J Neurophysiol 107: 3246–3255, 2012.
First published March 14, 2012; doi:10.1152/jn.00690.2010.

Orbitofrontal cortical activity during repeated free choice

Michael Campos, Kari Koppitch, Richard A. Andersen, and Shinsuke Shimojo
Division of Biology, Computation and Neural Systems, California Institute of Technology, Pasadena, California

Submitted 4 August 2010; accepted in final form 12 March 2012

Campos M, Koppitch K, Andersen RA, Shimojo S. Orbitofrontal cortical activity during repeated free choice. *J Neurophysiol* 107: 3246–3255, 2012. First published March 14, 2012; doi:10.1152/jn.00690.2010.— Neurons in the orbitofrontal cortex (OFC) have been shown to encode subjective values, suggesting a role in preference-based decision-making, although the precise relation to choice behavior is unclear. In a repeated two-choice task, subjective values of each choice can account for aggregate choice behavior, which is the overall likelihood of choosing one option over the other. Individual choices, however, are impossible to predict with knowledge of relative subjective values alone. In this study we investigated the role of internal factors in choice behavior with a simple but novel free-choice task and simultaneous recording from individual neurons in nonhuman primate OFC. We found that, first, the observed sequences of choice behavior included periods of exceptionally long runs of each of two available options and periods of frequent switching. Neither a satiety-based mechanism nor a random selection process could explain the observed choice behavior. Second, OFC neurons encode important features of the choice behavior. These features include activity selective for exceptionally long runs of a given choice (stay selectivity) as well as activity selective for switches between choices (switch selectivity). These results suggest that OFC neural activity, in addition to encoding subjective values on a long timescale that is sensitive to satiety, also encodes a signal that fluctuates on a shorter timescale and thereby reflects some of the statistically improbable aspects of free-choice behavior.

action selection; frontal cortex; reward; preferences; monkey

DECISION-MAKING IS an inherently subjective process. To understand decision-making, experiments can be devised that systematically vary the inputs to a decision (as in perceptual judgments) or that vary the expectation of the resulting rewards (as in economic choice). With these approaches a neuroscientist can characterize the neural correlates of the experimentally manipulated variables and delineate the progression from sensory input to observed choice behavior (Montague and Berns 2002). In the present study we take a complementary approach.

We developed a task that focuses exclusively on internal factors in decision-making. Nonhuman primate subjects repeatedly chose to press one of two buttons that were each associated with a different liquid reward. The reward magnitudes were kept constant throughout the experiment, and there were no experimental visual stimuli. In this simple but novel repeated free-choice paradigm, influences on decision-making could be unambiguously attributed to internal factors. Of particular interest were the monkeys' decisions to switch between rewards or to stay with the same reward for many consecutive choices. Because the switching behavior was internally motivated, our approach also contrasts with conventional reversal learning tasks (Isoda and Hikosaka 2007, 2008; Walton et al. 2004).

Neurons in the monkey orbitofrontal cortex (OFC) are known to encode the value of reward options (Kennerley et al. 2009; Padoa-Schioppa and Assad 2006; Roesch et al. 2007; Tremblay and Schultz 1999; Wallis and Miller 2003), but the role of the OFC in a repeated free-choice task is not clear (Walton et al. 2004). The subjective value interpretation of OFC activity implicitly assumes that the firing rates associated with a fixed amount of reward will remain constant or decrease monotonically over the course of an experimental session. Given that OFC also plays a role in switching between behaviors (Bechara et al. 2000; Fellows and Farah 2005; Kepecs et al. 2008), we hypothesized that at least some neurons might also fluctuate during the course of the session, in a manner that reflects the many individual switch and stay choices that give rise to matching-law behavior in aggregate. Here we characterized the activity of individual OFC neurons in a repeated free-choice task. We found that many of the OFC neurons reflected statistically improbable aspects of free choice, such as decisions to repeatedly choose one of the available rewards for long runs.

MATERIALS AND METHODS

Two monkeys (*Macaca mulatta*) participated in this study. The monkeys sat in a custom-designed chair in a dimly lit experimental room. A double-line juice tube that delivered apple juice or water was placed directly in front of the mouth. A two-button interface (McMaster Carr, Chicago, IL) was placed ~30 cm in front of the abdomen, arranged parallel to the axis of the shoulders, and spaced 3 cm apart from each other. Each button was associated with a particular liquid (apple juice or water), such that when the monkey pressed one button the associated liquid was delivered immediately via the juice tube. After each reward delivery (lasting 300 ms), the buttons were deactivated for 2.5 s. Button-reward mappings were stable for an entire liquid-reward period (10–25 min), and no stimuli indicated which reward was associated with each button. Between juice-reward periods (also 10–25 min) the monkeys could press the same buttons to watch short video clips. Button-reward mappings were reversed in successive liquid-reward periods.

To assess whether choice behavior could be considered "random," the observed series of choices were segmented into runs of the same choice and analyzed in terms of the distributions of run lengths for each liquid reward (Clauset et al. 2009; Eden and Kramer 2010). A linear-nonlinear-Poisson model describes a family of choice behavior models in which the choice selection step is a Poisson process—essentially a random selection between options that might have unequal likelihoods of selection (Corrado et al. 2005). If the random choice process is repeated multiple times, the distribution of run lengths will be a Poisson distribution. We tested whether the run length distributions observed in this study could be fit by a Poisson distribution using probability bounds for the observed Fano factor (Eden and Kramer 2010). The Fano factor for the run length distri-

Address for reprint requests and other correspondence: M. Campos, Qualcomm Research, 5775 Morehouse Dr., San Diego, CA 92121 (e-mail: camposm@qualcomm.com).

butions observed was calculated as the variance of all observed run lengths divided by the mean. These were calculated separately for runs of each reward option, for each monkey, in both the main and control tasks (described below).

Because the run length distributions were not well fit by a Poisson process, we also tested whether a power-law distribution provided a better fit (Fig. 1). Power-law distributions are partly characterized by a "heavy tail" (Barabasi 2005). In the present study, a heavy tail in the run-length distributions corresponded to a large number of long run lengths. Frequency histograms of idealized power-law distributions are linear on a log-log plot, but empirical data are usually observed to follow a power law above some minimum value (Clauset 2009). We therefore used an iterative procedure to first identify a putative minimum value and then attempted to fit the data to a power law. Frequency histograms of the observed run lengths were computed and plotted on log-log axes (see Fig. 1). An inflection point within this log-log plot, above which the frequency histogram was approximately linear, was identified by visual inspection. The corresponding run length values were then entered into an open-source fitting algorithm as the "xmin" value, and P values were computed (**http://tuvalu.santafe.edu/~aaronc/powerlaws/**). P values above 0.1 were considered significant by a method of fitting data to power-law distributions that is described in detail by Clauset (2009).

Cylindrical (18-mm inner diameter) recording chambers were surgically placed above the left principal sulcus of both monkeys, normal to the skull surface, and then fixed within a dental acrylic headcap. The stereotaxic coordinates of the chambers were (40A, 10L) for *monkey M* and (36A, 7L) for *monkey S*. Two headposts were embedded into each headcap for head immobilization during the experimental sessions. Anatomical MRIs were taken postoperatively, with a high-contrast agent (gallidium) inside of the recording chamber revealing the exact orientation of the chamber with respect to anatomical landmarks beneath. Recording chambers were placed to yield access to the areas 13m and 11m/l of the left OFC, between the medial and lateral orbital sulci (Carmichael and Price 1996), where neurons in OFC were shown to encode subjective, context-dependent reward values (Padoa-Schioppa and Assad 2006; Tremblay and Schultz 1999), but slightly anterior to the secondary taste cortex region of OFC, where very little evidence of the encoding of motor-related variables was found (Wallis and Miller 2003).

Recording sessions were guided by the landmarks visible in the anatomical MRI. Neural recording were performed with a five-channel microelectrode microdrive (Thomas Recording). Each electrode was positioned within one of five linearly arranged guide tubes and attached to a unique rotating motor. The guide tubes were fixed to each other and manually lowered to penetrate the dura. Each electrode was then independently driven with software that controlled the five individual motors. The voltage traces were monitored with speakers and with a visual display (Plexon, Dallas, TX) to detect transitions from gray matter to white matter. Depending on the location of the guide tubes, the electrodes might have passed through one or both banks of the principal sulcus (or neither), then the white matter, and finally the OFC. The depth at which the OFC was encountered varied depending on the anterior-posterior axis, ranging from ~11 mm (anterior) to ~19 mm (posterior). Waveforms were sorted off-line (Plexon). All procedures were approved by the Caltech Institutional Animal Care and Use Committee.

Neural firing activity was analyzed in the two 1-s intervals immediately preceding and following each button press. Because of the imposed 2.5-s time-out period between rewards, there was also at least a 500-ms gap between the data analysis interval following one choice and the data analysis interval preceding the subsequent choice. The 1-s interval preceding the first button press of a run of water, for example, contained firing rate data from the 1-s interval just before the monkey's switch from apple juice to water was expressed behaviorally. We assume that the monkey chose to switch to water prior to this behavioral expression. The choice may have occurred, for example, while the monkey was evaluating the receipt of apple juice from the previous choice, or even earlier.

The firing activity of many neurons was observed to systematically vary over the course of each run, with peaks at the beginning, middle, or end of the run. The peak of firing activity was not always restricted to the same choice within the run. Instead, neurons that were active in the beginning of a run, for example, may have been active for the first few trials, perhaps with a higher firing rate in the second or third choice compared with the first. Similarly, some neurons were maximally active for more than one of the last choices of a run that

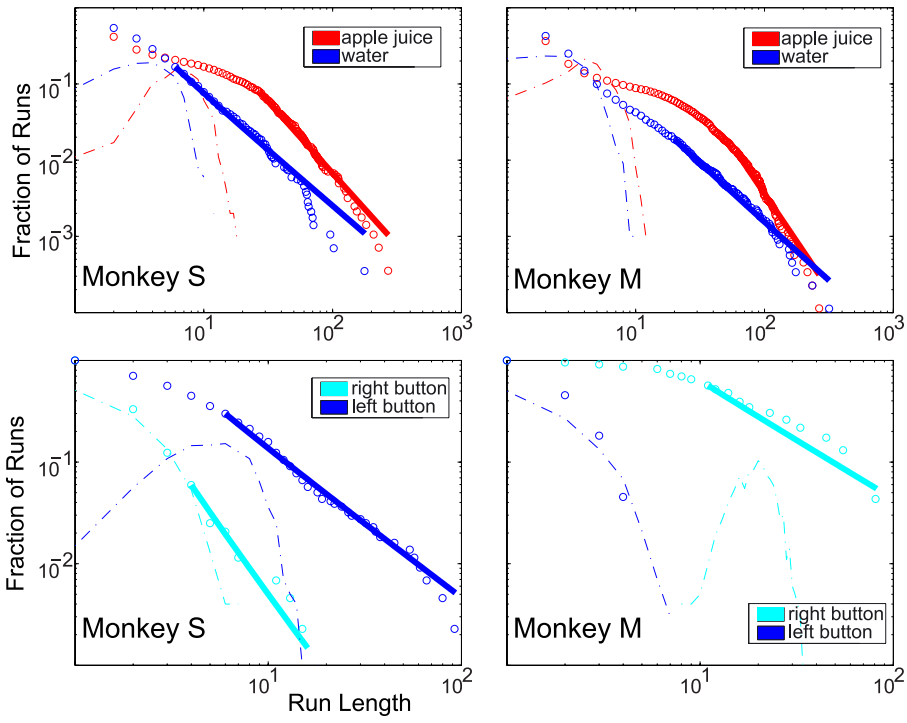

Fig. 1. Run length distributions. Frequency histogram of run lengths of each choice observed for each monkey over all recording sessions. *Top*: run length distributions from the main task (apple juice vs. water). Open circles indicate fraction of all runs (y-axis) observed at each run length (x-axis). Solid lines indicate power-law fit. Dot-dashed lines show expected curves of a Poisson process generated with the same average run length value. *Bottom*: run length distributions from the first control task (equal amounts of water reward associated with the right and left buttons).

preceded a switch. Many neurons reached a peak of activity in the middle of the run.

To quantitatively identify the peak of activity within runs, firing rates were binned according to the distance to switch choices. The first choice of a run was assigned a distance 1 from the switch event. The second and third choices were assigned distances 2 and 3. The firing rates were binned by the floor of the log (base 2) of these distances. This means that switch choices were assigned to the zeroth bin [log2(1) = 0]. The second and third choices of a run were assigned to the first bin [log2(2) = 1]. The fourth to seventh choices were assigned to the second bin [log2(4) = 2]. The eighth to fifteenth choices were assigned to the third bin [log2(8) = 3]. The choices that were sixteen or more from the previous switch event and more than eight from the next switch event were assigned to the fourth bin. Choices that were between four and seven choices before an upcoming switch were assigned to the fifth bin. Choices that were between two and three choices before an upcoming switch were assigned to the sixth bin. Choices that immediately preceded a switch were assigned to the seventh bin. In all, therefore, a choice could be assigned to one of eight bins (0 to 7). In cases in which a choice was at a distance of less than eight to both the preceding and next switch choices (which could occur if it belonged to a run of <16 consecutive choices), the choice was assigned to the bin associated with the shorter of these two distances. Choices that were equidistant to the preceding and next switch were assigned a distance relative to the preceding switch only. We excluded data from runs that were shorter than eight choices in length. This guaranteed that all data assigned to the second bin (choice distance 4–7 from the previous switch) were at least four choices removed from the subsequent switch, and, likewise, all choices assigned to the fifth bin were at least four choices removed from the previous switch. The decision to exclude runs of eight or less choices (as opposed to a higher threshold) represented a trade-off between statistical power and the ability to distinguish firing rate activity associated with the preceding versus subsequent switch choices.

The mean firing rates were then calculated separately for choices associated with each of the eight bins described above. These firing rate averages were assigned to points on a circle with 45° spacing. From these data, a population vector was calculated (vector sum of vectors pointing in these 8 directions with magnitudes equal to the associated firing rates) and the resulting angle used as an index of switch or stay selectivity. The population vector pointed toward 0° if cells were more active around the time of switch choices and pointed toward 180° if cells were more active in the middle of repeated runs of the same choice. For directional statistics tests, the firing rates were normalized such that the peak firing rate was assigned the value 10, and all other firing rates were scaled by the same factor. These values were then considered observations at the eight different angles. The nonuniformity of these data was assessed with the omnibus test for directional data (o-test; Berens 2009), and considered significant at $P < 0.05$. Neurons with significantly nonuniform firing activity were classified into one of four categories. If the direction of the population vectors calculated for the omnibus tests described above fell into quadrant I or IV (i.e., the half-circle centered at 0°) for runs of both liquid rewards, the neuron was classified as a Switch neuron. If the direction of the population vectors fell into quadrant II or III (i.e., the half-circle centered at 180°) for runs of both liquid rewards, the neuron was classified as a Stay neuron. If the direction of the population vector fell into quadrant II or III for runs of apple juice and fell into quadrant I or IV for runs of water, the neuron was classified as a Stay with AJ neuron. If the direction of the population vector fell into quadrant II or III for runs of water and fell into quadrant I or IV for runs of apple juice, the neuron was classified as a Stay with Water neuron.

In addition to neurons that exhibited some form of switch or stay selectivity, many OFC neurons were task selective without switch or stay selectivity by virtue of different firing rates associated with choices of the two liquid reward options. For each neuron, a separate ANOVA compared firing rates associated with all choices of apple juice and all choices of water. Neurons that exhibited significantly ($P < 0.05$) more firing associated with apple juice choices (water choices) but that did not reach significance for any of the switch and stay selectivity tests described above were classified as AJ Selective (H_2O Selective).

RESULTS

Two rhesus monkeys (*monkeys M* and *S*) participated in this study. They repeatedly chose between apple juice and water by pressing one of two buttons. The amount of apple juice or water delivered after a button press was equal (~0.25 ml). The choice behavior of both monkeys included several exceptionally long runs of a repeated choice, which was inconsistent with the prediction of a random choice-selection process. The responses of some neurons in the OFC were found to correlate with the statistically improbable long runs of a repeated choice. Firing activities of other OFC neurons correlated with other aspects of free-choice behavior, such as the decision to switch between options. Neurons were also observed to fire selectively for choices of one of the liquid rewards without selectivity for the location of those choices within a run of the same choice. These results are each described in detail in the following subsections.

Behavioral results. Over 56 recording days, *monkey M* completed 105 individual recording sessions (average of 1.9 recordings/day) and 185 individual work periods (average of 1.8 work periods/recording) in which he repeatedly chose between drops of water and apple juice for at least 10 min. Over 34 recording days, *monkey S* completed 84 individual recording sessions (average of 2.5 recordings/day) and 107 individual work periods (average of 1.3 work periods/recording).

Averaged over all experimental sessions, both monkeys preferred to drink apple juice to water. *Monkey M* worked for ~300 ml of liquid per day, choosing apple juice for 59% of all choices, or more often than water at a ratio of ~3:2 (mean 181 ml apple juice, 124 ml water). This corresponded to an average of 806 drops of apple juice and 489 drops of water per day. *Monkey M* chose apple juice more often than water on 77% of the recording days and 72% of the individual work periods. *Monkey S* worked for ~270 ml of liquid per day, choosing apple juice for 61% of all choices, or more often than water at a ratio of ~5:3 (mean 164 ml apple juice, 103 ml water). This corresponded to an average of 535 drops of apple juice and 280 drops of water per day. *Monkey S* chose apple juice more often than water on 83% of the recording days and 73% of the individual work periods.

The choice behavior of both monkeys included several exceptionally long runs of a repeated choice. The run length distributions were not consistent with a random Poisson process with unequal fixed probabilities. While both monkeys preferred apple juice, they also switched to drink water frequently, and occasionally for several consecutive choices. Choice behavior was further characterized as embedded within a sequence of choices (a run) of the same liquid. Each choice was assigned a position within a run, and the total number of choices within each run, termed the run length, was calculated. The distributions of run lengths segregated by choice are shown in Fig. 1.

The run length distributions are nearly linear on a log-log plot, which is a characteristic of a power-law distribution, such as a Pareto distribution (Barabasi 2005). They agree with

Pareto's 80-20 rule, such that 20% of the runs of apple juice contained 83% of all individual choices for *monkey S* and 79% for *monkey M*. The distributions were less skewed for water choices but still included a long "heavy" tail such that 64% of all water choices were contained in the longest 20% of runs of water for both monkeys. With proper selection of a minimum run length threshold, all of these run length distributions could be fit to a power-law distribution ($P > 0.1$, see MATERIALS AND METHODS). The resulting fits are shown in Fig. 1 as a solid line through the data points (open circles). The expected distributions of a Poisson process with the same average run length are shown as dot-dashed lines in Fig. 1. A Poisson distribution provides a very poor fit for all of the run length distributions in the main task ($P < <10\text{^}-5$; Eden and Kramer 2010). Taken together, the observed run length distributions are very different from what we would expect if the choices were random and dictated solely by unequal probabilities of choosing apple juice and water (i.e., a Poisson process), if one assumes that the relative values of the two rewards are fixed, that is, something other than random chance contributed to the monkey's choice behavior in the free-choice task.

Given that the values for the fits of run length distributions were extreme, we also computed a simple metric, the Fano factor, to assess how Poisson-like each distribution was. Poisson processes have a Fano factor of 1, and empirically derived data that fit a Poisson distribution are usually near 1 (Eden and Kramer 2010). The Fano factors for the run length distributions shown in Fig. 1 were 15 and above (*monkey M*, AJ: 38, Water: 24; *monkey S*, AJ: 42, Water: 15).

Behavioral control experiments. In the normal task, there were two external factors that influenced the monkey's choice on any given trial—the identity of the available liquids and the spatial location of each button. In the main task, both monkeys exhibited a slight preference for one of the two buttons that modulated the relative preference for the two liquid options. *Monkey M* preferred the right button to the left, choosing apple juice 66% of the time when it was associated with a right button press and 58% of the time when it was associated with a left button press. *Monkey S* preferred the left button to the right, choosing apple juice 69% of the time when it was associated with a left button press and 64% of the time when it was associated with a right button press. For both monkeys, therefore, the preference for apple juice dominated the preference for a particular button location in the main task.

Control experiments were run in which both buttons were associated with water to understand the effect of offering two qualitatively indistinguishable rewards in the main task. The control experiment reduced the external factors influencing choice behavior to spatial location alone. In this control task, the spatial preferences of each monkey were expected to dominate the choice behavior. Monkeys were run on three 20-min sessions for three consecutive days in this control experiment. *Monkey S* exhibited a substantial bias for one button starting from the first session, while *monkey M* only began to choose one of the buttons much more than the other on the second session of the second day (and therefore data were analyzed from this point on for *monkey M*).

In the control condition, the preferred and nonpreferred choices were defined by the spatial location of the button. For instance, for *monkey M* the nonpreferred button was the button on the left. The run length distributions for these control experiments are shown in Fig. 1, *bottom*. Both monkeys made very few or no runs of 10 or more for the nonpreferred button in the control task (*monkey M*: 0% of runs; *monkey S*: 1%), and the Fano factor for the distribution of run lengths for the nonpreferred button was close to 1 for both monkeys (*monkey M*: 0.55; *monkey S*: 1.2), suggesting Poisson-like randomness with respect to the decision to repeatedly choose the nonpreferred button. However, the run length distribution data were only significantly fit to a Poisson distribution for the nonpreferred button runs of *monkey M* ($P > 0.05$; Eden and Kramer 2010). All other distributions were not well fit by a Poisson distribution ($P < 0.005$) but were fit to a power-law distribution ($P > 0.1$). While Poisson-like randomness cannot be definitively ruled out for the nonpreferred button choices for *monkey M*, the non-Poisson choice behavior associated with the preferred buttons suggested that the choice behavior was nonrandom in the control task as in the main task, although less balanced. That is, in the control task, very long runs were observed for the preferred button only. The effect, therefore, of using two qualitatively different rewards in the main task may have served only to make the monkeys more likely to make long runs of both reward options.

In a separate control experiment a third external factor was manipulated—the magnitude of water reward—to test whether the monkeys' choice behavior was consistent with their experience of water as rewarding. After three days of water vs. water control experiments in which the reward amounts were equal, and the monkeys had begun to exhibit long runs of one of the buttons, the nonpreferred button was then associated with three times or five times as much water as the preferred button. Both monkeys chose the button associated with the larger magnitude of water substantially more frequently, demonstrating that the spatial bias was a weaker factor than the monkeys' sensitivity to liquid magnitude, and that greater amounts of water resulted in enhanced responding, typical of a rewarding stimulus. While we did not collect enough data in this final control task to generate meaningful run length distributions, the long runs of the larger-magnitude option again suggest nonrandom choice behavior when the monkeys are considering reward options that are identical with the exception of magnitude.

Neural response properties. The neural data set consisted of 344 neurons recorded from the left OFC of two monkeys (*monkey M*: 171, *monkey S*: 173). Several neurons were found to selectively fire around the time that the monkey switched from one choice to another (switch selective) or selectively during long runs of repeated choices of the same liquid (stay selective). Of all the neurons that were selective to switch or stay choices, some responded similarly for each choice and some exhibited opposite responses for the two choices or only responded selectively for portions of repeated runs of only one of the two choices. Other neurons were reward selective without switch or stay selectivity. Figure 2 details example neurons that exhibited these various features.

Figure 2, *top*, shows two neurons that were active when the monkey repeatedly chose apple juice or water. These neurons fired at an elevated rate when the monkey chose the same liquid reward multiple times but were relatively less active during long runs of the other liquid reward. This pattern can be observed in the raster plots in Fig. 2, *top left*. The firing activity appears elevated for several consecutive trials. For the neuron

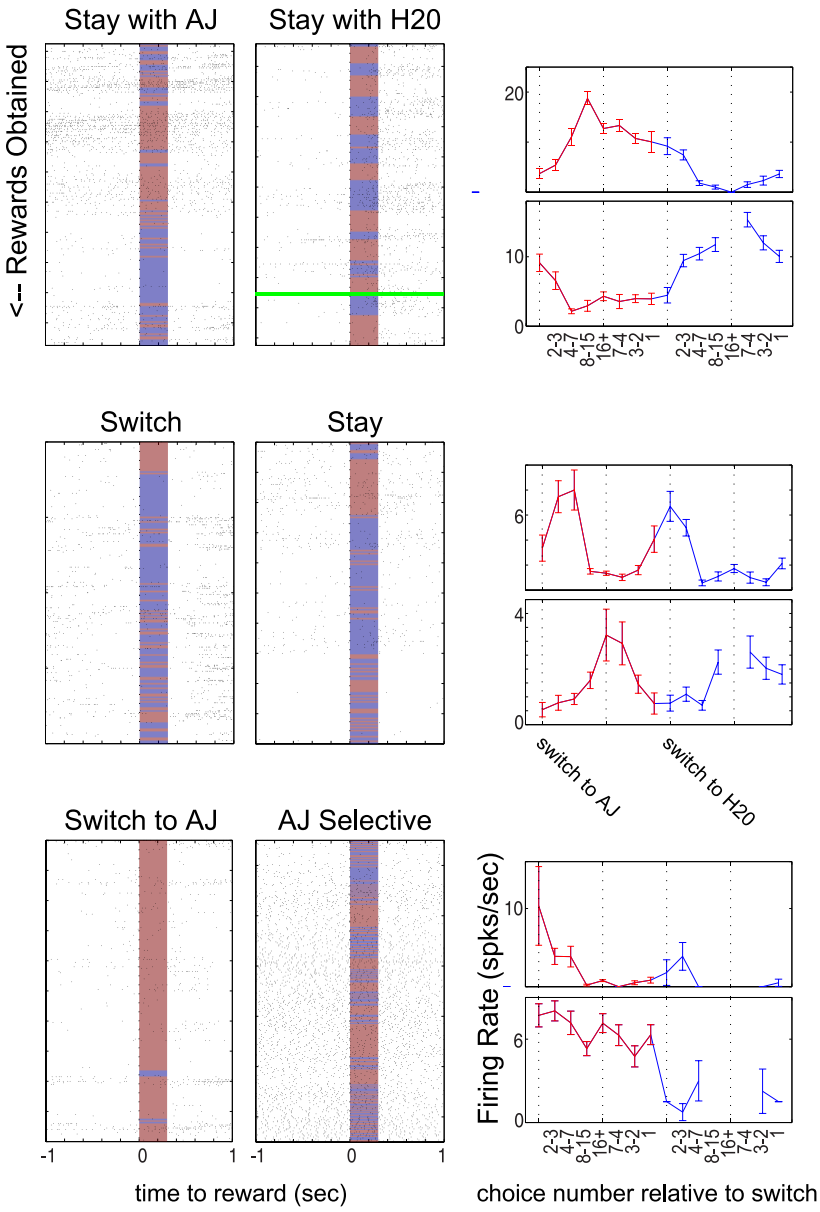

Fig. 2. Example neurons. Firing activity aligned to reward delivery (time = 0). Red indicates delivery of apple juice. Blue indicates delivery of water. Two hundred trials are arranged chronologically *top* to *bottom*. *Right*: mean ± SE binned firing rates with for choices at log2 distance from switch event. Switch choice bins indicated by labels. Two cycles of firing rates are plotted so that periodicity can be observed. *Top* (*bottom*) plots correspond to *left* (*right*) rasters. Firing rate histograms were calculated from the 1-s interval immediately following the button presses (post) for the neurons labeled "Stay with H_2O" and "Switch." All other histograms were calculated from the 1-s interval preceding the button presses (pre).

on the left, this elevated firing continues throughout the trial, with no apparent modulation around the reward delivery on each trial. For the neuron on the right, the elevated firing occurs in the postreward period. These periods of elevated firing are associated with long runs of a given choice (shown in the raster plot as unbroken blocks of red or blue).

The firing rate tuning curves in Fig. 2, *right*, show averaged firing rates associated with choices at various stages within runs of each choice for the neurons in the same row. The top firing rate tuning curve is associated with the left raster plot, and the bottom firing rate tuning curve is associated with the right raster plot. The tuning curves for the neurons labeled Stay with H_2O (*top right*) and Switch (*middle left*) in Fig. 2 were calculated from the 1-s period following the start of the button press. The remaining tuning curves in Fig. 2 were calculated from the 1-s interval just prior to the button press. Within the firing rate plots in Fig. 2, the blue portions of the curves are firing rates associated with water choices and the red portions of the curves are firing rates associated with apple juice choices. Choices of each reward were assigned to one of eight bins, with the first bin containing firing rates from the first choice of each run, the last bin containing firing rates from the last choice of each run, and intermediate bins containing firing rates associated with choices within a run plotted on a log scale (see MATERIALS AND METHODS).

For all of the recorded neurons, a directional analysis technique was used to identify the peaks of the firing rate curves. Two selectivity indexes were computed for each neuron, i.e., one for each liquid. As described in MATERIALS AND METHODS, neurons with Switch selectivity were defined by an index close to 0 (in quadrant I or IV) and neurons with Stay selectivity were defined by an index close to 180 (in quadrant II or III).

The indexes calculated for each liquid did not always agree. Figure 2, *top*, shows two examples of neurons that exhibited Switch selectivity for one liquid and Stay selectivity for the other. The "Stay with AJ" neuron shown in Fig. 2 had a high AJ index (177°, classified as "Stay" selectivity) and a low H_2O index (7°, classified as "Switch" selectivity). The term "Stay

with AJ" emphasizes the high firing activity observed during long runs of apple juice choices that frequently continued for a small number of choices into each run of water. A different naming convention that acknowledges the significant decrease of activity during long runs of water choices could have also been used. The tuning curve for the "Stay with H_2O" neuron was roughly opposite that of the "Stay with AJ" neuron. For the "Stay with H_2O" neuron the selectivity indexes (AJ: −3; H_2O: 259) were likewise roughly opposite those of the "Stay with AJ" neuron (AJ: 177; H_2O: 7).

We found that the firing patterns of OFC neurons, while striking and consistent, were only loosely correlated with individual switch and stay choices. While the neurons featured in Fig. 2, *top*, were significantly more active in the middle of long runs of apple juice and water, respectively, they both continued firing at an elevated rate even after the monkey had switched away from the neuron's preferred reward. The selectivity index that we developed was able to capture the tendency of each neuron to fire around the time of an internally motivated switch choice or, conversely, during the middle of a long run of one choice. While the neurons did not start or stop firing at a precise moment within a given choice sequence, the consistency of the firing patterns with respect to the longer timescale on which the monkey chose to switch between the available reward options is evident in the selectivity indexes as well as the error bars of Fig. 2, *right*.

The spike rasters from the Stay with H_2O neuron in Fig. 2 are shown from two separate work periods. A green horizontal bar on the raster plot marks the transition between the two periods. Work periods were separated by a 10- to 20-min break, after which the button-reward mappings were reversed. It can be observed that this neuron continued to fire selectively for runs of water in the second work period even though the button associated with that reward had changed. Typically, it was not possible to quantitatively assess spatial selectivity for the neurons recorded in our data set. The issue of spatial selectivity is discussed in more detail below.

The neurons in Fig. 2, *middle*, were more active at the beginning of each run, a "Switch" neuron (*left*), or the middle of each run, a "Stay" neuron (*right*). The firing rate plots for these neurons are complementary to each other as well. These neurons, however, have a periodic component that is twice as frequent as the neurons in Fig. 2, *top*. The Switch neuron was active at the beginning of runs of each liquid (most active in the first few bins of both liquids), and the Stay neuron was most active in the middle of runs of both liquids. These firing patterns are also reflected in the selectivity index values. The "Switch" neuron had a low index value for both liquids (AJ: 48; H_2O: 13). The "Stay" neuron had a high index for both liquids (AJ: 197; H_2O: 254).

Several OFC neurons exhibited increased firing when the monkey chose one liquid reward over the other. The Switch to AJ neuron in Fig. 2, *bottom left*, was similar to the Switch neuron shown in Fig. 2, *middle left*, but was more active during switches from water to apple juice. The selectivity indexes for this neuron were low for both liquids (AJ: 24; H_2O: 26). For the purposes of the population analysis described below, this neuron was grouped with the "Switch" neurons. We did not record a sufficient number of neurons to separate out classes of neurons along this axis.

Finally, several neurons exhibited task-related firing rate modulations without exhibiting any form of switch or stay selectivity. These neurons were more active for one of the two choices without additional modulations of firing rate within a run of the same choice. The firing rate tuning curves for these neurons resembled a step function. Figure 2, *bottom right*, shows a neuron that fired selectively for apple juice choices irrespective of the number of apple juice choices already made in the current run.

Population characteristics. Scatterplots of selectivity indexes are shown in Fig. 3 to give a sense of the distribution of cell types in the population. For the sake of comparison between neurons, Fig. 3 only includes data for neurons for which there were data points in each of the firing rate bins shown in Fig. 2, *right*. In practice, this restricted the data shown in Fig. 3 to the neurons recorded while the monkey made at least 1 run of 16 or more consecutive choices of both liquid rewards. Because the monkeys were free to choose between the two rewards, they did not always choose both rewards for long runs during each experimental session. In these cases, we did not have enough data to rule out the possibility that the neuron would have fired at an elevated rate if the monkey had chosen the neglected option for a long run in that experimental session. The "Stay with H_2O" neuron shown in Fig. 2, *top*, was such a neuron. This neuron was most active in medium-long runs of water, but the choice behavior was incomplete because the monkey did not make any long runs of water during that experimental session. Given the distributions of run lengths (Fig. 1), the incomplete choice behavior would lead to a bias toward "Switch" indexes if these neurons were included in the figure.

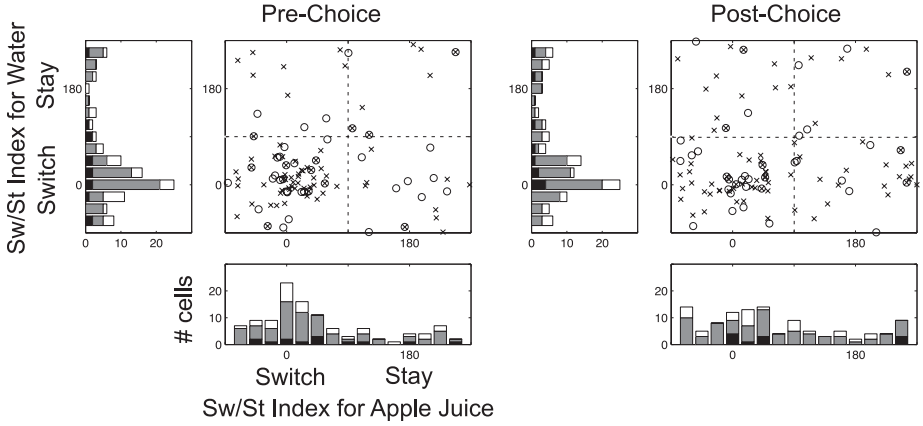

Fig. 3. Population trends. Scatterplots of switch/stay indexes for runs of both liquids. Data from both monkeys are combined. Data on *left* are from analysis of 1-s interval just before each choice. Data on *right* are from analysis of 1-s interval just after each choice. Histograms show cell counts with significant tuning for apple juice (*bottom*) and water (*left*) runs. Black portions of histograms and combined ○ and × markers denote cells with significant tuning for both liquids. Gray (white) portions of histograms and × (○) markers alone denote cells with apple juice (water) tuning only.

Table 1. *Cell counts*

Monkey	*N*	*N*-Full	Sw/St Selective	Quadrant	Liquid Selective
M—pre	171	63	31	(14, 7, 6, 4)	17
S—pre	173	33	18	(11, 2, 3, 2)	6
M—post	171	63	32	(13, 7, 6, 6)	14
S—post	173	33	19	(8, 4, 4, 3)	6

pre, Before reward; post, after reward; *N* = no. of neurons recorded for each monkey; *N*-Full = no. of recorded neurons for which there was a complete set of choice data. During a recording the monkey had to exhibit 1 run of at least 16 choices of both reward options to be considered a behaviorally complete data set. Sw/St Selective = no. of neurons of the set of *N*-Full with a significant Switch/Stay index. Quadrant refers to the 4 quadrants of Fig. 3, which correspond to different neuronal types [Q1: Switch (*bottom left*), Q2: Stay AJ (*bottom right*), Q3: Stay Water (*top left*), Q4: Stay (*top right*)]. Liquid Selective = no. of neurons of *N*-Full with significant selectivity for apple juice vs. water choices but without significant Sw/St selectivity.

The scatterplot in Fig. 3 is also restricted to neurons for which at least one of the tuning curves was significantly tuned (nonuniform, o-test; see MATERIALS AND METHODS). The cells with significant tuning in apple juice runs only are shown by "×." Cells with significant tuning in water runs only are shown by a circle. Cells with significant tuning for runs of both liquids have both a circle and an ×. For the neurons that were selective for only one liquid, the index for the other axis was not statistically significant. Cells that were not significant for runs of either liquid, such as the "AJ Selective" neuron in Fig. 2, are not shown in Fig. 3 because neither of the switch/stay indexes was significant. Cell counts are provided in Table 1.

Each panel in Fig. 3 can be split into quadrants to gauge the prevalence of switch or stay selectivity. The AJ selectivity indexes are plotted on the *x*-axis, and the Water selectivity indexes are plotted on the *y*-axis. The cell counts are summarized in the histograms on the bottom (AJ selectivity indexes) and left (H_2O selectivity indexes). The same analysis was applied to firing activity in the 1-s interval prior to the reward delivery (Fig. 3, *left*) and in the 1-s interval that started with the beginning of the reward delivery (Fig. 3, *right*). For both plots, the bottom left quadrant includes cells that have "switch" tuning for both liquids. This was the most common in the OFC population, comprising about half of the switch/stay-selective cells (see Table 1). This group includes cells that are most active just before (index just below 0) or just after (index just above 0) switch choices. An exemplar of this type of selectivity is shown in Fig. 2, *middle left* ("Switch"). There did not appear to be a strong separation of neurons active prior to switch events compared with after switch events.

As described above, the "Stay with AJ" neuron in Fig. 2 is an exemplar of neurons in the bottom right corner of the plots in Fig. 3. The "Stay with H_2O" neuron in Fig. 2 is an examplar of neurons in the top left corner in Fig. 3. The "Stay" neuron from Fig. 2 is an exemplar of the top right quadrant in Fig. 3. Population average histograms were calculated for all of the recorded neurons that fell into each of these four quadrants, regardless of whether they reached significance individually. These average histograms are shown in Fig. 4. In each of the Switch and Stay histograms there are two prominent peaks. These peaks are centered on the switch events for the Switch neurons and between the switch events for the Stay neurons. For each of the Stay with AJ and Stay with H_2O histograms there is only one prominent peak. These peaks are centered in the middle of the Apple Juice runs and Water runs, respectively.

Note on spatial selectivity. Our task was not designed to disambiguate selectivity for a given reward versus a button location, since the button-reward mappings were reversed only between liquid-reward periods and not on a trial-by-trial basis. The isolation quality remained high enough to justify the recording of a second work period for 92 (of 344) neurons (*monkey M*: 85/171; *monkey S*: 7/173). These neurons were tested in the same manner as above, but the resulting numbers of selective responses were small and effectively from only one subject. Analyzing the neural data with respect to spatial selectivity was made more difficult because of the switch and stay selective responses. These responses could only be assessed if the monkeys made enough long runs of both choices in each of the work periods. Because of these considerations, no definitive conclusions can be made from this data set with regard to spatial selectivity.

Previous studies have shown that the vast majority of OFC neurons were not selective for specific motor commands but rather the rewards that were associated with them (Tremblay and Schultz 1999; Wallis and Miller 2003). Consistent with this possibility, the neuron shown in Fig. 2, *top right*, continued to respond during long runs of water choices after the button-reward mapping had been reversed (above and below the green horizontal bar). It is possible, however, that the stable button-reward mappings employed in this task resulted in a neural code in OFC that did not always disambiguate a reward identity and the movement required to obtain it. This possibility would agree with results from a study of rodent OFC that

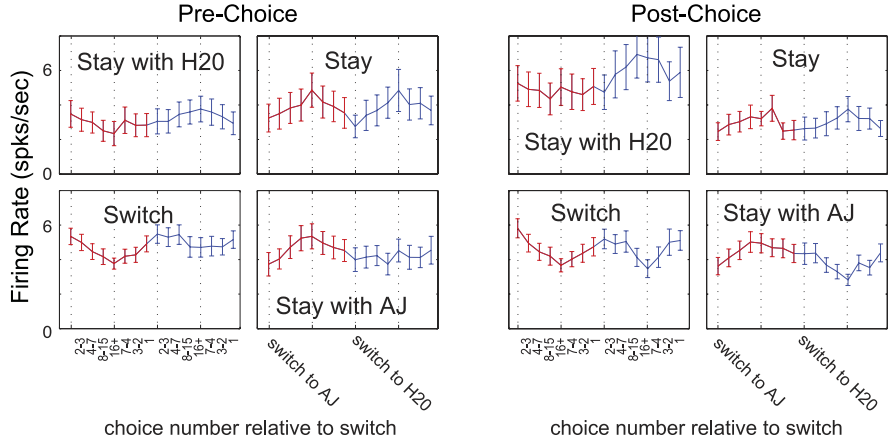

Fig. 4. Population average histograms. Population average firing rate histograms are shown for each of 4 cell type categories. Histograms were calculated for pre- and post-button press analysis intervals separately. Example neurons of each category are shown in Fig. 2. Mean firing rates by choice position were calculated for each neuron separately, as shown in Fig. 2, and then averaged.

found spatial selectivity in a task in which rewards and spatial locations were strongly associated (Feierstein et al. 2006). Future studies will be required to determine whether nonhuman primate OFC neurons are more likely to exhibit spatial selectivity in tasks that feature stable button-reward mappings or other forms of stable mappings between rewards and spatial locations.

DISCUSSION

We investigated the role of internal factors on choice behavior with a simple but novel free-choice task and simultaneous recording from individual neurons in nonhuman primate OFC. We found that, first, the observed sequences of choice behavior included exceptionally long runs of each of two available options, which were not well described by a purely random choice selection process of two unequal fixed values. Second, neurons in OFC encoded important features of the observed choice behavior. This includes neurons that were selectively active during exceptionally long runs of a given choice (stay selectivity) as well as neurons that encode switches between the choices (switch selectivity). These results suggest that OFC neural activity, in addition to encoding subjective values, reflects some of the statistically improbable aspects of free choice. These results could be interpreted either as a form of subjective value on a timescale that changes much faster than has been previously reported or as representing specific aspects of free choice, such as when to explore or exploit. Regardless of the interpretation, the quickly varying signals found in the OFC could be useful for efficient choice behavior when multiple rewards are simultaneously available.

Choice behavior is not random. The matching law formulates a robust observation that subjects choose between two rewarding options in proportion to their relative subjective value (Herrnstein 1961). It remains one of the most successful theoretical accounts of choice behavior to date. It specifically describes aggregate choice behavior but does not describe a mechanism that determines individual choices. For consistency with the matching law, a "null" mechanism for decision-making based on subjective values would be a simple random walk with unequal biases, in which the biases are defined by the relative subjective valuations of the options. This would result in Poisson-like distributions of run lengths (number of repeated choices of a given liquid). Considering that neurons in the nonhuman primate OFC have been shown to encode subjective values (Padoa-Schioppa and Assad 2006; Tremblay and Schultz 1999), it would be tempting to conclude that the role of OFC in decision-making is to represent the subjective values of the different options. These values could be considered the inputs for a subsequent decision step (Rangel and Hare 2010). The results presented here suggest that in order for the term "subjective value" to apply generally to the role of the OFC in decision-making, its meaning must be extended to accommodate a signal that fluctuates on a short timescale.

Assuming stable subjective values, the null model of choice behavior fails to capture the choice behavior that was actually observed in the free-choice task that we employed. As shown in Fig. 1, both of the monkeys that participated in this study exhibited choice behavior that was highly non-Poisson. In particular, both monkeys chose each liquid for exceptionally long runs. This suggests that the relative value of each choice, as determined by the proportion of overall choices allocated to each option, does not account for individual choices by simply performing a random selection process to those values. Instead, we suggest that one of three modifications to the random selection model must be made. First, the choices could continue to be a random selection between two subjective values, but these subjective value signals must be allowed to fluctuate in time, such that the monkey values one option over the other at a proportion that is exceptionally skewed during a long run of the same choice. Second, the random choice process could act upon a relatively stable subjective value representation after it has been modified by separate signals that represent a form of response inertia (i.e., a "stay" signal) and its complement (i.e., a "switch" signal). Third, the transformation from stable long-term values to fluctuating short-term values may not require a set of modulatory signals. Rather, the calculation of short-term fluctuations in reward value may be a by-product of repeated competition between the representations of available rewards in a network of neurons. The complementary firing activities of the neuron types shown in Fig. 2, *top* and *middle*, suggest competition at two levels—between the rewards themselves (Fig. 2, *top*) and between the more abstract decision to switch or stay with the current behavior (Fig. 2, *middle*). Future studies will be required to distinguish between these three explanations of the nonrandom choice behavior observed in repeated free choice between two unchanging rewards.

Neural correlate of switch and stay choices. The firing activity of several OFC neurons in our sample correlated with various features of the monkey's decisions to stay with a certain liquid or switch between liquids. This selectivity is not predicted by the subjective value interpretation of OFC activity that was established by observing changes in firing activity before and after selective satiety for a stimulus to which a neuron was selective (Critchley and Rolls 1996). The timescale of the changes in activity observed in this report was short compared with satiety effects and fluctuated frequently within a short experimental session, suggesting that these signals are qualitatively different from the signals previously described as encoding subjective value. The rapidly fluctuating signals that we observed could be interpreted as a representation of a subjective value signal that changes on a short timescale. In contrast, a subjective value signal that is only modulated by slow satiety mechanisms would be thought to change on a long timescale and could be considered stable over the course of a typical neurophysiology experiment. Recently, Padua-Schioppa and Assad measured relative subjective values by giving monkeys the opportunity to choose between varying amounts of two liquids. That study assumed that the relative preferences for two juices remained stable over the course of an experimental session (Padoa-Schioppa and Assad 2006). The results presented here call attention to the fact that this assumption is only valid if one averages over many similar choices. Over many repeated choices, the short-timescale fluctuations in subjective value might cancel each other, and the average could reflect the relative values of each option that is consistent with the matching law in aggregate. Alternatively, the neurons that encode subjective value on a long timescale, such that they would respond in a stable manner during the course of a short neurophysiology experiment (10–20 min), might have been the same as the neurons identified in our study as liquid selective. These neurons did not change appreciably on a short timescale. An example of such a neuron is shown in Fig. 2, *bottom right*.

The recognition of neural activity in OFC that fluctuates on a shorter timescale than could be attributed to satiety puts the OFC closer to an active role in decision-making. Whether these signals represent switch and stay choices themselves, or rapidly fluctuating subjective values that have a variable influence on a downstream decision-making process, must be addressed in future studies. In addition, we must emphasize that the role in decision-making remains indirect. Switch selective neurons were frequently active for several trials before or after a switch event, and not the switch choice exclusively. Stay-selective neurons likewise may have continued to fire even after the subject had made a new switch choice. The loose correlation between OFC activity and choice behavior might also be partly attributable to the relatively short intertrial interval employed in this study (2.5 s). If the monkeys had been required to wait a longer interval between choices, we might have observed greater correlation between the firing activity and choice behavior. The short intertrial interval employed in this study might be "oversampling" a decision process that naturally evolves more slowly.

Neural network underlying behavioral flexibility. Our results elucidate the role of OFC in behavioral flexibility. Other areas are also likely to contribute to the management of the competing drives of staying and switching, such as the ACC (Kennerley et al. 2006), the subthalamic nucleus and the pre-SMA (Isoda and Hikosaka 2007, 2008), the locus coeruleus (Aston-Jones and Cohen 2005), and the serotonergic system (Clarke et al. 2004).

The findings presented here bear similarities to elements of previous studies. OFC activity was observed to be precisely time-locked to choice behavior in rats, specifically at the start of a cluster of licks (Gutierrez et al. 2006). Switch-related activity found in our task by definition correlated with a cluster of a new behavioral choice. Switch-selective neurons might also be thought to encode novelty. When the monkey chooses a new liquid, it is relatively novel. Novelty-selective responses have been observed previously in OFC (Rolls et al. 2005). The switch-selective neurons reported here, however, could not have been the same as the novelty-selective neurons identified by Rolls and colleagues. Switch-selective neurons in the present study responded every time the monkey switched from one liquid to the other, even though both of these liquids were used exclusively for the entirety of the recordings, spanning months, and were therefore exceedingly familiar. We conclude that OFC contains neurons that are sensitive to novelty in the sense that the stimulus has never been experienced before (Rolls et al. 2005) and that it has not been experienced recently, as shown here.

Across the population, we observed more neurons with switch selectivity compared with stay selectivity and more neurons responding selectively for apple juice choices. This suggests that the portions of OFC that we studied were predominantly selective for changes in the current behavior, with possibly more neural tissue selective for the generally preferred options. Noting that the medial and orbital portions of area 10 in the OFC in monkeys have been identified with the human frontal pole area (Semendeferi et al. 2001), similar switch selectivity has been described previously based on functional MRI (fMRI) analysis in humans (Daw et al. 2006). fMRI studies, however, are not able to identify the smaller intermixed subpopulation of neurons that are selectively activated for stay choices.

"Clean" switching. Switch- and stay-selective neurons can account for clean switching (as opposed to dithering) in self-guided behavior. If deciding between two dependably available appetitive rewards, the subjective value of each reward changes in time as the organism becomes selectively satiated on the option that it chooses, while the value of the other option rises as a growing unmet need (or at least decreases more slowly). If subjective values of two juices drift into each other, a switch decision could be expected once the equivalence point is reached. This simple mechanism, however, leads to an ecological problem known as behavioral dithering, or repeatedly switching between two alternatives, when the subjective valuations of two options are nearly equal (McFarland 1989). While the chosen option might have a higher value at that moment, drinking the liquid will bring the value back down slightly, since the monkey is now becoming selectively sated on this new choice, which will soon result in a new switch and then repeated switching (or dithering). To handle the issue of dithering, theoretical accounts have posited the existence of a switch controller embedded within the action selection mechanism that could facilitate persistence with the newly chosen option (Prescott et al. 2006; Redgrave et al. 1999; Snaith and Holland 1990). The switch-selective neurons might push the nearly equal choice values to be more separated, leading to response stability following a change of the choice. Stay-selective neurons might also contribute to clean switching. We speculate that stay-selective neurons could provide "response inertia," so that the monkey continues to choose the same reward until its value drops well below the value of other options. Once a new choice is finally made, there would already be a separation in value between the two options, resulting in clean switching. This mechanism could help the organism efficiently exploit multiple simultaneously available reward sources by minimizing the costs associated with switching between them.

Recent evidence in humans showed that the OFC was especially activated when preference decisions were made between closely valenced options (Kim et al. 2007). This suggests that OFC is an essential component in free-choice behavior that depends on subjective valuation. The present results suggest that, when repeatedly choosing between closely valued options, the OFC has a role in allocating choice behavior to both options in a way that allows for extended sampling of each. If choice behavior were determined solely by long-term stable preferences, the monkeys would rarely switch between options. As has been formalized in the matching law, however, real choice behavior includes sampling of the available options and therefore switching between them. Individual choices remain difficult to predict, leading to the postulation of a random choice selection process to account for the choice behavior in individual trials.

Our results show that a purely random selection process does not account for individual choices. In particular, the tendency for very long runs of each choice could not be explained as a Poisson process, demonstrating that decision-making is not random. The run length distributions, however, could be fit with a power-law distribution. Power laws describe a variety of natural phenomena including, most famously, the distribution of wealth in human societies. One possible mechanism for power-law distributions is "preferential attachment" (Barabasi 2005). The power-law distribution of wealth could be explained as a feedback loop in which wealthy individuals are more likely to acquire more wealth (i.e., "the rich get richer"). In the present case, we suspect that the simple act of choosing one of two well-known options could make the same choice more likely in the near term. This could provide the feedback loop necessary for power-law behavior to emerge. But just as decision-making is not random, the tendency

to choose the known, nonpreferred option indicates that decision-making is not rigid. The observed free choice behavior was between the extremes of rigid value maximization and a whimsical random process.

ACKNOWLEDGMENTS

We thank Tomomi Kano and Tessa Yao for administrative support, Kelsie Pesja and Nicole Sammons for assistance with the animals, Victor Schebertayuk for technical support, Rick Paniagua for custom machine shop assistance, and Tim Heitzman for electrical shop assistance.

GRANTS

This work was supported by JST.ERATO Shimojo Implicit Brain Functions Project, JST.CREST Implicit Interpersonal Information Project, National Eye Institute (to R. A. Andersen), and Moore Foundation (to R. A. Andersen).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: M.C., R.A.A., and S.S. conception and design of research; M.C. and K.K. performed experiments; M.C. and K.K. analyzed data; M.C., K.K., R.A.A., and S.S. interpreted results of experiments; M.C. prepared figures; M.C. drafted manuscript; M.C., K.K., R.A.A., and S.S. edited and revised manuscript; M.C., K.K., R.A.A., and S.S. approved final version of manuscript.

REFERENCES

Aston-Jones G, Cohen JD. Adaptive gain and the role of the locus coeruleus-norepinephrine system in optimal performance. *J Comp Neurol* 493: 99–110, 2005.

Barabasi AL. The origin of bursts and heavy tails in human dynamics. *Nature* 435: 207–211, 2005.

Bechara A, Damasio H, Damasio AR. Emotion, decision making and the orbitofrontal cortex. *Cereb Cortex* 10: 295–307, 2000.

Berens P. CircStat: a MATLAB toolbox for circular statistics. *J Stat Software* 31: 10, 2009.

Carmichael ST, Price JL. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol* 371: 179–207, 1996.

Clarke HF, Dalley JW, Crofts HS, Robbins TW, Roberts AC. Cognitive inflexibility after prefrontal serotonin depletion. *Science* 304: 878–880, 2004.

Clauset A, Shalizi CR, Newman MEJ. Power-law distributions in empirical data. *SIAM Rev* 51: 661–703, 2009.

Corrado GS, Sugrue LP, Seung HS, Newsome WT. Linear-nonlinear-Poisson models of primate choice. *J Exp Anal Behav* 83: 581–617, 2005.

Critchley HD, Rolls ET. Hunger and satiety modify the responses of olfactory and visual neurons in the primate orbitofrontal cortex. *J Neurophysiol* 75: 1673–1686, 1996.

Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. *Nature* 441: 876–879, 2006.

Eden UT, Kramer MA. Drawing inferences from Fano factor calculations. *J Neurosci Methods* 190: 149–152, 2010.

Feierstein CE, Quirk MC, Uchida N, Sosulski DL, Mainen ZF. Representation of spatial goals in rat orbitofrontal cortex. *Neuron* 51: 495–507, 2006.

Fellows LK, Farah MJ. Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cereb Cortex* 15: 58–63, 2005.

Gutierrez R, Carmena JM, Nicolelis MAL, Simon SA. Orbitofrontal ensemble activity monitors licking and distinguishes among natural rewards. *J Neurophysiol* 95: 119–133, 2006.

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

Isoda M, Hikosaka O. Role for subthalamic nucleus neurons in switching from automatic to controlled eye movement. *J Neurosci* 28: 7209–7218, 2008.

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

Kennerley SW, Dahmubed AF, Lara AH, Wallis JD. Neurons in the frontal lobe encode the value of multiple decision variables. *J Cogn Neurosci* 21: 1162–1178, 2009.

Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF. Optimal decision making and the anterior cingulate cortex. *Nat Neurosci* 9: 940–947, 2006.

Kepecs A, Uchida N, Zariwala HA, Mainen ZF. Neural correlates, computation and behavioural impact of decision confidence. *Nature* 455: 227–231, 2008.

Kim H, Adolphs R, O'Doherty JP, and Shimojo S. Temporal isolation of neural processes underlying face preference decisions. *Proc Natl Acad Sci USA* 104: 18253–18258, 2007.

McFarland D. *Problems of Animal Behaviour*. London: Longman, 1989.

Montague PR, Berns GS. Neural economics and the biological substrates of valuation. *Neuron* 36: 265–284, 2002.

Padoa-Schioppa C, Assad JA. Neurons in the orbitofrontal cortex encode economic value. *Nature* 441: 223–226, 2006.

Prescott TJ, Montes González FM, Gurney K, Humphries MD, Redgrave P. A robot model of the basal ganglia: behavior and intrinsic processing. *Neural Netw* 19: 31–61, 2006.

Rangel A, Hare T. Neural computations associated with goal-directed choice. *Curr Opin Neurobiol* 20: 262–270, 2010.

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

Roesch MR, Calu DJ, Burke KA, Schoenbaum G. Should I stay or should I go? *Ann NY Acad Sci* 1104: 21–34, 2007.

Rolls ET, Browning AS, Inoue K, Hernadi I. Novel visual stimuli activate a population of neurons in the primate orbitofrontal cortex. *Neurobiol Learn Mem* 84: 111–123, 2005.

Semendeferi K, Armstrong E, Schleicher A, Zilles K, Van Hoesen GW. Prefrontal cortex in humans and apes: a comparative study of area 10. *Am J Phys Anthropol* 114: 224–241, 2001.

Snaith M, Holland O. An investigation of two mediation strategies suitable for behavioural control in animals and animats. In: *From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behavior, Paris, France*. Cambridge, MA: MIT Press, 1990.

Tremblay L, Schultz W. Relative reward preference in primate orbitofrontal cortex. *Nature* 398: 704–708, 1999.

Wallis JD, Miller EK. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur J Neurosci* 18: 2069–2081, 2003.

Walton ME, Devlin JT, Rushworth MFS. Interactions between decision making and performance monitoring within prefrontal cortex. *Nat Neurosci* 7: 1259–1265, 2004.