice and acquired the other target instead (probability to switch equal to 0.52). We quantified the relationship between the tendency to stay versus switch and the obtained reward by first normalizing the obtained reward so that most of its values lie between 0 (2.5 percentile of the reward sizes) and 1 (97.5th percentile). We then fitted a slope to the relationship between the binary variable indicating that a switch
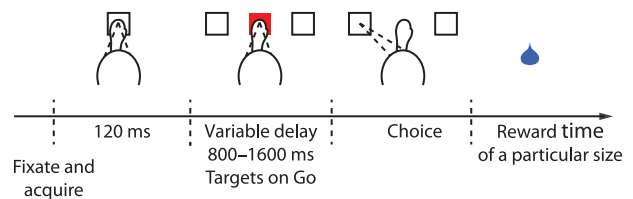


**Figure 1.** Choice task with variable reward size. Animals first fixated and put their hand on a central target. Following a short delay, 2 targets appeared in the periphery. One target was placed in the RF of the recorded neuron, the other target outside of the RF. The animals acquired one of the targets with either an eye or hand movement, depending on whether the central cue was red or blue, respectively. Each choice was followed by th delivery of a liquid reward of a particular size (see text).
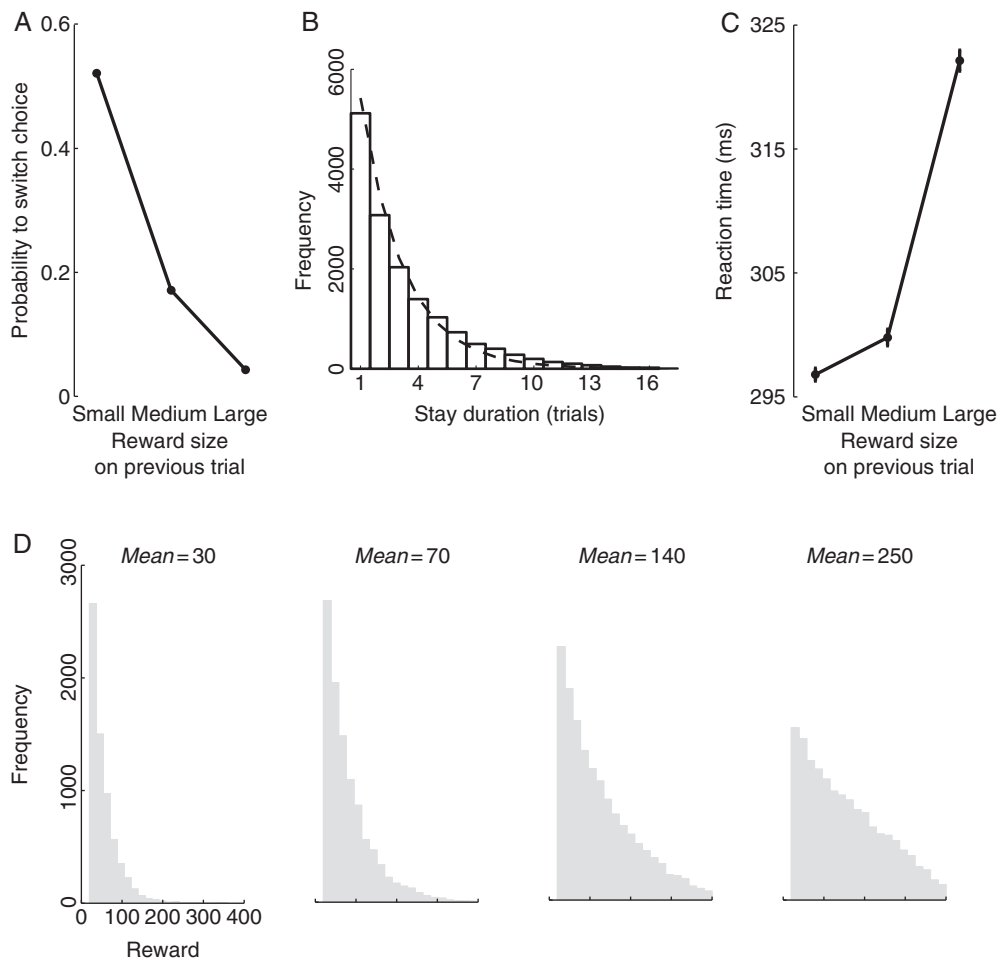
**Figure 2.** Choice behavior governed by reward size. (A) Mean ± SEM frequency of switching choice from one target to the other, as a function of the size of the reward obtained on the previous trial. For display purposes, the reward magnitudes were binned into bins of small (valve opening durations <80 ms), medium (valve opening durations ≥80 ms and <160 ms) and large rewards (valve opening durations ≥160 ms). The SEMs are smaller than the data point markers in this figure. (B) The histogram of the number of consecutive choices of the same target. The dashed line represents an exponential fit. (C) Mean ± SEM reaction time as a function of the magnitude of the reward size obtained on the previous trial. (D) Histograms of actual reward values (valve opening times) obtained by the monkeys for choosing an option from a distribution characterized by a particular mean value.

occurred and the normalized reward. The slope was −0.85 and −0.75 in monkey A and B, respectively, and the slopes were highly significant ($P < 0.0001$, monkey A: $t_{27,406} = -80.0$; monkey B: $t_{20,889} = -60.7$). The animals' behavior was similar around the time of a reward ratio transition (transition trial and 3 following trials; slope: −0.81, $P < 0.0001$, $t_{18,112} = -58.7$) and nontransition trials (all other trials; slope: −0.79, $P < 0.0001$, $t_{31,091} = -81.2$). We also quantified the relationship between the obtained reward size on the tendency to switch on an individual trial basis. In particular, the Pearson's correlation between the obtained reward size and the binary switch variable in each trial is $r = 0.42 \pm 0.06$ (mean ± SD, 125 sessions). These analyses indicate that the reward size informed each subsequent decision whether to repeat or switch the previous choice.

Due to the sensitivity of our animals to the reward size in this task, animals switched from one target to another often (Fig. 2B), on average about every third trial (probability of switching, $P = 0.31$). The stay duration histogram was well approximated by an exponential (Fig. 2B), which suggests (though it does not prove) that the choice an animal made on a given trial was independent of the choice the animal made on the previous trial.

The reward size modulated the choice reaction time (RT) in an unexpected way. The larger the reward on the previous trial, the

longer it took to make a choice on the subsequent trial (Fig. 2C). The slopes of the relationship between the RT and the normalized reward were 76.9 and 22.3 ms per the reward range in monkey A and monkey B, respectively, and the slopes were highly significant ($P < 0.0001$, monkey A: $t_{27,406} = 25.6$; monkey B: $t_{20,889} = 7.0$). Importantly, the effect persisted when we considered only "stay" trials (monkey A: 88.4 ms per reward range, $P < 0.0001$, $t_{18,895} = 24.8$; monkey B: 21.2 ms per reward range, $P < 0.0001$, $t_{14,445} = 5.9$) and when we considered only "switch" trials (monkey A: 86.1 ms per reward range, $P < 0.0001$, $t_{85,09} = 8.1$; monkey B: 66.0 ms per reward range, $P < 0.0001$, $t_{64,40} = 6.1$). This rules out the possibility that the animal is faster after a small reward because it switches its choice and anticipates, as a result of the switch, a larger reward. Many studies have shown that choosing an option associated with a larger "anticipated" reward typically results in a faster RT (Stillings et al. 1968; Hollerman et al. 1998; Tremblay and Schultz 2000; Cromwell and Schultz 2003; Roesch and Olson 2004). However, to our knowledge, previous studies have not investigated the effect of previous reward size on the subsequent RT. Our result therefore constitutes a novel effect of reward size on behavior.

The large effect that the reward size exerts on choice behavior suggests a correspondingly large effect on neuronal activity in
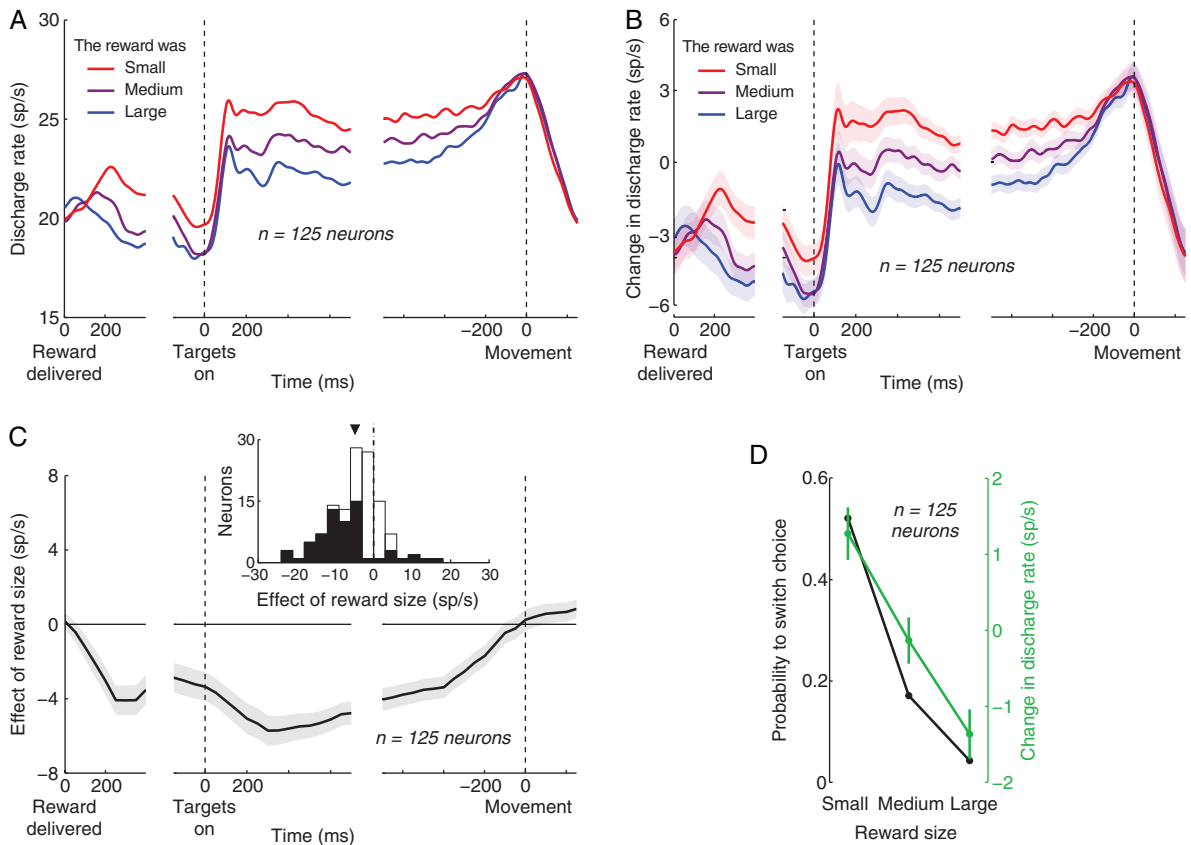
**Figure 3.** Parietal activity is strongly modulated by the size of the just received reward. (A) Mean discharge rate of 125 parietal neurons (60 LIP, 65 PRR) as a function of time, separately for the cases in which the immediately preceding reward size (time 0 on the left) was small, medium, or large (same discretization as in Fig. 2). Activity is aligned on the closing time of the valve (left), the onset of the choice targets (middle), and movement (right). The spike time histograms (1 ms resolution) were smoothed using a zero-lag Gaussian filter of standard deviation equal to 20 ms. (B) As in (A) with the exception that the mean discharge rate measured between the time of the reward offset and movement is subtracted from the discharge rate profile of each neuron, before the rates are averaged over the neurons. A similar result is obtained when the activity is divided by the mean rate. The shading represents the SEM over the 125 neurons. (C) Mean ± SEM slope of the line fitted to the relationship between the discharge rate and the normalized reward size. The discharge rate was measured in 300 ms windows sliding through the trial in 50 ms steps. The slope was computed in each of these windows. The reward size was normalized between 0 (2.5 percentile) and 1 (97.5th percentile). The inset shows the slope for the discharge activity measured between the time of the target onset and the go cue, separately for each neuron. The filled parts of the histogram represent neurons in which the slope was significant ($P < 0.05$, t-statistic). The triangle denotes the mean of the distribution. (D) Comparison of the the effects of reward size on the choice behavior (black; same as in Fig. 2A) and the parietal activity (green). As in the inset in (C), the discharge rate was measured between the time of the target onset and the go cue, and was normalized for each cell in the same way as in (B). The error bars represent the SEM; the black error bars are smaller than the symbols indicating the data points.

brain regions that are implicated in valuation and choice. We recorded from 2 such regions in the parietal cortex—the LIP and the PRR. These regions encode reward-related decision variables in tasks in which animals may move either to a target within the RF of the recorded neuron or to a target outside of the RF (Platt and Glimcher 1999; Dorris and Glimcher 2004; Musallam et al. 2004; Sugrue et al. 2004). LIP activity is more strongly modulated by choices made using saccades, whereas PRR activity is more strongly modulated by choices made using reaches (Calton et al. 2002; Dickinson et al. 2003; Cui and Andersen 2007). We collapsed the data over both choice effectors; in Figure 7 we investigate the effector specificity of the main effect reported in this study. We recorded activity of a total of 125 neurons in this task —60 neurons in LIP and 65 neurons in PRR.

We found that the size of the previous reward strongly modulates subsequent neuronal activity (Fig. 3). The effect points in a surprising direction (Fig. 3A). Small rewards (red) were followed by high activity, whereas large rewards (blue) were followed by low activity. We further normalized the neuronal activity by subtracting the mean discharge rate of each neuron, measured between the time of the reward offset and movement, before

averaging the activity over the neurons (Fig. 3B). The relatively small standard errors of the normalized mean activity (shading in Fig. 3B) indicate that reward size contributes with significant information to the discharge activity of parietal neurons. A significant distinction between the effects of small and large rewards ($P < 0.01$, two-sample t-test) is observed at 209 ms following the reward delivery offset. The difference prevails throughout the trial until it loses significance ($P > 0.01$) at 299 ms preceding the next response.

We quantified the effect of reward size on subsequent neuronal activity by computing the slope of the line fitted to the relationship between the raw neuronal discharge rate and the reward size (again normalized between 0 [2.5th percentile of the reward sizes] and 1 [97.5th percentile]). We performed this regression for discharge activity measured in 300 ms windows sliding through the trial by 50 ms (Fig. 3C). The figure confirms the impression of Figure 3A,B that the effect of the reward magnitude emerges soon following the reward delivery. Specifically, the effect gains significance ($P = 0.0048$, t-test, $n = 125$) in the window centered at 150 ms following the reward delivery offset. The effect reaches a maximum (–5.7 sp/s per the reward size range) in the window
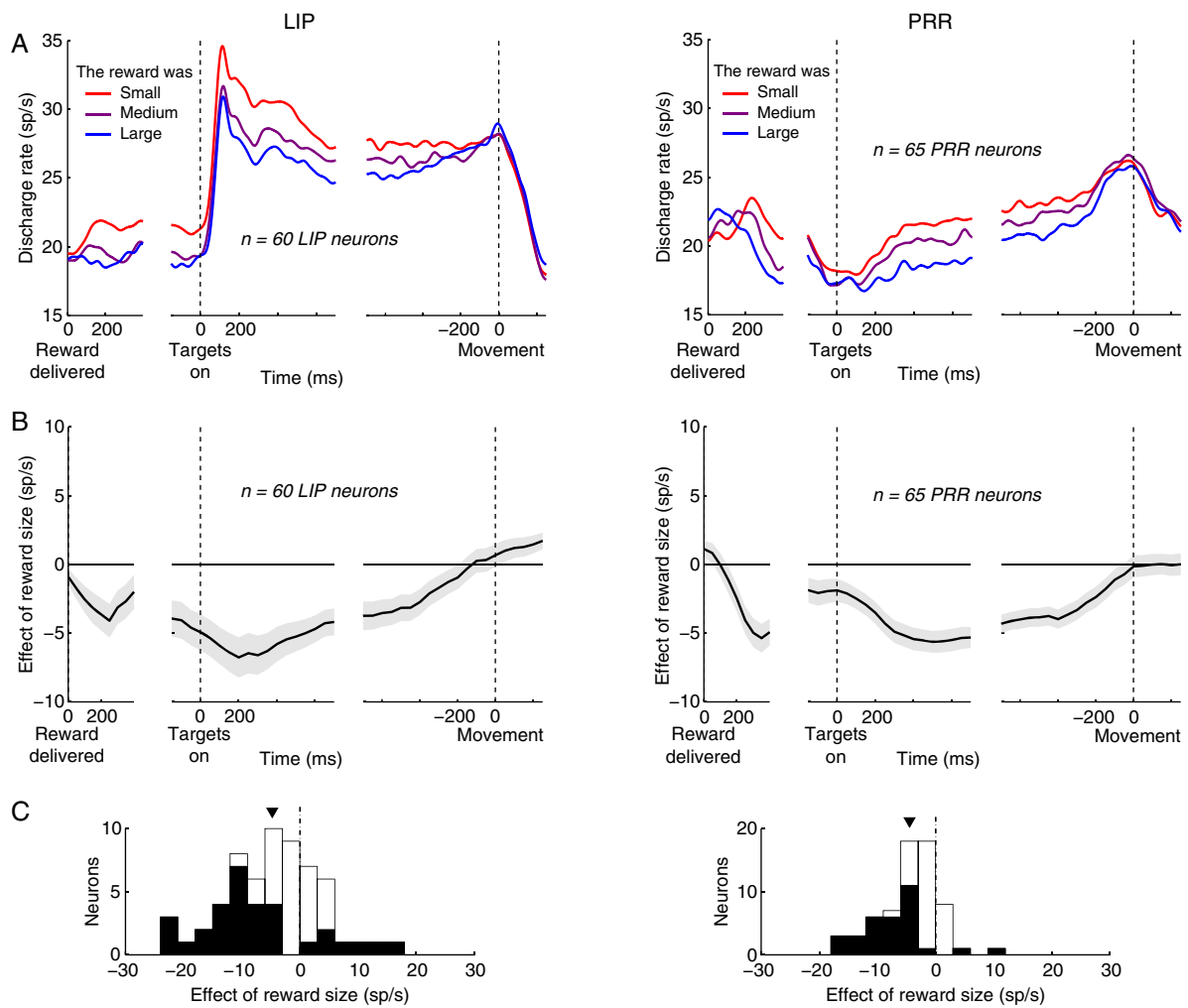
**Figure 4.** The effect of reward size in two regions of parietal cortex. Same analyses and similar format as in Figure 3. The histograms are now shown in a separate (C). The (B, D) of Figure 3 are not present in this figure. The data are displayed separately for area LIP (left column) and PRR (right column).

centered at 300 ms following the target onset. The effect vanishes ($P = 0.065$) in the window centered at 150 ms preceding the movement.

We quantified the effect of reward size for each individual neuron. To do that, we measured the discharge rate in the interval starting with the target onset and ending with the go cue, and computed the modulation (slope) of the discharge rate by the normalized reward size. The resulting distribution of the slopes over the individual neurons is shown in the inset of Figure 3C. The plot further corroborates the finding that the reward size has a strong negatively signed effect on the neuronal activity. The mean of the distribution (triangle above the distribution) is a −4.6 sp/s decrease of discharge activity over the range of the reward sizes. This effect is significant ($P < 0.0001$, $t_{124} = −7.5$; monkey A: −5.6 sp/s, $P < 0.0001$, $t_{73} = −6.4$; monkey B: −3.1 sp/s, $P < 0.001$, $t_{50} = −4.1$). A majority of the recorded neurons (98/125, 78%) showed a decrease of activity with the increasing reward size, and the effect was significant (significance of slope, $P < 0.05$, t-statistic) in a majority (56%) of these 98 neurons (black portions of the histogram). Within the 27 (22%) neurons which showed an increase of activity with increasing reward size, the effect was significant in only 33% of cases. The mean effect for all neurons that encoded reward size significantly (64/125, 51%) was a −7.5 sp/s decrease of discharge activity over the range of the

reward sizes ($P < 0.0001$, $t_{63} = −7.5$). For these neurons, the effect of reward size reached an average of −9.5 sp/s in the window centered at 300 ms following the target onset ($P < 0.0001$, $t_{63} = −7.6$).

The size of the obtained reward only influenced activity in the immediately following trial. There was no persistent activity over multiple trials, such as 2 trials following a reward ($P = 0.087$, $t_{124} = −1.73$).

We compared the effect of reward size on the choice behavior with its effect on the neuronal activity (Figure 3D). There is a match between the effect of reward size on the probability to switch choice (black) and the discharge rate (green). The larger the obtained reward, the less likely the animals were to switch their choice, and the less vigorously parietal neurons fired action potentials.

Next, we demonstrate the effect separately for LIP and PRR (Fig. 4). This figure rests on the same analyses as those performed in Figure 3. The histograms are now shown in a separate panel C, and the data are displayed separately for LIP (left column) and PRR (right column). The figure reveals that the effect is present in both areas, and to a similar extent. Perhaps the most salient difference between LIP and PRR is that regarding the raw discharge activity (Fig. 4A). LIP is much more responsive to the onset of the choice targets, indicated by time 0 in the middle panel (Kubanek et al. 2013). However, the dynamics of the grading

of neural activity due to the reward size (the individual colored traces in Fig. 4A) are similar in both LIP and PRR. This is apparent in Figure 4B. In both LIP and PRR, the effect rises gradually throughout the trial, reaches a peak briefly following the target onset (in the window centered at 200 [500] ms following target onset in LIP [PRR]), and vanishes shortly prior to a movement. Quantified in the same interval as previously (Fig. 4C), the effect in LIP (PRR) is a −4.7 (−4.4) sp/s decrease in neuronal activity over the range of the reward size, which is significant in both areas (LIP: $P < 0.0001$, $t_{59} = -4.4$; PRR: $P < 0.0001$, $t_{64} = -6.8$). There is no significant difference between the effect in the 2 areas ($P > 0.83$, $t$-test).

Prior studies have shown an effect of previous reward size that is linked to choice effects in the current trial (Platt and Glimcher 1999; Dorris and Glimcher 2004; Sugrue et al. 2004). We tested whether this previous result might explain the current data. In these studies, a larger reward on the previous trial leads to higher firing rates when, on the current trial, a target inside the RF is chosen, and leads to lower firing rates when a target outside the RF is chosen. We tested whether this prior result might explain the current data. Based on these prior studies, we would expect different results for stay trials in which both choices are inside the RF and stay trials in which both choices are outside the RF. When both choices are outside the RF, firing should vary inversely with the previous reward size. This is presumably because a large reward obtained for an out of RF choice increases the probability of repeating that choice (Fig. 2A), and a high probability of making an out of RF choice is associated with a decrease in activity (Dorris and Glimcher 2004; Sugrue et al. 2004). In fact, this was the case (Fig. 5, right column; mean, −3.9 sp/s, $P < 0.0001$, $t_{124} = -5.9$). In contrast, when both choices are in the RF, the prior studies suggest that firing should vary in proportion to the previous reward size. This is presumably because a large reward obtained for a RF choice increases the probability of repeating that choice (Fig. 2A), and a high probability of making a RF choice is associated with an increase in activity (Platt and Glimcher 1999; Dorris and Glimcher 2004; Sugrue et al. 2004). We instead observed a decrease (Fig. 5, left column; mean, −3.7 sp/s, $P < 0.0001$, $t_{124} = -5.3$). Thus, the effects of the previous reward size on subsequent firing rate cannot be explained by the prior report of value-related effects. This is not to say that the prior effect did not also occur. An analysis further below show that the two effects are independent of each other.

We also separately analyzed the effect of reward size in trials in which the animals switched choices, separately for switches into the RF and switches out of the RF (Fig. 5C, lower plots). Because switches occurred only in about a third of trials, this analysis may have limited power. Nonetheless, the same strong effect of reward size was observed for switches into the RF (lower left; mean, −6.3 sp/s, $P = 0.0027$, $t_{124} = -3.1$). The effect for switches out of the RF was half as large and in the opposite direction (mean, 3.5 sp/s, $P = 0.018$, $t_{124} = 2.4$). This likely reflects, at least in part, the coding of the reward-based quantities reported previously (Platt and Glimcher 1999; Dorris and Glimcher 2004; Sugrue et al. 2004). In this framework, a switch out of the RF is likely to occur when the RF reward is small and so when an animal expects that the out of RF target could be richer. A previous study (Sugrue et al. 2004) showed that when animals expect a large reward for a choice out of the RF, LIP neurons strongly suppress their activity. This way, a small reward obtained for the RF choice could lead to a strong suppression, i.e., the reward size could show a positive correlation with firing. This effect may partially override the effect of the reward size in this case.

We summarize the contribution of the individual variables considered above in a multiple regression analysis. In this analysis, the individual variables, and their interactions, figure in a linear regression model on the neuronal activity. The first factor was the primary effect reported in this study, "reward size". The second factor is a reward-based decision variable, the "RF target desirability" (Dorris and Glimcher 2004) (Materials and Methods). This variable has been found to positively correlate with parietal activity in previous studies (Platt and Glimcher 1999; Roesch and Olson 2003; Dorris and Glimcher 2004; Sugrue et al. 2004, 2005; Kubanek and Snyder 2015b). In our study, the Pearson's correlation between desirability and the animals' binary choices (left or right target) in each trial is $r = 0.56 \pm 0.07$ (mean ± SD, 125 sessions). Third, the binary variable whether an animal will select the same option of switch options ("switch") could have leverage on the neural activity. Fourth, there may be a neural effect of whether the animals were to chose the RF or the antiRF target (Fig. 5). We therefore considered the binary variable "choice" as an additional factor. Fifth, the neural effect could in part be explained by the animals' RT, because the reward magnitude positively correlated with the RT (Fig. 2C), and because parietal movement planning activity can decrease with increasing RT (Snyder et al. 2006). We included these factors and the important interactions (reward size × switch, reward size × choice, switch × choice) as additional regressors on the discharge rates of each cell on each trial. As in the previous analyses, the discharge rates were measured in the interval from the target onset to the go cue onset.

Figure 6 shows the weights of the individual factors in this linear model, computed separately for each cell. This analysis reveals that the effect of reward size (blue, mean weight, −2.3 sp/s) is significant ($P = 0.0021$, $t_{124} = -3.2$, two-sided $t$-test) even after accounting for multiple additional factors. This effect therefore cannot be simply explained by these additional factors.

The largest effect was due to choice (brown; −10.6 sp/s, $P < 0.0001$, $t_{124} = -15.4$). This effect is not surprising given that we specifically recorded from neurons that showed higher firing rates for movements directed into the RF compared with movement directed out of the RF (Materials and Methods).

Previous studies have shown that parietal neurons increase their activity with high reward (desirability) associated with the target placed in the RF (Platt and Glimcher 1999; Roesch and Olson 2003; Dorris and Glimcher 2004; Musallam et al. 2004; Sugrue et al. 2004, 2005). We replicated this finding (green bar in Fig. 6, +2.3 sp/s, $P < 0.001$, $t_{124} = 3.6$). The finding that both the effect of reward size and RF target desirability have significant leverage on the neural signal suggests that these factors describe independent phenomena. The idea that the effects of reward size and RF target desirability are independent phenomena is further supported by the finding that these factors point in the opposite direction, and the fact that the effect of reward size remains largely unchanged even if RF target desirability (green bar) is excluded from the regression (−3.6 sp/s, $P < 0.0001$, $t_{124} = -5.5$).

The effects of choice and desirability are consistent with our previous findings (Kubanek and Snyder 2015b), in which the effects of the "reward size" and its cross terms were not analyzed.

Furthermore, there was a significant interaction between the reward size and choice (−3.3 sp/s, $P < 0.001$, $t_{124} = -3.6$). This interaction captures the previous observation that reward size is encoded differentially for choices of the RF target and choices of the antiRF target (Fig. 5).

All the other factors or interactions considered in this analysis were nonsignificant ($P > 0.05$). Of particular interests is the lack of an effect ($P = 0.42$, $t_{124} = 0.8$) of switch, the binary variable indicating whether an animal was going to switch its choice or not. While the reward size and the tendency to switch are
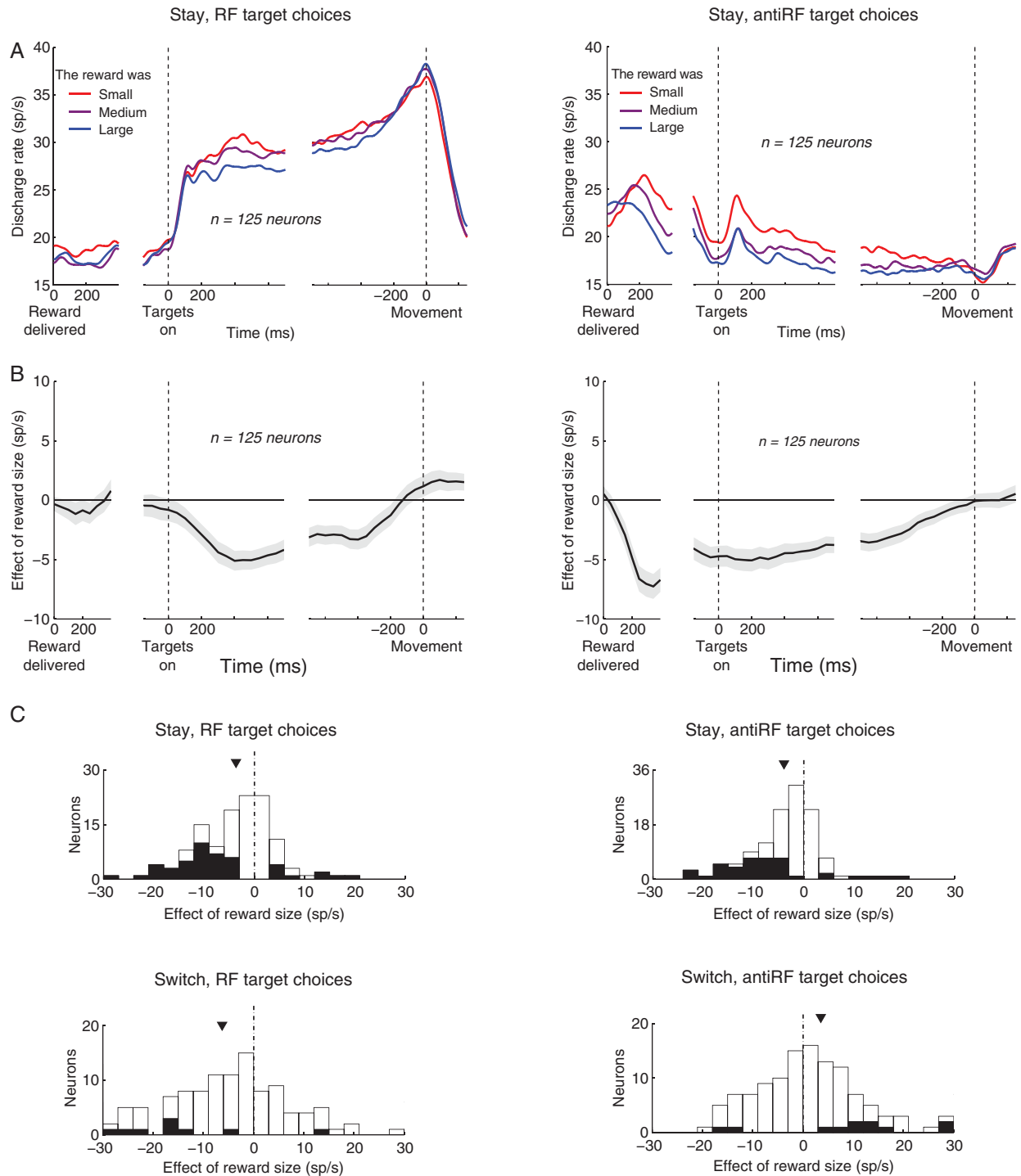
**Figure 5.** The effect of reward size during stay and switch trials. Same analyses and format as in Figure 4, for all neurons, all repeat (stay) trials, and separately for choices of the RF target (left column) and the antiRF target (right column). In C, effect for the individual neurons is shown also for switch trials.

confounded with one another (Fig. 2A), the correlation is only $r = -0.41$, so that the regression analysis can separate out the effects of these two factors. From this, we can conclude that it is the reward size—and not the plan to switch—that is encoded in the neural activity.

We validated the independence of desirability and reward size (Fig. 6) also at the behavioral level. To do so, we regressed the RF target desirability and the reward size on the binary variable whether the animals were going to switch choice or not. In this analysis, the normalized RF target desirability (value between 0 and 1) was re-formulated, without changing its explanatory power, such that instead of explaining the animals' choices (RF versus antiRF option), it explained whether the animal was going to stay or switch (i.e., it explained animals' choices relative to the previous trial). The weights assigned to the two factors in this regression were highly significant for both factors
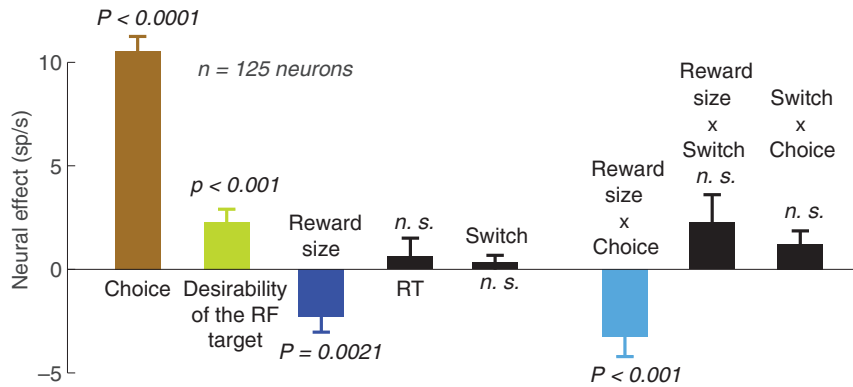
**Figure 6.** Contribution of individual variables to parietal activity. The figure shows the mean weights associated with the considered variables and their interactions in a linear regression on the neural activity. The regression was performed separately for each individual cell; weights averaged over the cells are shown. As in the previous analyses, the neural activity was measured in the interval from the target onset to the go cue onset. The values of all variables were normalized between 0 and 1 so that the weights magnitudes ("neural effects") are directly comparable. The significance values indicate the significance of the t-tests on the weights over the 125 neurons. n.s., $P > 0.05$.

(desirability:

$P < 0.0001$, $t_{124} = 70.7$; reward size: $P < 0.0001$, $t_{124} = -5.7$). Thus, RF target desirability and reward size are independent also at the behavioral level. It is possible that other forms of RF target desirability might influence this particular result. Nonetheless, notably, RF target desirability and reward size are distinct at the fundamental behavioral level: RF target desirability quantifies animals' tendency to choose the RF option or the opposite option, whereas reward size is a relativistic variable, informing the animal whether it should stay or switch. This simple decision variable might be of utility in foraging environments in which the reward associated with each option is difficult to predict.

Finally, we investigated whether the effect of reward size is a function of the effector (saccade, reach) with which the animals made a choice. We first investigated the behavioral effects shown in Figure 2A,C, separately for each effector. The slope of the relationship between the binary variable indicating whether a switch occurred and the normalized reward (Fig. 7A) was −0.79 and −0.82 for saccade and reach trials, respectively, and the slopes were highly significant ($P < 0.0001$). The slopes of the relationship between the RT and the normalized reward (Fig. 7B) were 27.3 and 76.1 ms per the reward range for saccade and reach trials, respectively, and the slopes were highly significant ($P < 0.0001$). Figure 7C shows the neural effects, separately for each effector. The figure demonstrates that both LIP and PRR encode the reward size during both plans to make a saccade and a reach (LIP saccades: mean −5.2 sp/s, $P < 0.0001$, $t_{59} = -4.6$; LIP reaches: −4.1 sp/s, $P < 0.001$, $t_{59} = -3.7$; PRR saccades: −2.2 sp/s, $P = 0.0018$, $t_{64} = -3.3$; PRR reaches: −6.8 sp/s, $P < 0.0001$, $t_{64} = -7.2$). Interestingly, nonetheless, the effect in PRR was significantly larger for reaches compared with saccades (mean difference, 4.6 sp/s) and this difference was highly significant ($P < 0.0001$, $t_{124} = 4.6$). The finding that the effect of reward size is larger on reach trials compared with saccade trials in PRR suggests that the effect of reward size is not a general response to the received reward, but instead has to do with a particular behavior and a particular parietal area.

## Discussion

Deciding whether to keep with the current option or switch to an alternative based on the previous outcome is an important behavioral capability in both the animal kingdom and in human settings. Surprisingly, little is known about the neural

mechanisms underlying this basic behavioral drive. Many previous studies used relatively complex reward-based decision tasks in which they concentrated on modeling the expected value of the available options (Rescorla and Wagner 1972; Sutton and Barto 1998a; Platt and Glimcher 1999; Dorris and Glimcher 2004; Sugrue et al. 2004, 2005; Samejima et al. 2005; Niv 2009; Kubanek and Snyder 2015b). The rich task structures and the model complexities somewhat complicate a focused study of the neural effects of each individual outcome.

Nonetheless, a few studies asked the simpler question of how each individual outcome contributes to the immediately following switching behavior and neural activity. A pioneering study in this direction was the study of Shima and Tanji (1998). In this study, monkeys selected one of two hand movements based on the amount of the immediately preceding reward. In this task, the monkeys kept choosing the same movement until the reward delivered for that movement was reduced; a trial or two following this reduction, the animals switched to produce the alternative movement. It was found, akin to what we report here, that a reduction of the reward induced a substantial increase of the activity of neurons in the regions of interest, the CMAs. In the rostral parts of the CMA (CMr), this effect was for many cells contingent upon switching the choice on that trial. Ventral parts of the CMA predominantly showed signals governing whether animals were going to switch or not. Silencing the CMr with muscimol impaired the proper switching behavior based on the experienced reward. It therefore appears that CMA is critical in properly parsing the information about the reward outcomes, for the purpose of making a proper choice leading to a higher reward.

In line with that study, in our study, obtaining a small reward resulted in a tendency to switch, and in an increase of neural activity. However, in parietal cortex, compared with CMA, we did not find signals related to the actual plan to switch (Fig. 6, "switch"). Thus, it appears that parietal cortex only reflects the reward-based decision variable guiding the decision whether to stay or switch. The command that dictates whether to stay or switch appears to be present in other brain regions. The regions in which such signals have been reported include the CMA (Shima and Tanji 1998), the pre-supplementary motor area (pre-SMA) (Matsuzaka and Tanji 1996; Isoda and Hikosaka 2007), and the anterior cingulate cortex (Quilodran et al. 2008). Neurons in these areas collectively show elevated activity on switch relative to stay trials.
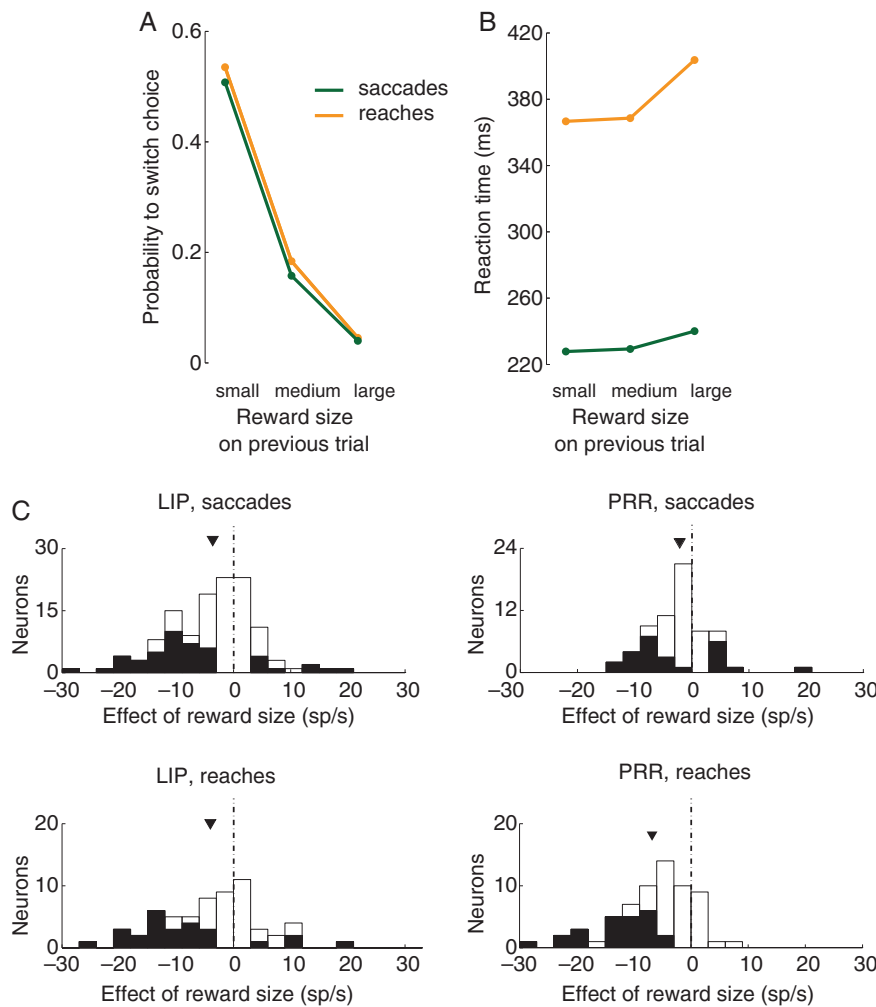
**Figure 7.** The effect of reward size as a function of the choice effector. (*A*,*B*) Same format as in Figure 2*A*,*C*, separately for each effector. (*C*) Same format as in Figure 4*C*, separately for each effector.

Furthermore, electrical microstimulation of the posterior cingulate and the pre-SMA (Isoda and Hikosaka 2007; Hayden et al. 2008) increases the likelihood of switching, therefore suggesting a casual role of these regions in switching behavior. Together, this evidence suggests that regions associated with the cingulate cortex, including the posterior cingulate and the CMAs, are important in mediating the decision between retaining the current choice or switching, based on the reward outcome. Given that ventral posterior cingulate cortex is reciprocally connected with the caudal part of the posterior parietal lobe (Cavada and Goldman-Rakic 1989a,b), it is likely that the effect that we report is of the same kind.

An analysis of a previous study indicates that reward size decreases neuronal activity also in frontal cortex (Seo and Lee 2009). In that study, gains induced choice repetition while losses induced choice avoidance (switch). Out of the 138 neurons pooled over dorsomedial frontal cortex (DMFC), dorsolateral prefrontal cortex (DLPFC), and dorsal anterior cingulate cortex that significantly encoded gain, 82 (59.4%) significantly decreased their firing following a gain (Table 2). This proportion is significantly different from 50% ($P = 0.033$, two-sided proportion test with Yates' correction). The effect was particularly evident in DMFC and DLPFC: in these 2 regions, 61 out of 91 (67%) of neurons that significantly encoded gain decreased their activity following a gain (and this proportion is significantly different from 50%, $P = 0.0017$).

We discuss specific candidate mechanisms that might underlie the effect.

One of the prominent engagements of parietal neurons are processes related to spatial attention or intention (Colby et al. 1996; Snyder et al. 1997; Gottlieb et al. 1998; Bisley and Goldberg 2003; Goldberg et al. 2006; Liu et al. 2010; Kubanek, Li, et al. 2015). The effect of reward size we report could therefore conceivably reflect an allocation of neuronal resources. In particular, high reward could increase attention or promote an intention related to a given target, and this increase in attention or intention could result in turn in an increase in the activity related to that target (Maunsell 2004; Schultz 2006). We observe instead a *decrease* of neuronal activity following high reward, regardless of whether the animal is about to choose the RF or the antiRF target (Fig. 5). Thus, it is difficult to explain the effect that we see as mediated by either attention or intention.

Previous studies showed that midbrain dopamine neurons signal the discrepancy between actually obtained and expected reward, the so-called reward prediction error (RPE) (Schultz et al. 1997; Bayer and Glimcher 2005). Our effects cannot be explained by a RPE effect. This is because the RPE (obtained reward minus predicted reward) increases with increasing obtained reward. Consequently, if the neurons in our task encoded RPE,

they would increase their activity with increasing obtained reward. We instead observe a decrease.

An alternative is that the effect of reward size might reflect an action mobilization. Poor rewards may alert the brain that an income is low, and that bodily and neural resources should be mobilized to seek richer reward. This process might increase firing rates in many regions of the brain. In light of previous studies, poor rewards are sensed by neurons in the limbic cortex (Shima and Tanji 1998; Hayden et al. 2008). These structures in turn activate motor-associated regions (Matsuzaka and Tanji 1996; Shima and Tanji 1998; Isoda and Hikosaka 2007), presumably with the goal to drive exploration for richer resources, or, at least, to change the current train of action (Matsuzaka and Tanji 1996; Shima and Tanji 1998; Isoda and Hikosaka 2007). The reward-size-sensitive signal that we report here in parietal cortex may reflect the reward-size-sensitive signals originating in the limbic cortex (Shima and Tanji 1998; Hayden et al. 2008), especially given that parietal cortex is tightly connected with the limbic regions (Cavada and Goldman-Rakic 1989a,b).

The effect of reward size cannot be explained by a modulation of an animal's general arousal level, because at least in PRR, the effect of reward size is a function of the effector with which a choice is to be made (Fig. 7). This supports the proposition that the effect has to do with a mobilization of action-related machinery with the goal to restore a favorable reward income.

With regard to effector specificity, it has been previously found that the coding of desirability in LIP is stronger for saccades than reaches, while in PRR it is essentially reach-specific (Kubanek and Snyder 2015b). The effects of previous size reported here show similar trends of effector selectivity, although the saccade preference in LIP in regard to this variable is relatively weak. This supports a notion that choice-related signals in LIP are partially saccade-specific and partially general, whereas PRR signals are more strongly embodied (de Lafuente et al. 2015; Kubanek, Li, et al. 2015). However, the findings of relatively weak saccade preference in LIP must be taken with care. In these studies (de Lafuente et al. 2015; Kubanek, Li, et al. 2015) animals were trained to reach to a target without looking at it. Despite the training, it is likely that animals planned (but did not execute) a covert eye movement to the target of an ensuing reach. As evidence of this idea, LIP activity becomes substantially more effector-specific when animals are explicitly required to plan an eye movement away from the arm movement target (Snyder et al. 1997). Thus it is highly likely that a covert eye movement to the reach target was planned in reach trials of the current study, as well as in de Lafuente et al. 2015. This limits the extent to which these studies provide an upper bound to the degree of effector specificity in LIP. (This issue does not apply to PRR, which is only weakly modulated by saccadic choices [de Lafuente et al. 2015; Kubanek, Li, et al. 2015] and responds almost identically to arm movements with or without an oppositely directed eye movement [Snyder et al. 1997]). The idea that LIP is in fact strongly effector-specific is reinforced by the fact that pharmacological inactivations of LIP elicit strictly saccade-specific deficits in choice behavior (Kubanek, Li, et al. 2015).

Our finding of an effect of reward size on the previous trial (Fig. 6, blue bar) does not supersede previous findings regarding how reward expectations (desirability of the receptive field, green bar) affect parietal activity (Platt and Glimcher 1999; Dorris and Glimcher 2004; Sugrue et al. 2004). Indeed, our neural effect of expected reward replicates these previous findings (Platt and Glimcher 1999; Dorris and Glimcher 2004; Sugrue et al. 2004; Kubanek and Snyder 2015b). However, the 2 effects were found to be independent, and even to have opposite signs (Fig. 6). In the raw data, this is apparent in Figure 5A, in which a choice of

the RF target, following a large reward from that target on the previous trial, results in a *decrease* of neuronal activity. Thus, the reported effect of reward size is a novel, independent variable that modulates neuronal activity in a reward-based choice task. Since the reward size in our task strongly guides the decision between staying and switching (Fig. 2A), this variable can be thought of as a decision variable pertaining to the decision between staying and switching. Parietal neurons strongly encode this decision variable.

The finding that a reward of a small size leads to an increase in LIP (and PRR) activity must be reconciled with the findings that LIP neurons increase their firing with expected value (Platt and Glimcher 1999; Dorris and Glimcher 2004; Sugrue et al. 2004) as well as anticipated punishment (Leathers and Olson 2012). A critical difference is that these previous studies described the influence of the anticipated outcome of the current trial on parietal activity, while our new finding describes the influence of the previous trial's outcome. These are two orthogonal influences on firing rate. Specifically, our study shows that the effect of the anticipated outcome of the current trial (the effect of desirability) is accompanied by a strong effect of the outcome of the previous trial. These effects may point in the same or in the opposite direction. For example, if the animal receives a small reward and repeats its choice, the small reward devaluates the targets' current desirability, and so the effects of desirability and the size of the previously obtained reward will influence firing in opposite directions. However, a small reward that results in a switch with the goal to attain a richer target will result in the two effects influencing firing in the same direction.

The finding of a new, independent reward-related signal in parietal cortex has important implications for future investigations into value-related signals in the brain. Value-related signals reported in parietal cortex previously pertain to the decision to direct a saccade or a reach to a particular visual target (Platt and Glimcher 1999; Dorris and Glimcher 2004; Sugrue et al. 2004). In contrast, the reward-size-related signal reported here pertains to a relativistic decision—the decision between retaining the current choice or switching. The relative weights of the two kinds of signals are likely to be a function of the particular brain region and the particular task. For instance, it is conceivable that in a difficult or a new task, an animal might adopt a win-stay lose-shift strategy (Worthy et al. 2013). In such scenarios, a relativistic approach to guiding choice might generate neural signals of the sort reported here and in the limbic system.

In summary, we found that poor outcomes drove the tendency to change choice and resulted in an increase in parietal activity. Favorable outcomes produced choice repetition and relatively suppressed the neuronal activity. Thus, reward size is a variable that guides the decision between repeating the previous and producing an alternative choice, and this decision variable is strongly encoded in the activity of parietal neurons. The effect is present also in frontal regions of the brain based on an examination of a previous study. These novel value-related signals in the parietal and frontal regions likely reflect similar effects reported in the limbic structures (Shima and Tanji 1998; Hayden et al. 2008; Quilodran et al. 2008) that have been shown to be crucial for mediating proper repeat-switch decisions based on recently obtained reward.

## Authors' Contributions

J.K. and L.H.S. developed the concept. J.K. and L.H.S. designed the task. J.K. trained the animals, collected the data, and analyzed the data. J.K. and L.H.S. wrote the paper.

## Funding

## Notes

## References

Abe H, Lee D. 2011. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. Neuron. 70:731–741.

Bayer HM, Glimcher PW. 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron. 47:129–141.

Bisley JW, Goldberg ME. 2003. Neuronal activity in the lateral intraparietal area and spatial attention. Science. 299:81–86.

Calton JL, Dickinson AR, Snyder LH. 2002. Non-spatial, motor-specific activation in posterior parietal cortex. Nat Neurosci. 5:580–588.

Cavada C, Goldman-Rakic PS. 1989a. Posterior parietal cortex in rhesus monkey: I. parcellation of areas based on distinctive limbic and sensory corticocortical connections. J Comp Neurol. 287:393–421.

Cavada C, Goldman-Rakic PS. 1989b. Posterior parietal cortex in rhesus monkey: Ii. evidence for segregated corticocortical networks linking sensory and limbic areas with the frontal lobe. J Comp Neurol. 287:422–445.

Colby CL, Duhamel JR, Goldberg ME. 1996. Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. J Neurophysiol. 76:2841–2852.

Cromwell HC, Schultz W. 2003. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. J Neurophysiol. 89:2823–2838.

Cui H, Andersen RA. 2007. Posterior parietal cortex encodes autonomously selected motor plans. Neuron. 56:552–559.

de Lafuente V, Jazayeri M, Shadlen MN. 2015. Representation of accumulating evidence for a decision in two parietal areas. J Neurosci. 35:4306–4318.

Dickinson AR, Calton JL, Snyder LH. 2003. Nonspatial saccade-specific activation in area lip of monkey parietal cortex. J Neurophysiol. 90:2460–2464.

Dorris MC, Glimcher PW. 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. Neuron. 44:365–378.

Goldberg ME, Bisley JW, Powell KD, Gottlieb J. 2006. Saccades, salience and attention: the role of the lateral intraparietal area in visual behavior. Prog Brain Res. 155:157–175.

Gottlieb JP, Kusunoki M, Goldberg ME. 1998. The representation of visual salience in monkey parietal cortex. Nature. 391:481–484.

Hassani OK, Cromwell HC, Schultz W. 2001. Influence of expectation of different rewards on behavior- related neuronal activity in the striatum. J Neurophysiol. 85:2477–2489.

Hayden BY, Nair AC, McCoy AN, Platt ML. 2008. Posterior cingulate cortex mediates outcome- contingent allocation of behavior. Neuron. 60:19–25.

Hollerman JR, Tremblay L, Schultz W. 1998. Influence of reward expectation on behavior-related neuronal activity in primate striatum. J Neurophysiol. 80:947–963.

Isoda M, Hikosaka O. 2007. Switching from automatic to controlled action by monkey medial frontal cortex. Nat Neurosci. 10:240–248.

Kable JW, Glimcher PW. 2009. The neurobiology of decision: consensus and controversy. Neuron. 63:733–745.

Kennerley SW, Behrens TE, Wallis JD. 2011. Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. Nat Neurosci. 14:1581–1589.

Kubanek J, Li JM, Snyder LH. 2015. Motor role of parietal cortex in a monkey model of hemispatial neglect. Proc Natl Acad Sci. 112: E2067–E2072.

Kubanek J, Snyder LH. 2015a. Matching behavior as a tradeoff between reward maximization and demands on neural computation. F1000Research. 4:147–157.

Kubanek J, Snyder LH. 2015b. Reward-based decision signals in parietal cortex are partially embodied. J Neurosci. 35:4869–4881.

Kubanek J, Snyder LH, Abrams RA. 2015. Reward and punishment act as distinct factors in guiding behavior. Cognition. 139:154–167.

Kubanek J, Wang C, Snyder LH. 2013. Neuronal responses to target onset in oculomotor and somatomotor parietal circuits differ markedly in a choice task. J Neurophysiol. 110:2247–2256.

Lau B, Glimcher PW. 2008. Value representations in the primate striatum during matching behavior. Neuron. 58:451–463.

Leathers ML, Olson CR. 2012. In monkeys making value-based decisions, lip neurons encode cue salience and not action value. Science. 338:132–135.

Leon MI, Gallistel C. 1998. Self-stimulating rats combine subjective reward magnitude and subjective reward rate multiplicatively. J Exp Psychol Anim Behav Process. 24:265.

Lerman DC, Vorndran CM. 2002. On the status of knowledge for using punishment: Implications for treating behavior disorders. J Appl Behav Anal. 35:431–464.

Liu Y, Yttri EA, Snyder LH. 2010. Intention and attention: different functional roles for lipd and lipv. Nat Neurosci. 13(4):495–500.

Matsuzaka Y, Tanji J. 1996. Changing directions of forthcoming arm movements: neuronal activity in the presupplementary and supplementary motor area of monkey cerebral cortex. J Neurophysiol. 76:2327–2342.

Maunsell JHR. 2004. Neuronal representations of cognitive state: reward or attention? Trends Cogn Sci (Regul Ed). 8:261–265.

Musallam S, Corneil BD, Greger B, Scherberger H, Andersen RA. 2004. Cognitive control signals for neural prosthetics. Science. 305:258–262.

Neuringer AJ. 1967. Effects of reinforcement magnitude on choice and rate of responding. J Exp Anal Behav. 10:417–424.

Niv Y. 2009. Reinforcement learning in the brain. J Math Psychol. 53:139–154.

Platt M, Glimcher P. 1999. Neural correlates of decision variables in parietal cortex. Nature. 400:233–238.

Quilodran R, Rothé M, Procyk E. 2008. Behavioral shifts and action valuation in the anterior cingulate cortex. Neuron. 57:314–325.

Rescorla RA, Wagner AR. 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. Classical Conditioning II: Current Research and Theory. 2:64–99.

Roesch MR, Olson CR. 2003. Impact of expected reward on neuronal activity in prefrontal cortex, frontal and supplementary eye fields and premotor cortex. J Neurophysiol. 90:1766–1789.

Roesch MR, Olson CR. 2004. Neuronal activity related to reward value and motivation in primate frontal cortex. Science. 304:307–310.

Samejima K, Ueda Y, Doya K, Kimura M. 2005. Representation of action-specific reward values in the striatum. Science. 310:1337–1340.

Schultz W. 2006. Behavioral theories and the neurophysiology of reward. Annu Rev Psychol. 57:87–115.

Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. Science. 275:1593–1599.

Seo H, Barraclough DJ, Lee D. 2007. Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. Cereb Cortex. 17(Suppl 1):i110–i117.

Seo H, Barraclough DJ, Lee D. 2009. Lateral intraparietal cortex and reinforcement learning during a mixed-strategy game. J Neurosci. 29:7278–7289.

Seo H, Lee D. 2009. Behavioral and neural changes after gains and losses of conditioned reinforcers. J Neurosci. 29:3627–3641.

Shima K, Tanji J. 1998. Role for cingulate motor area cells in voluntary movement selection based on reward. Science. 282:1335–1338.

Skinner BF. 1953. Science and human behavior. New York: Simon and Schuster.

Skinner BF. 1963. Operant behavior. Am Psychol. 18:503.

Snyder LH, Batista AP, Andersen RA. 1997. Coding of intention in the posterior parietal cortex. Nature. 386:167–170.

Snyder LH, Dickinson AR, Calton JL. 2006. Preparatory delay activity in the monkey parietal reach region predicts reach reaction times. J Neurosci. 26:10091–9.

So N, Stuphorn V. 2012. Supplementary eye field encodes reward prediction error. J Neurosci Nurs. 32:2950–2963.

Stillings NA, Allen GA, Estes W. 1968. Reaction time as a function of noncontingent reward magnitude. Psychon Sci. 10:337–338.

Strait CE, Blanchard TC, Hayden BY. 2014. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. Neuron. 82:1357–1366.

Sugrue LP, Corrado GS, Newsome WT. 2004. Matching behavior and the representation of value in the parietal cortex. Science. 304:1782–1787.

Sugrue LP, Corrado GS, Newsome WT. 2005. Choosing the greater of two goods: neural currencies for valuation and decision making. Nat Rev Neurosci. 6:363–375.

Sutton RS, Barto AG. 1998a. Introduction to reinforcement learning. Cambridge: MIT Press.

Sutton RS, Barto AG. 1998b. Reinforcement learning: an introduction. IEEE Trans Neural Netw. 9:1054.

Thorndike EL. 1898. Animal intelligence: An experimental study of the associative processes in animals. Psychological Monographs: General and Applied. 2:1–109.

Thorndike EL. 1911. Animal intelligence: Experimental studies. London: Macmillan.

Thorndike EL. 1927. The law of effect. Am J Psychol. 39(1/4): 212–222.

Tremblay L, Schultz W. 2000. Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex. J Neurophysiol. 83:1864–1876.

Tversky A, Kahneman D. 1986. Rational choice and the framing of decisions. J Bus. 59(4):S251–S278.

Worthy DA, Hawthorne MJ, Otto AR. 2013. Heterogeneity of strategy use in the iowa gambling task: a comparison of win-stay/lose-shift and reinforcement learning models. Psychon Bull Rev. 20:364–371.