Behavioral/Systems/Cognitive

Adaptation of Reward Sensitivity in Orbitofrontal Neurons

Shunsuke Kobayashi, Ofelia Pinto de Carvalho, and Wolfram Schultz
Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3DY, United Kingdom

Animals depend on a large variety of rewards but their brains have a limited dynamic coding range. When rewards are uncertain, neuronal coding needs to cover a wide range of possible rewards. However, when reward is likely to occur within a specific range, focusing the sensitivity on the predicted range would optimize the discrimination of small reward differences. One way to overcome the trade-off between wide coverage and optimal discrimination is to adapt reward sensitivity dynamically to the available rewards. We investigated how changes in reward distribution influenced the coding of reward in the orbitofrontal cortex. Animals performed an oculomotor task in which a fixation cue predicted the SD of the probability distribution of juice volumes, while the expected mean volume was kept constant. A subsequent cue specified the exact juice volume obtained for a correct saccade response. Population responses of orbitofrontal neurons that reflected the predicted juice volume showed adaptation to the reward distribution. Statistical tests on individual responses revealed that a quarter of value-coding neurons shifted the reward sensitivity slope significantly between two reward distributions, whereas the remaining neurons showed insignificant change or lack of adaptation. Adaptations became more prominent when reward distributions changed less frequently, indicating time constraints for assessing reward distributions and adjusting neuronal sensitivity. The observed neuronal adaptation would optimize discrimination and contribute to the efficient coding of a large variety of potential rewards by neurons with limited dynamic range.

Introduction

Adaptation is a ubiquitous property of the brain that enables efficient processing of diverse physical events by systems with limited dynamic coding range (Fairhall and Bialek, 2002; Dean et al., 2005). For example, retinal ganglion cells increase sensitivity to amplify weak sensory inputs in a dark environment and decrease sensitivity to prevent strong inputs from saturation in a bright environment (Hosoya et al., 2005; Dunn et al., 2007). Adaptation to the statistics of environmental stimuli is known to improve discrimination and thus increase the effective operating range of the neural system.

The biological value of stimuli in the world is also highly diverse and statistical. Foraging animals encounter different ranges of rewards depending on various factors such as place and season. To assure optimal chances for survival, the brain's reward system needs to discriminate a variety of possible rewards. Thus it is fundamental to understand how efficient neuronal reward coding is when the statistical properties of reward distributions change.

The orbitofrontal cortex is a key reward structure of the brain. Monkeys with orbitofrontal lesions respond abnormally to changes in reward contingencies (Iversen and Mishkin, 1970; Dias et al., 1996) and show altered reward preferences (Baylis and

Gaffan, 1991). Orbitofrontal neurons are sensitive to different types and magnitudes of reward (Thorpe et al., 1983; Tremblay and Schultz, 1999; Hikosaka and Watanabe, 2000; Wallis and Miller, 2003), and thus translate various reward features into a scalar measure of reward value.

An important question is whether the reward signals in the orbitofrontal cortex adapt to changes in contexts and statistical distributions of reward. Tremblay and Schultz (1999) demonstrated that value-coding responses in the orbitofrontal cortex shifted their reference according to the available rewards that changed in every block of trials. Neuronal sensitivity tuning was optimized to the center of currently expected reward values, suggesting neuronal adaptation to the mean, first moment, of reward distributions.

Adaptation to the second statistical moment, represented by variance or SD, is critical for efficient neuronal coding. Unlike mean adaptation, SD adaptation requires changes in neuronal sensitivity because different ranges of input signals have to be mapped on naturally limited neuronal encoding range. In this study, we aimed to examine whether orbitofrontal neurons adapt to the SD of reward distributions provided they were known in advance. We used specific visual cues that predicted different SDs while keeping the means of the distributions constant. We measured neuronal response to a subsequent value cue which specified the exact upcoming juice volume. To confirm the animals' learning and preferences of the cues, we examined behavioral choices between the cues associated with different juice volumes.

Materials and Methods

Experimental design
We used two reward distributions with different SDs, and each distribution consisted of three equiprobable juice volumes. For each distribution, a specific geometric picture predicted the SD (SD cue),
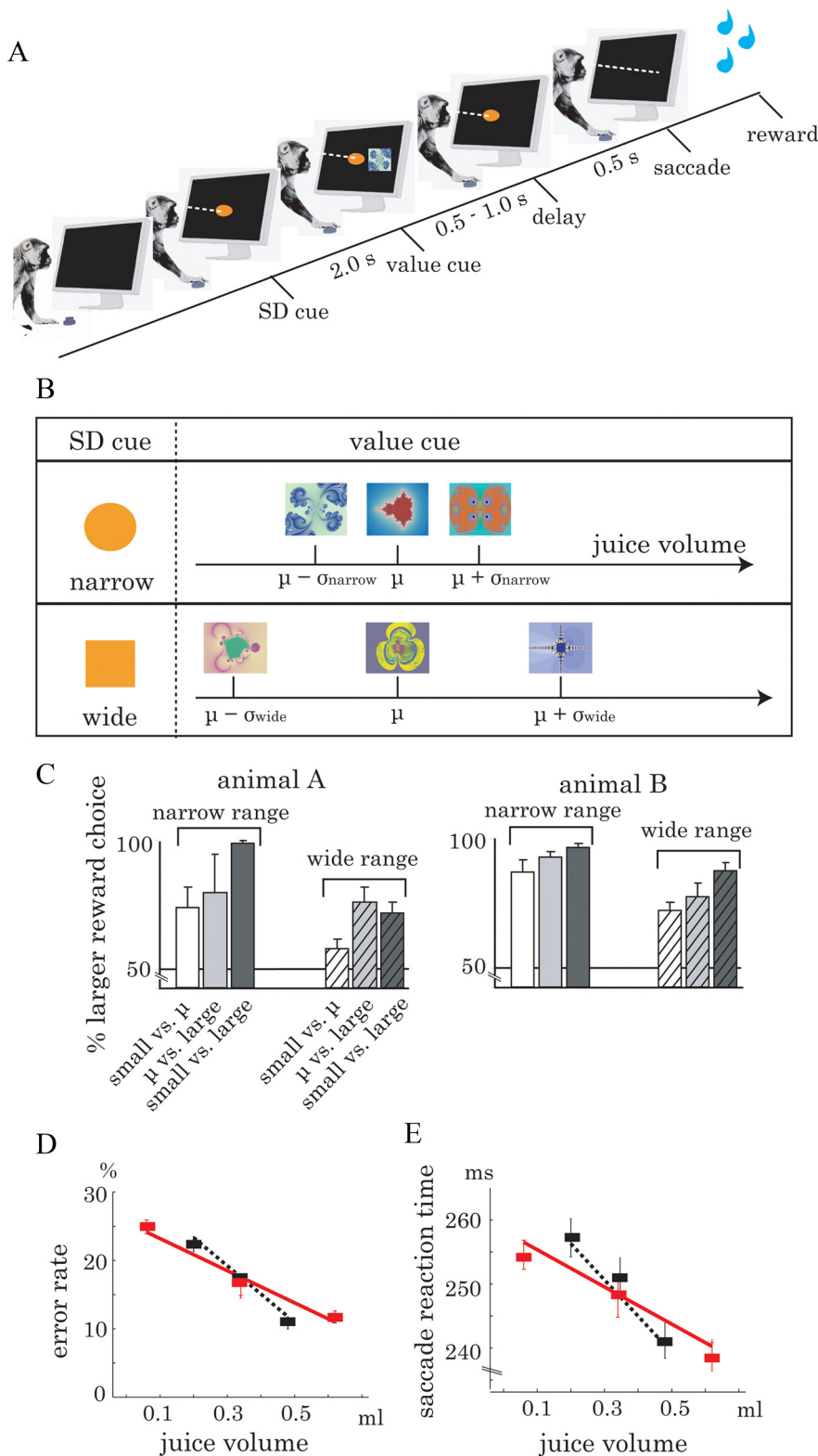
Figure 1. Experimental design and behavioral results. **A**, Imperative saccade task used for neuronal recording. A trial started when the animal touched an immobile key and gazed at the central cue that indicated the SD of the outcome reward distribution (SD cue). If the animal maintained eye fixation for 2.0 s, a peripheral picture (value cue) was presented briefly indicating the location of future saccade and the volume of upcoming juice reward. Disappearance of the SD cue signaled the monkey to make a saccade to the previously cued location. Successful saccades were followed by juice delivery at the predicted volume. **B**, Stimulus–reward mapping. The SD cue (left) predicted the SD of the reward distribution. Different fractal pictures (value cues, right) indicated different juice volumes. Two different SDs of reward were tested (σ_{narrow}, σ_{wide}) with the same mean (μ). **C**, Animal choice behavior. Preferences to different value cues were tested in the choice saccade task. Both animals chose the cues associated with larger volumes of juice regardless of small or large SD. Error bars represent SD. **D, E**, Behavioral adaptation to the predicted reward

and a subsequent specific fractal picture predicted the reward volume (value cue) (Fig. 1A,B). The two SD cues and the six value cues were counterbalanced across the two animals used. Mean reward volume was identical in each distribution.

Subjects and surgery

We used two adult male monkeys (*Macaca mulatta*), weighing 10 and 14 kg. Before the recording experiments started, we implanted under general anesthesia a head holder and a chamber for access to the brain via a small opening in the cranium while keeping the dura intact. All experimental protocols were approved by the Home Office of the United Kingdom.

Behavioral paradigm

Imperative saccade task. Animals sat in a primate chair, at 45 cm in front of a computer display (Fig. 1A,B). An immovable, touch-sensitive resting key was mounted on the right hand side in front of the animal. Each trial started with the SD cue (2.6° visual angle, presented at monitor center). The animal touched the resting key and maintained eye fixation at the center for 2.0 s. Then the value cue (7.3°) appeared briefly (0.5–1.0 s) at either the left or right at a distance of 10.6° from the center of the monitor. After 0.5 s the SD cue disappeared, triggering a saccade to the location of the value cue. Following a correct saccade, a red circle (1.1°) appeared at target location. The animal maintained eye fixation on the target for 1.5 s until the red circle changed to green, which signaled to release the resting key. One second after key release, blackcurrant squash juice was delivered at the volume predicted by the value cue. Trials were immediately aborted after premature break of fixation, inaccurate saccades, or premature release of the resting key.

We tested each neuron with two different reward distributions. Variations in reward volume consisted either of different juice volumes delivered as a single shot (animal A, 0.06–0.62 ml) or of different numbers of successive shots of fixed juice volumes (animal B, 1–9 shots, 1 shot = 75 μl). For animal A, we used juice volumes of 0.20, 0.34, and 0.48 ml in the narrow distribution (SD, σ_{narrow} = 0.114 ml) and 0.06, 0.34, and 0.62 ml in the wide distribution (SD, σ_{wide} = $2\sigma_{narrow}$ = 0.228 ml). For animal B, we used juice volumes of 3, 5, and 7 drops (0.23, 0.38, and 0.53 ml) in the narrow distribution (SD, σ_{narrow} = 1.63 drops = 0.122 ml) and 1, 5, and 9 drops (0.08, 0.38 and 0.68 ml) in the wide distribution (SD, σ_{wide} = $2\sigma_{narrow}$ = 3.26 drops = 0.244 ml). The method of juice delivery, single shot with variable volume or

---

distribution during the imperative saccade task. Error rate (**D**) and saccadic reaction time to the value cue (**E**) are plotted against the predicted juice volume. The shifts of the regression slope between narrow (dotted black line) and wide (solid red line) reward distributions suggest scaling of both behavioral measures to reward range. Error bars represent SEM.

Table 1. Number of task-related, value-coding, and adaptive responses in each task period

	Value cue	Delay	Saccade	Reward	Total
Task-related responses	149	97	87	97	430
	A: 101	A: 65	A: 50	A: 60	A: 276
	B: 48	B: 32	B: 37	B: 37	B: 154
Value-coding responses	63 (42.3%)	40 (41.2%)	19 (21.8%)	27 (27.8%)	149 (34.7%)
	A: 40 (39.6%)	A: 27 (41.5%)	A: 10 (20.0%)	A: 11 (18.3%)	A: 88 (31.9%)
	B: 23 (47.9%)	B: 13 (40.6%)	B: 9 (24.3%)	B: 16 (43.2%)	B: 61 (39.6%)
Adaptive responses	22 (34.9%)	12 (30%)	1 (5.3%)	3 (11.1%)	38 (25.5%)
	A: 16 (40%)	A: 11 (40.7%)	A: 1 (10.0%)	A: 2 (18.2%)	A: 30 (34.1%)
	B: 6 (26.1%)	B: 1 (7.7%)	B: 0 (0%)	B: 1 (6.3%)	B: 8 (13.1%)

Top, Task-related responses. Responses significantly different from baseline (intertrial interval) activity were counted as task-related responses in each task period ($p < 0.01$ post hoc Scheffé test). Middle, Value-coding responses. Responses significantly correlated with predicted juice volume were counted as value-coding responses in each task period ($p < 0.05$, Spearman's rank correlation, corrected for multiple comparisons). Numbers in parentheses indicate proportions of value-coding responses in task-related responses during each task period. Bottom, Adaptive responses. Neuronal adaptation was examined by comparing the regression slopes between narrow and wide reward distributions using regression models (Eqs. 1, 2). Numbers in parentheses indicate adaptive responses as percentage of value-coding responses during each task period. A and B refer to the two animals used.

multiple shots of fixed volume, did not significantly affect the proportion of value coding responses within the task-related responses or the proportion of neurons that adapted to reward distributions ($p > 0.5$, χ^2 test). Thus, we collapsed neuronal data sampled with two different methods of reward delivery. We routinely calibrated the solenoid valve to assure that valve opening times corresponded exactly to the set juice volumes.

To examine the speed of neuronal adaptation, each sampled neuron was tested with one of three schedule types (random, mini block, or large block schedule). In the random trial schedule, the reward distribution changed pseudorandomly in every trial (see Fig. 5A, top). In the mini block schedule, the specific reward distribution used was fixed for a small number of trials (4–13 trials, mean 6.4 trials); narrow and wide distributions alternated frequently (Fig. 5A, middle). In the large block schedule, we kept the same reward distribution constant for a large number of trials (14–93 trials, mean 28.1 trials); thus reward distributions alternated infrequently (Fig. 5A, bottom). The SD cue always indicated the specific distribution used in each trial. We used the three schedules randomly and independent of responses and anatomical coordinates of sampled neurons.

Choice saccade task. A choice task served to assess the animals' learning and preferences of the different volumes of juice reward. This task was identical to the imperative saccade task, except that two value cues appeared simultaneously at left and right monitor positions (10.6° from center). The animal chose one target by a saccadic eye movement. In each trial we presented randomly two value cues from each reward distribution (small versus intermediate, intermediate versus large, small versus large). Positions of the two value cues were randomized. Behavioral preferences were expressed as probability of choosing the larger volume of reward in each cue combination (Fig. 1C). Behavioral testing with the choice task alternated in trial blocks with the imperative task during neuronal recordings.

Recording procedures

Conventional techniques of *in vivo* extracellular recordings served to study the activity of single orbitofrontal neurons. Animal A provided neuronal data from the left hemisphere, the recording chamber being centered at A 34.5 and L −11. Animal B provided data from both hemispheres, centered at A 36 L9 and A33 L −9. A stainless steel guide tube (0.8 mm diameter) served to insert a single tungsten microelectrode into the brain (125 μm diameter, 1–5 MΩ initial impedance at 1 kHz; Frederick Haer). Although the guide tube conceivably caused more damage to the overlying dorsolateral cortex and white matter compared to solid microelectrodes alone, it permitted the use of thinner microelectrodes causing very little damage to the orbitofrontal cortex itself. A hydraulic micromanipulator (MO-95, Narishige) advanced the microelectrode vertically in the stereotaxic plane. Discharges from neuronal perikarya were amplified, filtered (300 Hz to 2 kHz) and monitored with oscilloscopes. An adjustable Schmitt trigger converted neuronal discharges into standard digital pulses which were continuously monitored on a digital oscilloscope together with the original waveforms. Custom-made software on a Macintosh IIfx computer (Apple) controlled the behavioral

task. An infrared eye tracking system monitored eye position with 200 Hz (5 ms) resolution (ETL200; ISCAN).

Data analysis

We analyzed neuronal data from the imperative saccade task in several consecutive steps. First, we identified task-related activities by comparing activity between five task periods: intertrial interval (from 1.0 to 0.5 s before the SD cue), value cue (from 0 to 0.5 s after value cue onset), delay (from 0.5 to 1.0 s after value cue onset), saccade (from 0.3 s before to 0.2 s after saccade onset), and reward (from 0 to 0.5 s after reward onset) ($p < 0.01$, one-way ANOVA). For those neurons that showed significant task relation, we further identified the task periods in which responses were significantly different from baseline activity during the intertrial interval ($p < 0.01$, post hoc Scheffé test).

In the second step, we searched for value-coding responses in all task-related responses identified by the ANOVA. We used the nonparametric Spearman's rank correlation coefficient to assess a possible relationship to juice volume ($p < 0.05$, corrected for multiple comparisons).

In the third step, we assessed neuronal adaptation to the SD of reward distributions in all value-coding responses identified by Spearman's correlation. We tested the hypothesis that adaptation of neuronal coding to SD consisted in a change of value-response slope, being steeper with the narrow compared to the wide distribution. We used the following two linear regressions as main model that allowed us to obtain and compare directly the response slopes between the two distributions:

$$Y = \beta_0 + \beta_{\text{narrow}}(X_{\text{narrow}} - \mu), \qquad (1)$$

$$Y = \beta_0 + \beta_{\text{wide}}(X_{\text{wide}} - \mu), \qquad (2)$$

where β_0, β_{narrow}, and β_{wide} were unstandardized regression coefficients. X_{narrow} and X_{wide} were juice volumes with equiprobable three elements ($X_{\text{narrow}} = \{\mu - \sigma_{\text{narrow}}, \mu, \mu + \sigma_{\text{narrow}}\}$, $X_{\text{wide}} = \{\mu - \sigma_{\text{wide}}, \mu, \mu + \sigma_{\text{wide}}\}$). Y is discharge rate in a given task period in response to juice at volume X. The terms μ and σ are expected mean and SD of the reward distributions, respectively. Mean juice volume μ was identical in both reward distributions; therefore it should produce similar neuronal responses. Based on this assumption, the two models were constrained to the point (μ, β_0), which simplified the subsequent analyses. The regression coefficients, β_{narrow} and β_{wide}, predicted neuronal response changes per unit juice volume (impulses/s/ml); thus they reflected the neuronal response slope, i.e., neuronal sensitivity to reward. Responses with different signs of β_{narrow} and β_{wide} were excluded from all analysis (8.5% of whole value-coding responses).

To test for neuronal adaptation to the two reward distributions, we compared the two response slopes (β) in each individual value-coding neuron with the normalized-t approximation method (DeShon and Alexander, 1996) with the null hypothesis of $\beta_{\text{narrow}} = \beta_{\text{wide}}$. A significantly steeper slope with the narrow compared to the wide distribution would reject the null hypothesis and indicate adaptive neuronal coding. We used two-tailed t test to assess both adaptive ($|\beta_{\text{narrow}}| > |\beta_{\text{wide}}|$) and inverse ($|\beta_{\text{narrow}}| < |\beta_{\text{wide}}|$) slope changes ($p < 0.05$). However, inverse
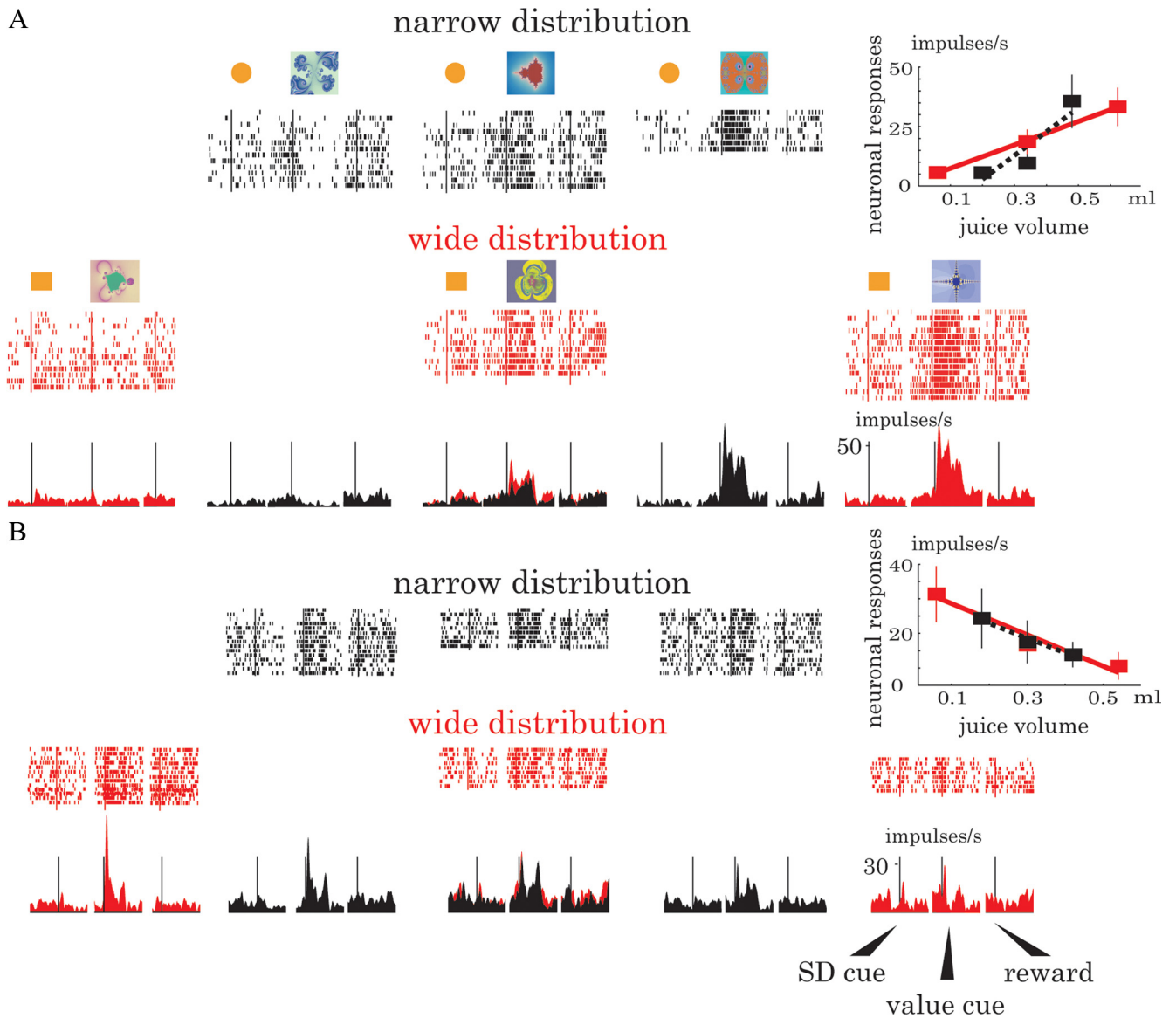
Figure 2. Examples of two value-coding orbitofrontal neurons. *A*, Adaptation in a neuron whose response increases with increasing juice volume. The slopes in the top right regression plot show the relationships between neuronal responses (ordinate, impulses/s) and predicted juice volume (abscissa, ml), separately for small (black) and large (red) SDs. The slope changes indicate adaptation of reward sensitivity to predicted reward distribution. *B*, Lack of adaptation in a neuron whose response decreases with increasing juice volume. The regression lines for the two reward distributions were parallel, indicating graded coding across all five reward volumes and thus lack of adaptation. Error bar, SEM. For each raster, the sequence of trials runs from top to bottom. Vertical lines in rastergrams indicate onsets of SD cue (left), value cue (center) and reward (right). Tick marks in rastergrams indicate neuronal impulses, histograms below rastergrams display mean discharge rates (black, small SD; red, large SD).

slope changes were never found to be significant. We quantified the degree of neuronal adaptation by an adaptation score, defined as the ratio between β_{narrow} and β_{wide} ($\beta_{narrow}/\beta_{wide}$) obtained from Equations 1 and 2. Adaptation score > 1.0 indicated steeper slopes in the more narrow distribution ($|\beta_{narrow}| > |\beta_{wide}|$).

Our main regression model (Eqs. 1, 2) estimated and compared directly the response slopes for the two distributions, thus reflecting straightforwardly the experimental rationale for adaptive coding. To confirm the results of the main model, we conducted separate tests with hierarchical regression models (supplemental Eqs. S1–S3) and an interactive regression model (supplemental Eqs. S4, S5 in supplemental material, available at www.jneurosci.org). The data obtained with the main and the two supplementary regression models were almost identical (Table 1; supplemental Tables S1, S2, available at www.jneurosci.org as supplemental material). The main description of the results will be largely based on the main model because of its direct and intuitive comparison of response slopes as substrate of adaptation.

We used mutual information to measure the effects of adaptation on neuronal discrimination between juice volumes. Predictable information of juice volume associated with neuronal responses (I) was quantified as decrease in entropy of stimulus occurrence $H(J)$

$$I = I(J;X) = H(J) - H(J|X)$$

$$= \sum_J -p(j)\log(p(j)) - \left\langle \sum_J -p(j|x)\log(p(j|x)) \right\rangle_x, \quad (3)$$

where J is a set of value cues j, X is a set of neuronal responses x, $p(j|x)$ is the conditional probability of a value cue j given an observed impulse count x, and $p(j)$ is the a priori probability of value cue j. We corrected for potential bias in the values of mutual information caused by the limited number of trials and uneven distribution of data samples (Treves and Panzeri, 1995; Kobayashi et al., 2002). We calculated encoded reward information in narrow and wide reward distributions separately in sliding time windows of 200 ms width that moved in 5 ms steps.
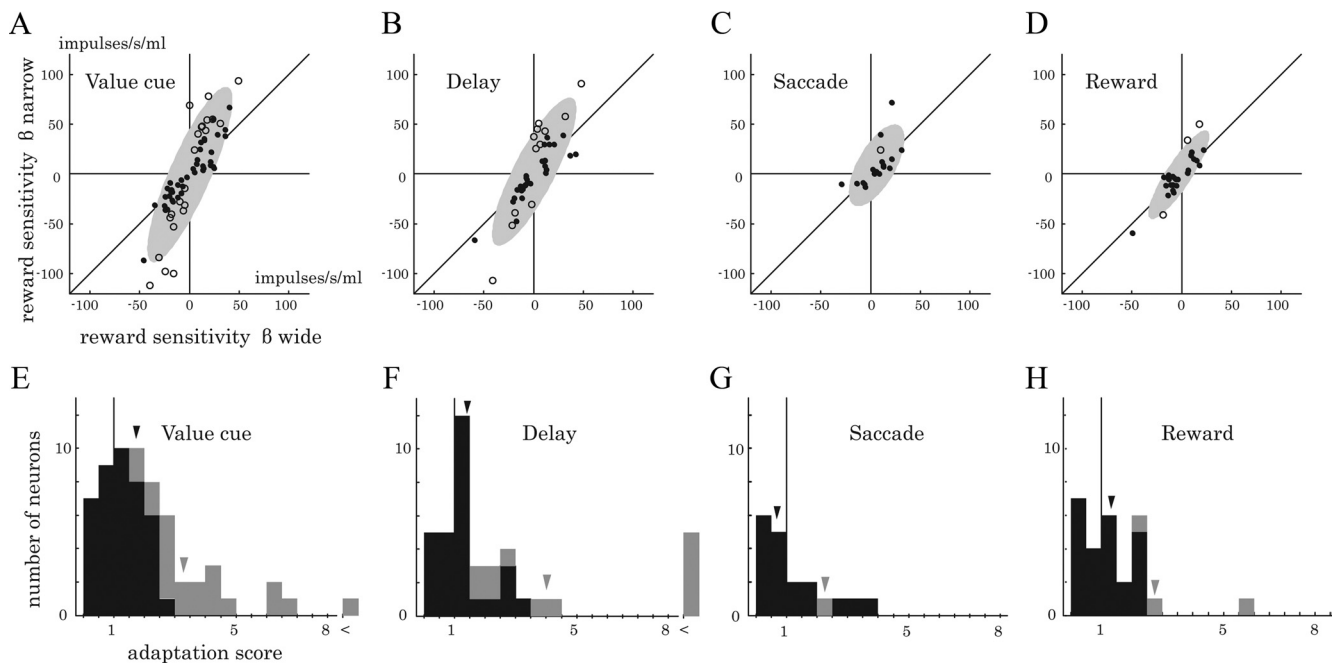
Figure 3. Adaptations of orbitofrontal reward sensitivity to predicted reward distribution. **A–D**, Plots of response slopes for large versus small SD (abscissa vs ordinate). Slopes (β) of value-coding responses were estimated by linear regression models (Eqs. 1, 2) for each neuron in each task period and reflect discharge rate per unit juice volume. Each circle indicates significant (open) or insignificant (filled) slope change between the two reward distributions from each value-coding response ($p = 0.05$). Symbols above the diagonal unit line in the upper right quadrant and below the unit line in the lower left quadrant indicate steeper reward slope with smaller compared to larger SD ($|\beta_{narrow}| > |\beta_{wide}|$). Shaded ellipses delineate the distribution contours (2 SD) of all value-coding responses. **E–H**, Histograms of adaptation scores. The adaptation score quantified the degree of adaptation and is defined as $\beta_{narrow}/\beta_{wide}$. Distributions of adaptation scores were significantly shifted to >1.0 for responses to value cue and delay periods ($p < 0.05$, t test), indicating adaptation to SD. Black and gray arrowheads indicate median scores of all value-coding activities and statistically significant adaptive responses, respectively.

Off-line analysis used MATLAB for Windows (version 7.5, Math-Works). All impulse analyses used only trials in which the animals made correct behavioral responses.

Recording positions
During the last recording sessions with each animal, we placed small marking lesions by passing negative currents (5–10 μA for 5–20 s) through the microelectrode, while positioning larger lesions (20 μA for 20 or 60 s) at a few locations higher in the same tracks. This procedure resulted in distinct patterns of vertically oriented histological marks. Animals were killed with an overdose of sodium pento-barbital (90 mg/kg, iv) and perfused with 4% paraformaldehyde in 0.1 M phosphate buffer through the left ventricle of the heart. Frozen coronal sections were cut on a cryotome at every 50 μm parallel to the recording microelectrode tracks. The sections were stained with cresyl violet.

Results

Behavior in choice saccade task
We measured choice preferences to assess the animals' discrimination between the fractal pictures associated with different juice volumes. In the saccadic choice task, animals chose between small and intermediate, intermediate and large, or small and large volumes (animal A, 562 trials; animal B, 1581 trials). Overall correct task performance was 73% (animal A) and 88% (animal B). Animals chose the pictures indicating the larger of two juice volumes significantly more often compared to the smaller volume, both with the narrow distribution (animal A, $\chi^2 = 31.3$; animal B, $\chi^2 = 221$) and the wide distribution (animal A, $\chi^2 = 41.3$; animal B, $\chi^2 = 200$, 1581 trials; $p < 0.001$ for all, χ^2 test) (Fig. 1C). Thus, both animals appropriately preferred the pictures associated with the larger juice volumes, indicating good discrimination between the reward predicting pictures.

Behavior in imperative saccade task
The animals performed the imperative task correctly in 71.8% of trials, overall. Typical errors were premature fixation breaks and inaccurate saccades. Animals made fewer errors when the value cue predicted larger juice volumes. The regression slopes in Figure 1D depict the relationship between error rate and predicted juice volume. Interestingly, the slope was steeper when the reward was given in the narrow range (dotted black line; slope $= -0.41$, $r^2 = 0.14$, $p < 0.001$) compared with reward in the wide range (solid red line; slope $= -0.23$, $r^2 = 0.15$, $p < 0.001$).

We found similar behavioral adaptation in saccade responses to the value cue in correct trials (Fig. 1E). Saccade reaction time was generally faster when larger volume of juice was predicted, as indicated by the negative regression slopes. Reaction time changed within a similar range (240–260 ms) for both reward distributions. As a result, Δ reaction time per unit juice volume was larger (56.6 ms/ml, dotted black line) when juice volume varied in the narrow range compared with the wide range (29.0 ms/ml, solid red line).

The shifts of regression slopes in the measures for error rate and saccade reaction time indicate that behavioral responses adapted to the range of juice distribution predicted by the SD cue (see supplemental material for further behavioral analysis, available at www.jneurosci.org). The results suggest that animals used the cue information to adjust their behavioral responses and discrimination.

Adaptation of neuronal value signals
We recorded the activity of 876 single neurons from the orbito-frontal cortex of two monkeys during task performance (animal
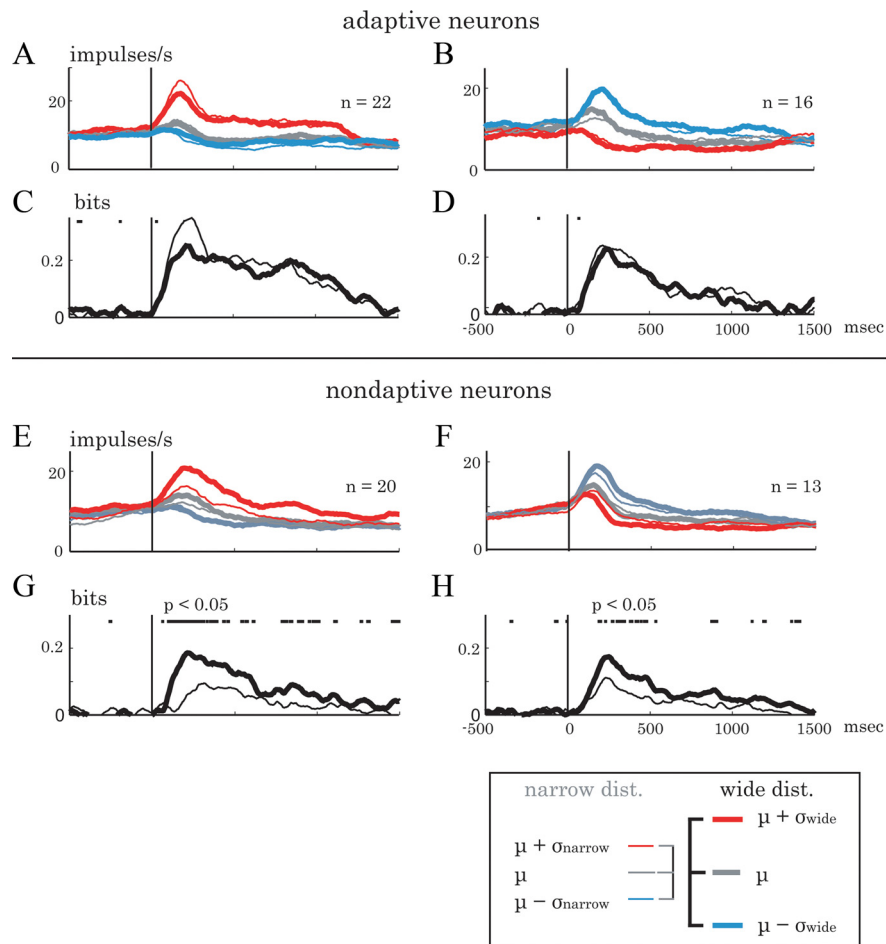
Figure 4. Population histograms of discharge rate and reward information. **A, B, E, F**, Average responses to value cues that showed significant (**A, B**) or insignificant (**E, F**) adaptation to SD of reward distributions. Responses varied positively (**A, E**) or negatively (**B, F**) with reward volume. Thick lines refer to large SD and thin lines to small SD of reward volume. Blue, gray, and red lines indicate small, intermediate, and large juice volumes. Juice volume increased according to thick blue < thin blue < thick gray = thin gray < thin red < thick red (inset). With adaptive responses (**A, B**), thick and thin lines of same color largely overlapped, indicating slope adaptation to reward range. In the population lacking significant adaptation, responses increased (**E**) or decreased (**F**) monotonically across all five physical juice volumes used. **C, D, G, H**, Population-averaged reward information. Thick black line, large SD of reward volume; thin gray line, small SD. Adaptive responses carried similar amount of reward information in two reward distributions (**C, D**). In contrast, nonadaptive responses lost reward information with small SD compared to large SD (**G, H**). Horizontal ticks indicate periods during which reward information was higher in the wide compared with narrow reward distribution ($p < 0.05$, 2-tailed paired t test, uncorrected). The mutual information was calculated using a sliding window (duration, 200 ms; step size, 5 ms) and averaged across neurons.

A, 464 neurons; animal B, 240 neurons). Of these, 189 neurons (26.8%) showed activity significantly related to the task (animal A, 123 neurons [26.5%]; animal B, 66 neurons [27.5%]; $p < 0.01$, one-way ANOVA). We examined the sensitivity of these neurons to the volume of predicted or received reward separately in four task periods (value cue, delay, saccade, and reward periods). A total of 149 responses from 86 neurons showed significant relationships to the juice volume, which we termed value coding (Table 1) ($p < 0.05$, Spearman's correlation, corrected for multiple comparisons). The total number of responses was larger than the total number of neurons, as some neurons showed value sensitivity in multiple task periods.

The neuron displayed in Figure 2A shows an example of adaptive value coding with two different reward distributions tested in separate blocks of trials. In the first block (black raster-histograms), an orange circle (SD cue) predicted the narrow distribution of juice volume (0.20–0.48 ml) and the subsequent

fractal picture (value cue) specified the exact juice volume for each trial (0.20, 0.34, and 0.48 ml from left to right). The neuron showed higher and more sustained responses when the value cue predicted larger juice volume (activity after the value cue, second vertical line). In the second block (red raster-histograms), the orange square predicted the wider distribution of juice volume (0.06–0.62 ml). Similar to the value-coding responses observed in the first block, the response of this neuron showed positive correlation with predicted juice volume. However, despite the twofold difference in SD, the response varied within the same range (5–35 impulses/s) in both reward distributions. The similar responses to the different minimal and maximal juice volumes in the two distributions indicated neuronal adaptation to the predicted distributions.

To examine adaptive value coding more closely, we estimated the relationship between the neuronal response and the predicted juice volume with linear regressions (Eqs. 1, 2). The regression slope (β) reflected the change in impulse activity per unit juice volume (impulses/s/ml), which is a direct physiological measure of reward sensitivity. For this particular neuron, the regression slope of the value cue response was steeper for the narrow compared to the wide juice distribution [100.0 impulses/s/ml (black dotted line) vs 49.3 impulses/s/ml (red solid line)] (Fig. 2A, top right). The slope change indicated higher reward sensitivity in the more narrow reward distribution as a result of the adaptation.

In contrast, Figure 2B shows the response of an orbitofrontal neuron that failed to adapt to the predicted reward distribution. The response to the value cue decreased monotonically across the five associated reward volumes. The largely overlapping regression lines indicate that the reward sensitivity of this neuron remained constant despite the change in reward distribution [-42.3 impulses/s/ml (black dotted line) vs -45.3 impulses/s/ml (red solid line)] (Fig. 2B, top right).

Population analysis of neuronal adaptation

To assess adaptation of neuronal sensitivity to reward distributions, we compared the response slopes between the two reward distributions in each individual neuron, using the slope parameters β obtained from our main regression model (Eqs. 1, 2).

Figure 3A–D shows response slopes from the narrow distribution (ordinate) plotted against slopes from the wide distribution (abscissa) for the four task periods. Responses showing positive slopes with juice volume ($\beta > 0$) appear in the upper right quadrants. Importantly, many of these positive value-coding responses were found above the diagonal line, indicating

steeper slopes with the narrow compared to the wide reward distribution. Responses showing negative relations with juice volume ($\beta < 0$) appear in the lower left quadrant. These responses were often below the diagonal line in the lower left quadrant, indicating again steeper slopes with the narrow distribution. The differences in slope between the two distributions were significant in 38 of the 149 value-coding responses (Fig. 3A–D: open circles; Table 1) ($p < 0.05$; normalized t-approximation method comparing Eqs. 1 and 2). The remaining 111 of 149 responses showed insignificantly different slopes that fell close to the diagonal line, indicating lack of significant adaptation (closed circles).

The two other regression models described in the supplemental material confirmed and extended these results (available at www.jneurosci.org). The hierarchical regression model allowed us to identify both presence and absence of adaptation (supplemental Eqs. S1–S3, supplemental Table S1, available at www.jneurosci.org as supplemental material) and we found slightly more adaptive ($n = 42$) than nonadaptive ($n = 33$) responses. The remaining unassigned responses had intermediate characteristics (neither adaptive nor nonadaptive), adding to an overall adaptive population response. We also tested adaptation as an interaction between reward volume and SD and identified 43 adaptive responses (supplemental Eqs. S4, S5, supplemental Table S2, available at www.jneurosci.org as supplemental material). Together, the largely overlapping results from all three regression models suggest that a sizeable fraction of the sampled value-coding neurons showed adaptation of reward-response slope to the SD of reward distributions.

The distribution ellipse of response slopes (gray shade) was tilted counterclockwise away from the diagonal unit line, particularly for value cue and delay periods (Fig. 3A, B) ($p < 0.01$, paired t test). These data indicated a net adaptation of reward sensitivity in our population of value-coding orbitofrontal neurons.

We further quantified the degree of adaptation for each neuron by taking the ratio of response slopes from the two reward distributions (adaptation score $= \beta_{\text{narrow}}/\beta_{\text{wide}}$) (Fig. 3E–H). Activities with high adaptation score were observed mainly during the value cue and delay periods. The unimodal distributions of adaptation scores indicate that adaptability was continuously graded across the value-coding population rather than being restricted to a distinct group of neurons.

Informational aspects of adaptive value coding
Previous studies on visual processing suggested that neuronal adaptation to probability distributions of inputs resulted in more efficient coding with increased amount of information transmission (Wainwright, 1999; Brenner et al., 2000). Adaptation in reward systems may improve coding efficiency by adjusting the distribution of neuronal responses to the expected reward distribution. To address this issue, we examined input (juice volume)/
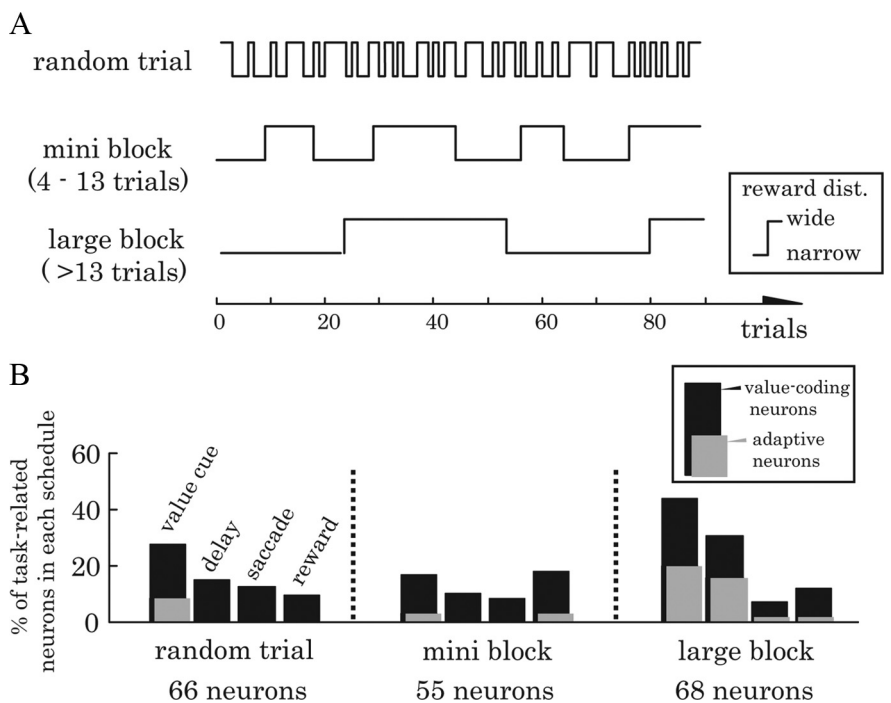

Figure 5. Neuronal adaptation to reward distribution during schedules of different volatility. **A**, Top, Narrow and wide reward distributions changed pseudorandomly in every trial. Middle, Reward distribution changed every 4–13 trials. Bottom, Reward distribution changed only between large blocks of trials (>13 trials). **B**, Proportions of value-coding (black bars) and adaptive (gray bars) responses in total task-related neurons sampled in each schedule (left, random trial; middle, mini block; right, large block) during different task periods (value cue, delay, saccade, reward from left to right). Total numbers of task-related neurons sampled in three schedule types are shown below schedule labels.

output (impulse rate) matching in the populations of adaptive and nonadaptive neurons.

We grouped neuronal responses with statistically significant adaptation according to their positive (Fig. 4A, 22 responses) or negative (Fig. 4B, 16 responses) relationships to juice volume. The dynamic range of the population response in these neurons was nearly the same with the narrow (thin line) and wide (thick line) reward distributions (minimum of 4–5 impulses/s and maximum of 18–23 impulses/s). As a consequence, the neurons appeared to discriminate juice volumes equally well in the two distributions, even with the rather smaller volume differences in the more narrow distribution.

Coding efficiency can be characterized by using mutual information theory (Golomb et al., 1997), and the method has been applied to neurophysiological data (Gershon et al., 1998; Kobayashi et al., 2002). We quantified neuronal discrimination of juice volume as bits of information, separately for the two reward distributions. The average reward value information carried by the responses with statistically significant adaptation peaked at ~0.2 bits (Fig. 4C, positive value-coding; D, negative value-coding). This value was similar for the narrow (thick line) and wide (thin line) reward distributions ($p > 0.05$, two-tailed paired t test). Thus, as a result of sensitivity adaptation, these neuronal responses carried the same amount of information regardless of the distribution of inputs.

These adaptive responses contrasted with responses lacking adaptation according to the supplementary regression model (supplemental Eqs. S1–S3 in supplemental material, available at www.jneurosci.org). When juice volume varied within a wider range, neuronal responses also changed within a wider range (thick lines in Fig. 4E,F), as shown by the distinct differences for the large (red line), intermediate (gray line) and small (blue line)
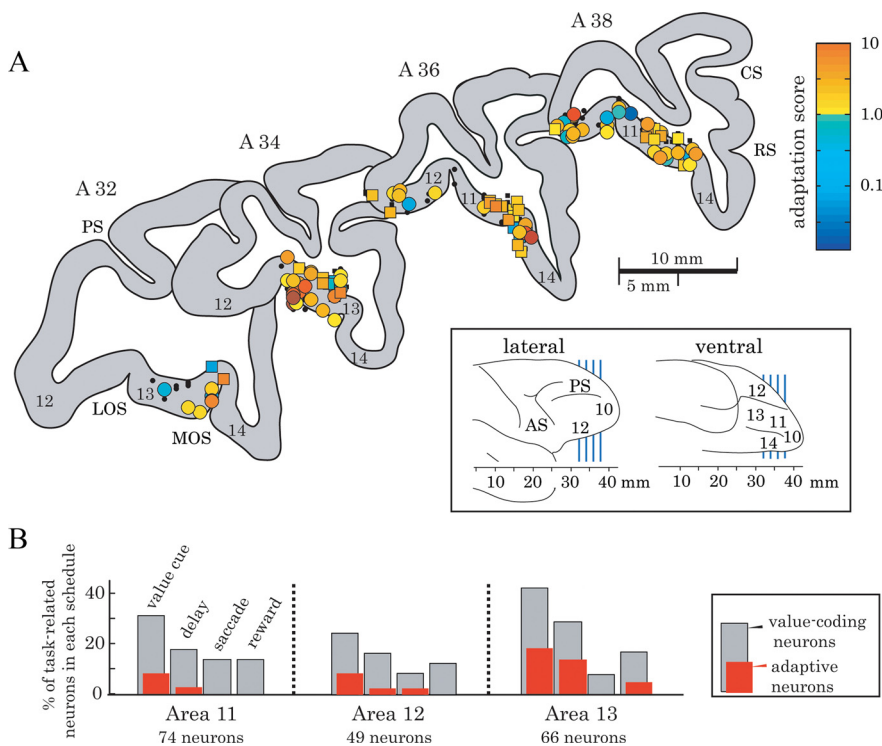
Figure 6. Anatomical locations of sampled orbitofrontal neurons. **A**, Locations of single neurons are marked with colors reflecting the adaptation score (ratio of regression slope: $\beta_{narrow}/\beta_{wide}$; compare right color scale). Positions of neurons sampled from three hemispheres of two monkeys are superimposed and mapped on four coronal sections. A 32, A 34, A 36, and A 38 denote stereotaxic rostrocaudal coordinates, indicated by blue vertical lines in the inset. Circles, Animal A. Squares, Animal B. Small black symbols, Neurons related to task but not coding value. AS, Arcuate sulcus; PS, principal sulcus; LOS, lateral orbital sulcus; MOS, medial orbital sulcus; RS, rostral sulcus; CS, cingulate sulcus. Numbers on gray areas and in the inset refer to Walker's cytoarchitectonic areas. **B**, Proportions of value-coding (gray bars) and adaptive (red bars) responses in the whole task-related neurons sampled in each orbitofrontal subarea during different task periods (value cue, delay, saccade, and reward periods from left to right). Numbers of all task-related neurons sampled in the three areas are shown below the area labels.

juice volumes. For the narrow reward distribution, the range of neuronal responses was also narrow (thin lines in Fig. 4 E,F). Correspondingly, the population response showed a loss of discrimination with more narrow distributions (Fig. 4 E, 20 responses; F, 13 responses). These neurons lost information when coding the narrow compared to the wide distribution (horizontal ticks above the histograms indicate $p < 0.05$ by two-tailed paired t test, uncorrected) (Fig. 4 G,H). Together, nonadaptive neuronal responses processed reward value less efficiently than adaptive responses.

Speed of neuronal adaptation

Each neuron described above was tested in one of the three adaptation schedules that varied the two reward distributions at different speeds. In the first, random trial schedule, distributions changed pseudorandomly in every trial (Fig. 5 A, top). Among 66 task-related neurons tested with this schedule, 47 responses encoded juice volume in at least one of the four task periods. Only six of the 47 value-coding responses (12.8%) showed significant adaptation to the reward distribution. The second schedule involved slower switches between the two distributions. It consisted of mini blocks with a small number of trials (4 –13 trials, mean 6.4 trials) within which the reward distribution was fixed (Fig. 5 A, middle). Thus narrow and wide distributions alternated less frequently than in the random trial schedule. In 55 task-related neurons tested with this schedule, only 5 of 32 value-coding responses showed significant adaptation (15.6%). In the third

schedule, we created a more stable situation by keeping a reward distribution fixed for a relatively large block of trials (14 –93 trials, mean 28.1 trials) (Fig. 5 A, bottom). Thus, neuronal adaptation could occur not only in response to the explicit cue in every trial, but also could be based on the reward distribution sampled during the preceding trials. In 68 task-related neurons tested with this schedule, 27 of 70 value-coding responses showed significant adaptation to reward distribution (38.6%). Thus, less frequent changes in value distributions resulted in higher incidence of adaptive coding.

Figure 5 B summarizes the proportion of adaptations in value-coding responses, separately for each schedule. Neuronal adaptations were uneven among the three schedules ($p = 0.01$, χ^2 test). Post hoc tests revealed that adaptations occurred more frequently during the large block schedule compared to both random trial and mini block schedules ($p < 0.01$, χ^2 tests with correction for multiple comparisons). The incidence of adaptation did not differ significantly between random and mini-block schedules ($p > 0.05$). Thus, the slower switches and longer stability of each reward distribution with the large block schedule appeared to favor neuronal adaptations.

Positions of neurons

Histological reconstruction of recording sites revealed that the sampled orbitofrontal neurons were located in areas 11 (74 neurons), 12 (49 neurons), and 13 (66 neurons) (Fig. 6 A). Clusters of neurons in caudal area 11 and rostral area 13 showed value-coding responses mainly during the value cue and delay periods. The percentage of value-coding responses was highest in area 13, although the difference against the other two areas was insignificant (Fig. 6 B, gray bars) ($p > 0.1$, χ^2 test with correction for multiple comparisons). Adaptive neurons were distributed unevenly across the three anatomical areas (Fig. 6 B, red bars) ($p = 0.02$, χ^2 test). Pairwise post hoc tests revealed that adaptive neurons were more common in area 13 than area 11 or 12 ($p < 0.05$). The distribution of adaptive neurons differed insignificantly between areas 11 and 12 ($p > 0.1$, χ^2 test with correction for multiple comparisons).

Discussion

The present data show adaptation of response gain to predicted reward distributions in a population of orbitofrontal neurons. The tested distributions varied in SD but had constant mean. Adaptive neurons coded the more narrow range of reward with steeper response slopes, thus "zeroing in" on the currently valid range. Slope adaptation led to maximal reward discrimination within each distribution and thus optimal coding efficiency. In contrast, nonadaptive orbitofrontal neurons used smaller coding ranges with more narrow distributions, resulting in fixed reward sensitivity, reduced reward discrimination and less transmitted information. Thus, adaptation
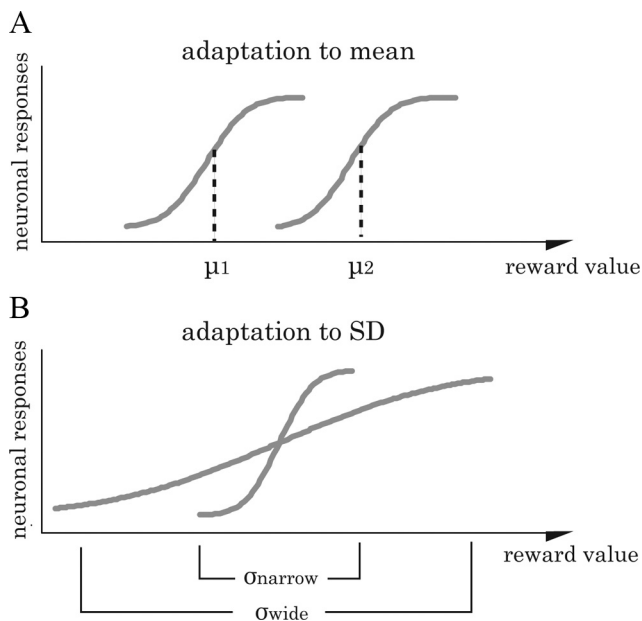
A

B

Figure 7. Schematic forms of adaptation to reward distributions in orbitofrontal neurons. **A**, Adaptation to mean reward distribution (approximated from data by Tremblay and Schultz, 1999). Neuronal response slopes shift into the predicted distribution, rather than stretching across the full range of the two distributions combined. **B**, Adaptation to SD of reward distribution (current data). Neuronal response slopes become steeper with more narrow distributions, and flatten with wider distributions. The two forms of adaptation refer to different parameters of distributions but represent the same phenomenon, namely matching of neuronal responses to predicted and currently used reward distributions. The slopes reflect the quasilinear part of reward response slopes.

consisted of efficient matching of the neuronal encoding range onto the predicted reward range. Such neuronal mechanisms would allow animals to forage successfully for essential resources in variable environmental situations and thus increase their chances for survival.

Behavioral adaptation to reward distributions
Both monkeys discriminated the value cue reliably during the saccade choice task based on the size of associated juice reward (Fig. 1C). Also, error rate and saccade reaction time during the neuronal recordings varied monotonically with the juice volume predicted by the value cue. Importantly, the animals scaled these behavioral responses to the juice range predicted by the SD cue. These results indicate behavioral sensitivity and adaptation to the explicit reward cues.

Nature of adaptive coding
Earlier studies showed that reward-related responses reflect the animals' relative preference among the available rewards rather than physical reward properties (Tremblay and Schultz, 1999; Cromwell et al., 2005; Hosokawa et al., 2007). Relative preference coding could be interpreted as neuronal adaptation to the mean of reward distributions matching the centers of input/output ranges; when the reward distribution shifted, the same output (neuronal response) range covered the new input (reward value) range (Fig. 7A). This type of adaptation does not necessarily involve changes in response slopes.

The present data demonstrate neuronal adaptation to the SD of reward distributions, which essentially involves changes in reward sensitivity slope (Fig. 7B). We kept the mean constant and varied the width of reward distributions. However,

adaptation to the mean and SD could happen at the same time. Formally, the normalization to the mean and SD of reward distributions resembles the statistical z-score [(value − mean)/SD].

Adaptive coding may appear inconsistent with previous studies that showed cardinal (number-like) value coding independent of changes in contexts (Padoa-Schioppa and Assad, 2006, 2008). The discrepancy among the studies might be explained by differences in the examined neuronal populations within the orbitofrontal cortex. Indeed, our adaptive neurons were more commonly located in agranular area 13 than in dysgranular areas 11 and 12 (Fig. 6). Another explanation might lie in different task designs. Tremblay and Schultz (1999) found neuronal adaptation using a block design. We found that adaptation to the SD increased when reward distributions changed less frequently (see below for further discussion). Padoa-Schioppa and Assad (2008) used a random trial design. They required animals to choose between two different kinds of juice reward (called menu). The menu and the juice quantity changed randomly in every trial. The maximum juice quantity was adjusted such that the juice value varied in largely overlapping ranges across menus. Thus, the reward distributions were nearly fixed across menus, which would not have helped to reveal value adaptation. A relatively slow speed of adaptation may explain why Padoa-Schioppa and Assad (2008) failed to see adaptive coding with the random trial design. In a recent follow-up study, Padoa-Schioppa (2009) compared the activity of orbitofrontal neurons tested with different value distributions. The experiment varied the maximum value across separate blocks of trials while keeping the minimum value constant at zero. In this situation, the range of orbitofrontal responses stayed constant despite the changes in range and mean of the value distribution. This adaptation is consistent with the present results.

Taking the results of the present and the previous studies together (Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2008; Padoa-Schioppa, 2009), value signals of adaptive neurons would show parametric relationships within the specific reward distributions but not beyond these bounds. Adaptive value signals would not reflect value on a common scale across an unlimited range. Adaptive signals would rather provide accurate discrimination and allow transitivity within the bounds of given reward distributions. In contrast, nonadaptive neurons would allow for more stable and transitive value coding over wider value ranges, as discussed below.

Temporal requirements for adaptation
In the present study, the initial SD cue predicted a particular distribution with three possible juice volumes, and the subsequent value cue predicted the specific volume delivered in each trial. When trials with different distributions varied pseudorandomly between trials, neuronal adaptation depended solely on the prediction given by the SD cue. Less than 15% of orbitofrontal responses showed adaptation to the value cue ~2 s later. However, we found substantially more adaptive coding when we switched distributions less frequently using the large block schedule (Fig. 5). The block design afforded additional time and additional reward prediction by the block context. These data suggest that reward adaptations in the orbitofrontal neurons take time in the order of seconds to tens of seconds. In contrast to these orbitofrontal processes, dopamine reward prediction error responses adapted within short delays of 2 s to predicted means and SDs (Tobler et

al., 2005). Thus, reward adaptations in different brain structures may operate at different time scales.

Neuronal adaptations to environments that change without explicit cues require sampling and estimation of stimulus distributions (Wark et al., 2007). The sampling and estimation process will primarily determine the adaptation speed. Thus, it should be noted that our task conceivably reduced the sampling time by providing explicit predictive stimuli. This is not a fundamentally different procedure, as explicit predictive stimuli induced similar visual adaptations as changing environments (Hosoya et al., 2005). Despite faster sampling time, adaptations in the present study were slower than the fastest visual adaptations with switches between known distributions (Fairhall et al., 2001), indicating different time courses between visual and reward adaptations or between the involved brain structures, or both.

Coding efficiency

Adaptive coding contributes importantly to the efficiency of neuronal information processing (Barlow, 1961; Laughlin, 1981; Brenner at al., 2000; Maravall et al., 2007). Our data suggest that adaptation maintains the amount of information transmitted by neuronal responses, whereas lack of adaptation results in information loss. Too narrow a sensitivity would result in miscoding or outright missing of peripheral values, and too wide a sensitivity would lead to unnecessarily flat slopes with poor discrimination within the used input range. Rescaling the response slope to match the relevant reward range maximizes information transmission within the limited dynamic range of neuronal responses and thus contributes to efficient reward coding.

Neuronal adaptation occurs in all major sensory systems (Laughlin and Hardie, 1978; Dean et al., 2005; Maravall et al., 2007). The observed adaptations in reward neurons might simply reflect the known adaptations of sensory inputs that give rise to reward value signals. However, reward value is a more abstract parameter defined by behavior rather than being derived simply from sensory stimulation. Reward value is often subjective, as shown by temporal discounting (Roesch et al., 2007; Kobayashi and Schultz, 2008), and includes probability that reflects the frequency of occurrence rather than direct sensory stimulation (Fiorillo et al., 2003). Therefore, the observed neuronal adaptations to probability distributions of reward value may constitute a proper mechanism of reward systems.
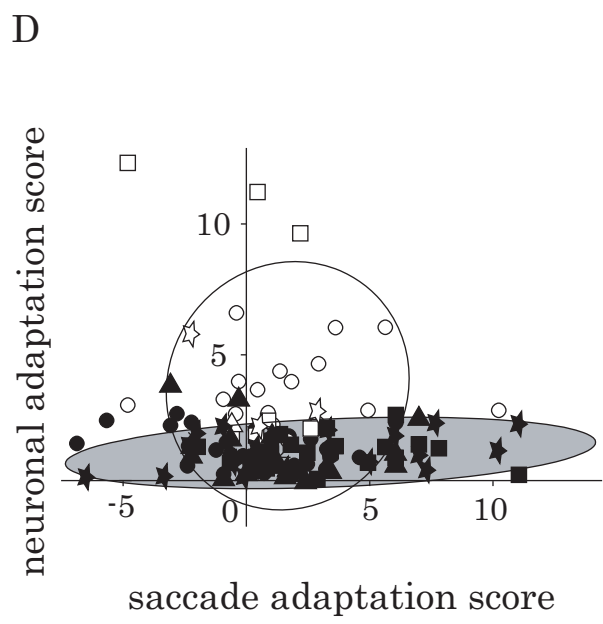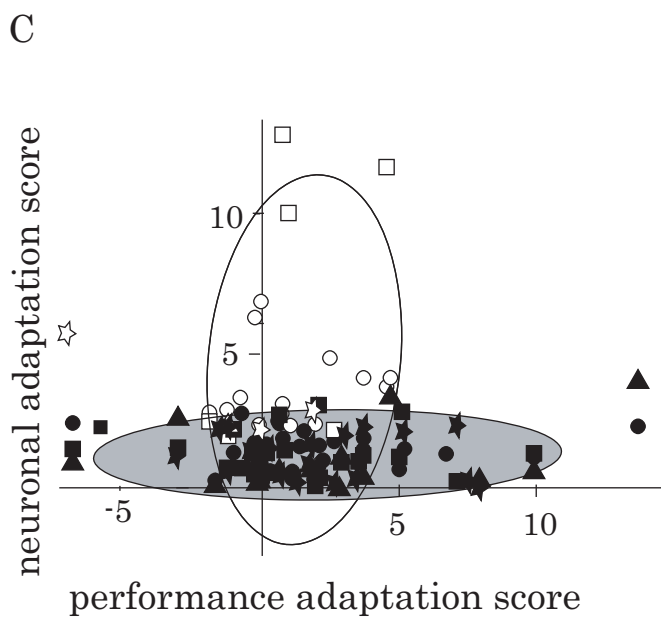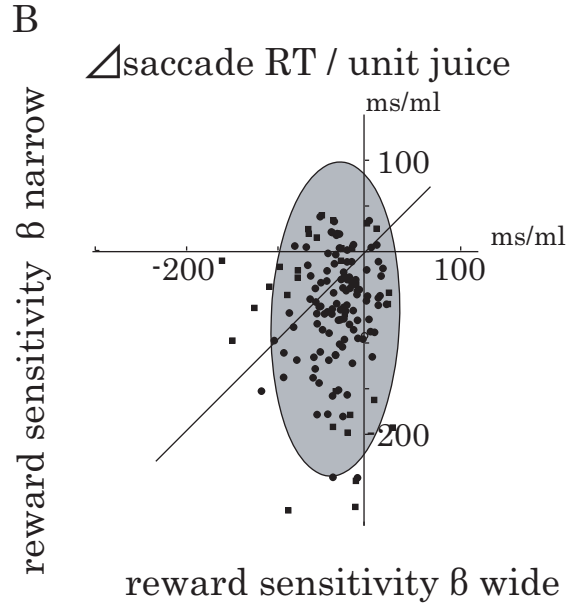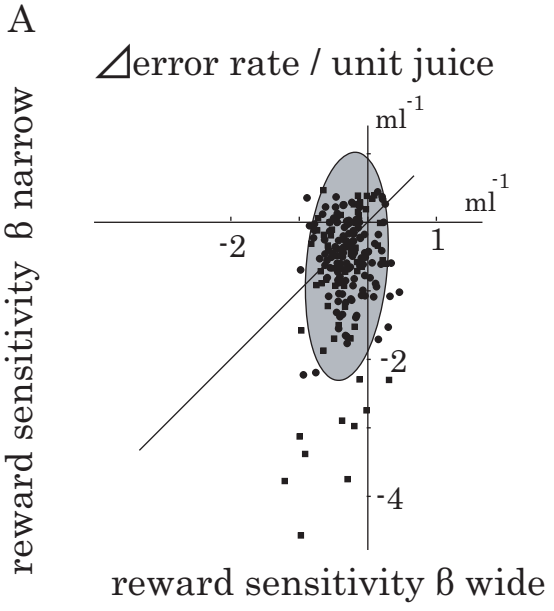
Value coding without adaptation

A sizeable fraction of orbitofrontal neurons maintained response slopes constant despite changes in reward distribution. These neurons reflected reward value by changes in discharge rate on a fixed scale. As a result, value discrimination and transmitted information decreased with more narrow value distributions. Substantial nonadaptive coding occurred also in primary sensory systems (Hosoya et al., 2005; Maravall et al., 2007). Nevertheless, nonadaptive coding provides several advantages for reward coding. It assesses reward value on a constant scale that allows comparisons and transitivity across the wide range of possible reward values regardless of specific distributions (Padoa-Schioppa and Assad, 2008). Furthermore, nonadaptive coding may serve as anchoring reference for adaptive coding. It may also constitute a substrate for ensemble coding of a wide range of reward values in subpopulations of neurons devoted to particular reward aspects (Schoenbaum and Eichenbaum, 1995; van Duuren et al., 2008; Simmons and Richmond, 2008). There may be two mechanisms

for the coding of reward value in orbitofrontal cortex; nonadaptive coding that allows reference of reward values across a wide range and regardless of context, and adaptive coding that allows optimal discrimination within contexts defined by specific probability distributions.

References

Barlow H (1961) Possible principles underlying the transformation of sensory messages. In: Sensory communication (Rosenblith WA, ed), pp 217–234. Cambridge, MA: MIT.

Baylis LL, Gaffan D (1991) Amygdalectomy and ventromedial prefrontal ablation produce similar deficits in food choice and in simple object discrimination learning for an unseen reward. Exp Brain Res 86:617–622.

Brenner N, Bialek W, de Ruyter van Steveninck R (2000) Adaptive rescaling maximizes information transmission. Neuron 26:695–702.

Cromwell HC, Hassani OK, Schultz W (2005) Relative reward processing in primate striatum. Exp Brain Res 162:520–525.

Dean I, Harper NS, McAlpine D (2005) Neural population coding of sound level adapts to stimulus statistics. Nat Neurosci 8:1684–1689.

DeShon RP, Alexander RA (1996) Alternative procedures for testing regression slope homogeneity when group error variances are unequal. Psychol Methods 1:261–277.

Dias R, Robbins TW, Roberts AC (1996) Dissociation in prefrontal cortex of affective and attentional shifts. Nature 380:69–72.

Dunn FA, Lankheet MJ, Rieke F (2007) Light adaptation in cone vision involves switching between receptor and post-receptor sites. Nature 449:603–606.

Fairhall AL, Bialek W (2002) Adaptive spike coding. In: The handbook of brain theory and neural networks (Arbib MA, ed), pp 90–94. Cambridge MA, MIT.

Fairhall AL, Lewen GD, Bialek W, de Ruyter Van Steveninck RR (2001) Efficiency and ambiguity in an adaptive neural code. Nature 412:787–792.

Fiorillo CD, Tobler PN, Schultz W (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. Science 299:1898–1902.

Gershon ED, Wiener MC, Latham PE, Richmond BJ (1998) Coding strategies in monkey V1 and inferior temporal cortices. J Neurophysiol 79:1135–1144.

Golomb D, Hertz J, Panzeri S, Treves A, Richmond B (1997) How well can we estimate the information carried in neuronal responses from limited samples? Neural Comput 9:649–665.

Hikosaka K, Watanabe M (2000) Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. Cereb Cortex 10:263–271.

Hosokawa T, Kato K, Inoue M, Mikami A (2007) Neurons in the macaque orbitofrontal cortex code relative preference of both rewarding and aversive outcomes. Neurosci Res 57:434–445.

Hosoya T, Baccus SA, Meister M (2005) Dynamic predictive coding by the retina. Nature 436:71–77.

Iversen SD, Mishkin M (1970) Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. Exp Brain Res 11:376–386.

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

Kobayashi S, Lauwereyns J, Koizumi M, Sakagami M, Hikosaka O (2002) Influence of reward expectation on visuospatial processing in macaque lateral prefrontal cortex. J Neurophysiol 87:1488–1498.

Laughlin S (1981) A simple coding procedure enhances a neuron's information capacity. Z Naturforsch C 36:910–912.

Laughlin S, Hardie RC (1978) Common strategies for light adaptation in the peripheral visual systems of fly and dragonfly. J Comp Physiol 128:319–340.

Maravall M, Petersen RS, Fairhall AL, Arabzadeh E, Diamond ME (2007) Shits in coding properties and maintenance of information transmission during adaptation in barrel cortex. PLoS Biol 5:e19.

Padoa-Schioppa C (2009) Range-adapting representation of economic value in the orbitofrontal cortex. J Neurosci 29:14004–14014.

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

Padoa-Schioppa C, Assad JA (2008) The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. Nat Neurosci 11:95–102.

Roesch MR, Calu DJ, Schoenbaum G (2007) Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. Nat Neurosci 10:1615–1624.

Schoenbaum G, Eichenbaum H (1995) Information coding in the rodent prefrontal cortex. II. Ensemble activity in orbitofrontal cortex. J Neurophysiol 74:751–762.

Simmons JM, Richmond BJ (2008) Dynamic changes in representations of preceding and upcoming reward in monkey orbitofrontal cortex. Cereb Cortex 18:93–103.

Thorpe SJ, Rolls ET, Maddison S (1983) The orbitofrontal cortex: neuronal activity in the behaving monkey. Exp Brain Res 49:93–115.

Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. Science 307:1642–1645.

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

Treves A, Panzeri S (1995) The upward bias in measures of information derived from limited data samples. Neural Comput 7:399–407.

van Duuren E, Lankelma J, Pennartz CM (2008) Population coding of reward magnitude in the orbitofrontal cortex of the rat. J Neurosci 28:8590–8603.

Wainwright MJ (1999) Visual adaptation as optimal information transmission. Vision Res 39:3960–3974.

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

Wark B, Lundstrom BN, Fairhall A (2007) Sensory adaptation. Curr Opin Neurobiol 17:423–429.

A

⊿error rate / unit juice

reward sensitivity ß narrow

ml⁻¹
ml⁻¹
-2
1
-2
-4

reward sensitivity ß wide

B

⊿saccade RT / unit juice

reward sensitivity ß narrow

ms/ml
100
ms/ml
-200
100
-200

reward sensitivity ß wide

C

neuronal adaptation score

10
5
-5
5
10

performance adaptation score

D

neuronal adaptation score

10
5
-5
0
5
10

saccade adaptation score

Fig. S1 Kobayashi et al.

Supplementary material

Behavioral adaptation to reward distributions

We evaluated adaptation of animal behavior to reward distributions by quantitatively measuring the sensitivity of behavioral responses to the predicted juice volume in two reward distributions. First, we analyzed the errors made after the value cue onset (premature fixation breaks and inaccurate saccades). We subjected the error rate data during each recording session to the same linear regression model as used for the neuronal responses (equations 1, 2 in the main text) and obtained a pair of regression slopes from two reward distributions. The regression slopes were mostly negative for both β_{narrow} (ordinate) and β_{wide} (abscissa, Fig. S1A), indicating that the animals made fewer errors when the value cue predicted the larger juice volume. Importantly, the slopes were significantly steeper in the narrow as compared with wide distribution ($p < 0.001$, 2-tailed paired t-test).

We also analyzed the saccade reaction time to the value cue in correctly performed trials. Saccade reaction time served as the dependent variable Y in the linear regression model (equation 1, 2 in the main text).

The regression slope was mostly negative, indicating shorter reaction times after prediction of larger juice volumes (Fig. S1B). Importantly, regression slopes were generally steeper in the more narrow reward distribution ($p < 0.001$, 2-tailed paired t-test).

These results suggest that behavioral measures of both error rate and saccade reaction time were graded by the predicted size of reward. The reward impact on these behavioral responses was normalized to the reward distributions, indicating behavioral adaptation.

Relationships between behavioral and neuronal adaptations

A sizeable fraction of the sampled neurons failed to adapt to the changes in reward distributions (Fig 4E, F; see also Table S1). However, it is possible that those neurons did have the capacity to adapt but did not exhibit it because the behavior of the animal lacked adaptation during those particular recordings. To examine this possibility, we analyzed the relationship between behavioral and neuronal adaptations by using the adaptation scores. We defined behavioral adaptation score as the ratio between the two regression slopes (β_{narrow} / β_{wide}) based on error rate and

saccade reaction time analyses, analogous to the neuronal adaptation score (see Materials and Methods in the main text). Figure S1C, D plots neuronal versus behavioral adaptation scores (C, error rate; D, saccade reaction time). As expected, neuronal adaptation scores of nonadaptive neurons were low (filled symbols). However, behavioral adaptation scores corresponding to these nonadaptive neurons were variable as shown by the horizontal stretch of their distributions (gray ellipses). Therefore, lack of adaptation in these neurons appears to have little direct relation to behavioral adaptation. The result supports the view that orbitofrontal cortex initially encodes the reward value objectively (filled symbols), and adaptive processes modify the coding into more subjective value expressed in animal behavior (open symbols).

Positive identification of adaptive and nonadaptive responses by hierarchical regression models

The main regression analysis classified responses as adaptive when their regression slopes differed significantly between the two reward distributions (equations 1 and 2 in the main text, Table 1). However, insignificant slope

change would not necessarily indicate significant lack of adaptation. To test neuronal responses for significant nonadaptation, we used the following models:

Full model: $Y = \alpha + \beta_1(X - \mu) + \beta_2(X - \mu) / \sigma$ (S1)

Nonadaptive reduced model: $Y = \alpha + \beta(X - \mu)$ (S2)

Adaptive reduced model: $Y = \alpha + \beta(X - \mu) / \sigma$ (S3)

where Y is the discharge rate in each task period, X is an independent variable that represents reward volume, σ is standard deviation (SD) and μ is mean in each reward distribution, and α, β, β_1, β_2 are coefficients. The models are constrained to the point (μ, α), because μ is the common juice volume between the two reward distributions, and value-coding neurons should produce similar activity (α) at this point. Equation S3 models the statistical z-score.

Using a hierarchical approach we compared the full model against a reduced model with the F test ($p < 0.05$) (Snedecor and Cochran, 1989). If neuronal activity adapted to the SD of reward distributions, the full model (S1) should fit the data significantly better than the nonadaptive reduced

model (S2) lacking the adaptive term ($\beta_2 \times X / d$). By contrast, if neuronal

activity reflected the physical juice volume without adapting to the predicted

standard deviation, the full model (S1) should fit the data better than the

adaptive reduced model (S3) lacking the physical volume term ($\beta_1 \times X$).

Responses that showed neither significant adaptation nor significant

nonadaptation were classified as unassigned.

The results of this analysis replicated the results of the main

regression model, showing almost identical numbers of adaptive responses in

each task period (Table S1, compare with Table 1). As additional result of the

hierarchical approach, a number of significantly nonadaptive responses was

positively identified (Table S1). These explicitly nonadaptive responses were

slightly less frequent than the adaptive responses, particularly in the earlier

task periods.

	Cue	Delay	Saccade	Reward	Total
# value-coding responses	63	40	19	27	149
Adaptive responses (% of value-coding responses)	23 (36.5%)	13 (32.5%)	2 (10.5%)	4 (14.8%)	42 (28.2%)

Nonadaptive responses	12	6	7	8	33
(% of value-coding responses)	(19.0%)	(15%)	(36.8%)	(29.6%)	(22.1%)
Unassigned responses	28	21	10	15	74
(% of value-coding responses)	(44.4%)	(52.5%)	(52.6%)	(55.6%)	(49.7%)

Table S1. Adaptation of value-coding activity to reward distributions tested by hierarchical regression methods (equations S1–S3). The table shows explicit adaptive responses (equations S1 versus S2, $p < 0.05$), explicit nonadaptive responses (equations S1 versus S3, $p < 0.05$, F test), and unassigned value coding responses without significant adaptive and nonadaptive properties ($p > 0.05$).

Adaptation assessed by regression with interaction term

As a separate test to examine the effects of reward distributions on response slopes, we constructed the following two models

$$Y = \beta_0 + \beta_1 X + \beta_2 M \qquad (S4)$$

$$Y = \beta_0 + \beta_1 X + \beta_2 M + \beta_3 X{\cdot}M \qquad (S5)$$

where Y is the dependent variable (neuronal response), X is an independent variable (juice volume), and X belongs to two dichotomous groups M (low

and high SDs, modeled as 0 and 1, respectively). Equation S5 has an interaction term between juice volume and reward distribution. Neuronal adaptation to SD would become manifest as interaction between juice volume and distribution (the $X \cdot M$ product term in equation S5) (Zar, 1999). When β_3 is significant and equation S5 gives significantly better fit than equation S4 ($p < 0.05$, F test), we concluded that interaction is significant hence adaptation. This test resulted in very similar numbers of significant adaptations of neuronal responses (Table S2) as the main regression model with normalized-t approximation (Table 1) and the hierarchical regression model (Table S1).

	Cue	Delay	Saccade	Reward	Total
# value-coding responses	63	40	19	27	149
Adaptive responses (% of value-coding responses)	22 (34.9)	13 (32.5)	2 (10.5)	6 (22.2)	43 (28.9)

Table S2 Adaptation of value-coding activity to reward distributions tested by the interaction regression model (equations S4 and S5).

Supplementary figure legends

Figure S1. Behavioral and neuronal adaptations **A**, **B**, Behavioral adaptation to predicted reward distributions. Plots of regression slopes for small versus large standard deviations (SD) (ordinate versus abscissa). Slopes (β) of error rate (**A**, $\triangle$error rate / ml) and saccade reaction time (**B**, $\triangle$ms / ml) were estimated by linear regression models (equations 1, 2 in the main text) for each behavioral test (circles, animal A; squares, animal B). Most symbols were found in the lower left quadrants, reflecting negative relationships between the behavioral measures and the predicted juice volume. Shaded ellipses delineate the distribution contours (2SD) of the plots. **C, D,**

Relationships between behavioral and neuronal adaptation scores. Neuronal adaptation scores are plotted against behavioral adaptation scores for task performance (error rate) (**C**) and saccade reaction time to the value cue (**D**). Open symbols, adaptive neurons; filled symbols, nonadaptive neurons. Shapes of the symbols indicate the task period in which neuronal adaptation scores were measured (circles, value cue period; squares, delay period; triangles, saccade period; asterisks, reward period). Shaded and open ellipses delineate the distribution contours (2SD) of nonadaptive and adaptive neurons, respectively.

Supplementary references

Snedecor GW, Cochran WG (1989) Statistical Methods (chapter 17.4). Eighth Edition. Ames: Iowa State University Press.

Zar JH (1999) Biostatistical Analysis (chapter 20.12). Fourth Edition. Upper Saddle River, NJ: Prentice Hall.