Behavioral/Systems/Cognitive

Single-Cell and Population Coding of Expected Reward Probability in the Orbitofrontal Cortex of the Rat

Esther van Duuren,[1,2] **Geoffrey van der Plasse,**[3] **Jan Lankelma,**[1] **Ruud N. J. M. A. Joosten,**[2] **Matthijs G. P. Feenstra,**[2] **and Cyriel M. A. Pennartz**[1]

[1]Swammerdam Institute for Life Sciences, Center for Neuroscience, University of Amsterdam, 1098 SM Amsterdam, The Netherlands, [2]Netherlands Institute for Neuroscience, 1105 BA Amsterdam, The Netherlands, and [3]Department of Anatomy and Neuroscience, Vrije Universiteit University Medical Center, 1081 BT Amsterdam, The Netherlands

The orbitofrontal cortex (OFC) has been implicated in decision-making under uncertainty, but it is unknown how information about the probability or uncertainty of future reward is coded by single orbitofrontal neurons and ensembles. We recorded neuronal ensembles in rat OFC during an olfactory discrimination task in which different odor stimuli predicted different reward probabilities. Single-unit firing patterns correlated to the expected reward probability primarily within an immobile waiting period before reward delivery but also when the rat executed movements toward the reward site. During these pre-reward periods, a subset of OFC neurons was sensitive to differences in probability but only very rarely discriminated on the basis of reward uncertainty. In the reward period, neurons responded during presentation or omission of reward or during both types of outcome. At the population level, neurons were characterized by a wide divergence in firing-rate variability attributable to expected probability. A population analysis using template matching as reconstruction method indicated that OFC generates a distributed representation of reward probability with a weak dependence on neuronal group size. The analysis furthermore confirmed that predictive information coded by OFC populations was quantitatively related to reward probability, but not to uncertainty.

Introduction

One of the key factors in decision-making is the probability of future rewards resulting from voluntary actions. Behavioral studies in humans have shown that a certain reward is generally preferred over an uncertain or probabilistic reward of the same amount: in a process called probability (or odds) discounting, the value of probabilistic rewards is degraded as the reinforcer becomes less probable (Rachlin et al., 1991). The choice between small, likely rewards and large, unlikely rewards has been found to activate the orbitofrontal cortex (OFC) (Rogers et al., 1999; Ernst et al., 2004), an area of the prefrontal cortex that has been strongly implicated in the assessment of reward value (O'Doherty et al., 2001, 2003) and in the planning of actions leading to immediate rewards (Tanaka et al., 2004). Additional support for an involvement of the OFC in risky decision-making comes from studies with humans suffering orbitofrontal damage. These patients perform poorly on tasks involving uncertainty, such as the Iowa gambling task, by continuing to choose high-risk decks of cards, whereas normal subjects bias their choice behavior toward low-risk decks (Bechara et al., 1996, 1997). Furthermore, recent brain imaging studies suggest a positive correlation between orbitofrontal activity and the unpredictability of reward (Hsu et al., 2005). Thus far, rodent studies have produced somewhat conflicting results on the precise role of OFC in risky decision-making. Using a probability discounting paradigm, Mobini et al. (2002) demonstrated that orbitofrontal-lesioned rats preferred the smaller, more probable reinforcer over the larger, but infrequent reward. In contrast, Pais-Vieira et al. (2007), using an alternative probability discounting paradigm more similar to the gambling tasks used in humans, showed that animals with orbitofrontal lesions preferred the larger but less probable reward, which is in accordance with stronger risk-taking behavior as demonstrated by data from patients with prefrontal lesions. The contrasting results of these two rodent studies are likely caused by differences in experimental design and methods, but altogether, these animal studies and the human imaging data do suggest that the OFC is involved in assessing the value of rewards on the basis of their probability.

Despite this body of evidence, it is still unknown whether and how value based on reward probability is coded in OFC and whether coding based on probability is distinguishable from coding based on uncertainty. Studying this specific topic may also shed light on the more general question how neuronal populations represent the probability of any behaviorally relevant variable, be it sensory, motor, or motivational (Knill and Pouget, 2004; Daw et al., 2005). To examine how the firing activity of orbitofrontal neurons is affected by a varying future reward prob-

ability, we trained rats on an olfactory discrimination task in which odors were predictive of the probability of a pellet reward. During this task, we examined both single-unit and population activity and expected neural activity predictive of future reward probability to be found in multiple task periods (Schoenbaum et al., 1998; van Duuren et al., 2007).

Materials and Methods

Subjects
All experiments were approved by the Animal Experimentation Committee of the Royal Netherlands Academy of Arts and Sciences and were performed in accordance with the National Guidelines for Animal Experimentation. Data were collected from four male Wistar rats (Harlan), weighing 375–425 g at the time of surgery. Animals were socially housed in standard type 4 Makrolon cages, weighed and handled daily, and kept under a reversed 12 h light/dark cycle (dimmed red light at 7:00 A.M.). Animals were maintained on 90% of their free-feeding body weight (16 g of standard rat food chow per day per rat), with water available *ad libitum*. After surgery, the animals were housed individually in a larger cage (1 × 1 × 1 m) under the same conditions.

Behavior
Apparatus. The recording chamber (40 × 37 × 41.5 cm), placed in a sound-attenuated and electrically shielded box, had a black interior with straight walls. The front panel contained on the right side a light signaling trial onset and an odor sampling port, and the left side had a food trough. Behavioral events and data collection were controlled and registered by a computer. Both sampling port and food trough contained an infrared beam transmitter and detector port inside to detect the responses made by the animals. Odor delivery was controlled by a system of solenoid valves and flow meters (van Duuren et al., 2007) with separate delivery lines for each odor to prevent mixture of odors in the system. Two pellet dispensers were present (ENV-203 Magazine Type, 45 mg; MED Associates): one for pellet delivery (45 mg sucrose pellets; Bioserve Biotechnology) and one empty dispenser used to mimic the sound of the dispenser during unrewarded trials. The odorants (Tokos BV) were separated into different families, i.e., fruity, floral, herbal, woody, and citrus. For each discrimination session, four distinct odors were used, each odor from a different family. Furthermore, no single family of odors was preferentially associated with a particular trial outcome.

Behavioral paradigm. After habituation, animals were progressively trained on the behavioral procedure of the four-odor probability task. Four new odors were used in each discrimination session, each odor associated with a specific reward probability, i.e., $p = 100\%$, $p = 75\%$, $p = 50\%$, and $p = 0\%$. Initially, $p = 25\%$ was inserted in the task paradigm as well, but we removed this reward condition because animals were less willing to perform the task, which was probably caused by the lower availability of rewarded trials. Animals were first trained to make a nose poke in the odor sampling port, which was sufficient to immediately obtain reinforcement by visiting the food trough. In the next phase, animals learned to make an odor poke with a minimal duration of 1.5 s. In the final stage of shaping, a waiting period of 1.5 s was introduced during the poke in the food trough and before the pellet was delivered.

Once animals were familiar with the behavioral procedure of the task, two different four-odor discrimination problems were consecutively presented to the animal to provide additional training. After rats learned new odor–reward probability associations (as visible by withholding responses toward or at the food trough after sampling the odor predictive for the null probability), they were implanted with a head stage containing an array with individually movable tetrodes ("hyperdrive"), and recordings started. During each recording session, a new set of four odors was presented, which were chosen pseudorandomly. During the task, trial onset was indicated by the trial light switching on, after which the animal had 15 s to make an odor poke. If no odor poke was made, the trial light turned off, and the intertrial interval (ITI) (with a variable duration of 10–25 s) started. Whenever a prolonged odor poke was made, the trial light switched off after 0.25 s, followed 0.25 s later by the presentation of an odor. This period was included to prevent the animal from moving during cue sampling. Odor sampling itself was required to last at least 1 s. After retraction of the animal's nose out of the odor sampling port or whenever a maximal duration for odor sampling (10 s) was exceeded, odor presentation was terminated. Premature retraction from the odor sampling port (odor pokes shorter than the minimal duration of 1.5 s) resulted in the start of the intertrial interval. After the waiting period in the food trough of 1.5 s, a pellet was delivered during the reward trials, and, 5 s later, the intertrial interval started. The behavioral sequence comprising the departure from the sampling port to the food trough, including nose entry and waiting period in the food trough, will be referred to as the "go" response.

Surgery and electrophysiology
Animals were anesthetized with 0.08 ml/100 g Hypnorm intramuscularly (0.2 mg/ml fentanyl and 10 mg/ml fluanison) and 0.04 ml/100 g Dormicum subcutaneously (midazolam 1 mg/kg) and mounted in a David Kopf Instruments stereotaxic frame. After exposure of the cranium, five small holes were drilled to accommodate surgical screws, one of which served as ground. Another hole was drilled over the OFC in the left hemisphere (center of the hole was 3.6 mm anterior, 3.2 mm lateral to bregma according to Paxinos and Watson, 2005). The dura was opened, and the exit bundle of the hyperdrive was lowered onto the exposed cortex, after which the hole was filled with a silicone elastomer (Kwik-Sil; World Precision Instruments), and the hyperdrive was anchored to the screws with dental cement. The hyperdrive, which was custom built, contained an array of 12 individually drivable tetrodes and two reference electrodes (13 μm nichrome wire; Kanthal), spaced apart by at least 310 μm (Gray et al., 1995; Gothard et al., 1996). Immediately after surgery, all tetrodes and reference electrodes were advanced 1 mm into the brain; in the course of the next 3 d, the tetrodes were gradually lowered until the OFC was reached. Animals were allowed to recover at least 7 d before the start of the recordings. To record different units during each recording session, all tetrodes were lowered at the start of a recording day with increments of 40 μm. Once the tetrodes were lowered, the animal was left to rest in his home cage for at least 2 h in view of unit recording stability, after which the experimental session started.

Electrophysiological recordings were performed using a Cheetah recording system (Neuralynx). Signals from the individual leads of the tetrodes were passed through a low-noise unity-gain field-effect transistor preamplifier, insulated multiwire cables, and a 72 channel commutator (Dragonfly) to digitally programmable amplifiers (gain, 5000 times; bandpass filtering, 0.6–6.0 kHz). Amplifier output was digitized at 32 kHz and stored on a Windows NT station. The occurrence of task events in the behavioral chamber was recorded simultaneously.

After finishing experiments with a given rat, tetrode positions were marked by passing a 10 s, 25 μA current through one of the leads of each tetrode. Animals were perfused transcardially ~24 h after the lesions were made, using a 0.9% saline solution followed by 10% Formalin. After removal from the skull, brains were stored in a 10% Formalin solution for several days before sectioning. Brain sections of 40 μm were cut using a vibratome and were Nissl-stained to reconstruct the tracks and final positions of the tetrodes. This showed that recording sites ranged from 2.7 to 4.7 mm anterior to bregma and were limited to the ventral and lateral orbital regions of the OFC. Recording depth ranged from approximately −3 to −5.5 mm (Paxinos and Watson, 2005) (Fig. 1).

Data analysis
Behavior. Behavioral data was analyzed using SPSS for Windows (version 11.0). Unless otherwise stated, results are expressed as mean ± SEM values. "Movement time" was defined as the interval between nose retraction from the odor port and nose entry into the food trough, whereas the "overall response time" was defined as the duration of the behavioral sequence starting with odor sampling and ending with the nose poke in the food trough. The mean response times per reward probability were obtained from all trial types associated with a particular probability from all sessions. These measures were compared across different trial types with the nonparametric Kruskal–Wallis test ($p < 0.05$), followed by a *post hoc* Mann–Whitney *U* test ($p < 0.05$).
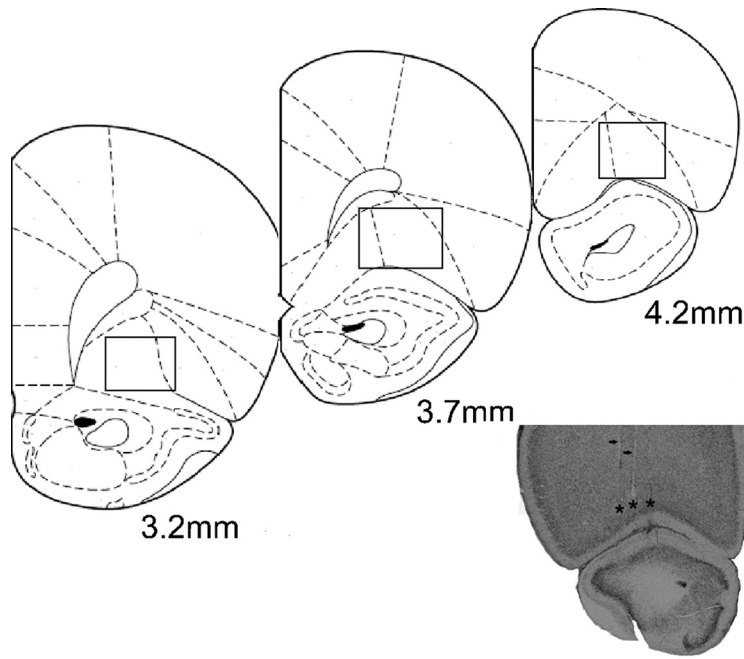
Figure 1. Localization of tetrode recording sites. As indicated by rectangles, recordings in all rats were localized in the ventral and lateral regions of the OFC, between 2.7 and 4.7 mm anterior from bregma. Recording depth ranged from approximately −3 to −5.5 mm (Paxinos and Watson, 2005). Several tetrode tracks are visible, as indicated by the black arrows. Black asterisks mark the lesion sites showing the final position of three tetrodes.

Variability in the population code for reward probability. To examine the variability in responses within the population toward reward probability in more detail, we calculated two different measures of response variability. These measures have been used previously to indicate sparseness of neural coding (Rolls and Tovee, 1995; Perez-Orive et al., 2002). Parameter variability (V_{par}), which is indicative of the response variability of a single cell attributable to differences in reward probability, was calculated as follows:

$$V_{par} = \frac{N}{N-1} \frac{\overline{r^2} - (\bar{r})^2}{\overline{r^2}},\qquad (2)$$

with

$$\bar{r} = \frac{1}{N} \sum_{j=1}^{N} r_j$$

and

$$\overline{r^2} = \frac{1}{N} \sum_{j=1}^{N} r_j^2$$

where N indicates the total number of reward probabilities ($N = 3$, namely 50, 75, or 100%; the $p = 0\%$ condition was not part of this analysis because of the lower amount of trials, but see below), and r_j is the mean firing rate per cell per probability. V_{par} assumes a value of 1 when the cell under scrutiny fires selectively for only one reward probability and does not fire at all for the other two conditions, whereas $V_{par} = 0$ when spike counts are equal for all three reward probabilities. In addition, we calculated the population variability (V_{pop}), which is indicative of the variability in the mean firing rate of single cells across the population, regardless of the probability of reward. This measure was calculated in a similar manner, but r_j now indicates the mean firing rate of neuron j during a particular trial phase, averaged across all three reward probabilities, and N is the number of units recorded in a given session. Thus, $\bar{r}$ now represents the mean firing rate in the population and $\overline{r^2}$ is the mean squared firing rate. Values ranged between 0 and 1, with 1 representing the maximal variability attainable.

Ensemble analysis of reward probability coding. Representation of expected reward probability by ensemble activity was examined using template matching as reconstruction method. For an extensive explanation of this method, see van Duuren et al. (2008). Briefly, the aim of this analysis was to reconstruct or decode the probability of reward from spike trains recorded from an ensemble of OFC neurons, because usually multiple spike trains provide more information about an encoded variable than a spike train generated by a single neuron. In this procedure, the series of firing-rate values of all cells of the ensemble are conceptualized as a population vector for each reward probability, containing the spike counts of each cell within a specified trial phase (e.g., the waiting period). From each session, two population vectors were created, denominated as $\mathbf{x} = (x_1, x_2, \ldots x_N)$ and $\mathbf{y} = (y_1, y_2, \ldots y_N)$, with x_i and y_i indicating the spike count of cell i averaged across trials. One vector is used for the encoding part of the procedure, which determines the "template," i.e., the response profiles or "tuning curves" of the cells toward reward probability. These response profiles consist of a list of the spike counts of all cells pertaining to the different reward conditions. The other vector is used for the decoding part of the procedure, in which the spike counts, specific for reward probabilities, are taken from the same cells but now from the first part of the session. The decoding vector is then compared with the encoding vector. Thus, these vectors are used to calculate the decoding score, which is the percentage of correctly identified reward

Single units. Single units were isolated and analyzed as described previously by van Duuren et al. (2007). In short, spike sorting was done offline using standard cluster cutting procedures (BBClust/MClust 3.0). Perievent time histograms were constructed to examine correlations between events in the task and changes in firing rate. Neural responses during trials were statistically assessed with the nonparametric Wilcoxon's matched-pairs signed-rank (WMPSR) test ($p < 0.01$). These task-related neural correlates were considered significant if firing rates during trials, quantified per bin, were significantly different from a fixed control (baseline) period during the intertrial interval. This control period consisted of five consecutive bins, and any of the bins in the trial period tested for a significant change in firing was required to differ significantly from each of these five control bins. In addition, responses had to be significant for two bin sizes to be considered as such. These bin sizes were 100 and 1500 ms for all task periods, except for the movement period, for which 100 and 700 ms were adopted (700 ms corresponded to the approximate mean duration of this period). The use of a nonparametric test avoids the assumption of normally distributed spike counts inherent to parametric tests (cf. van Duuren et al., 2007). Once the WMPSR test indicated a significant deviation in firing rate with respect to baseline, the nonparametric Kruskal–Wallis test ($p < 0.05$) and a *post hoc* Mann–Whitney U test ($p < 0.05$) were used to compare the different PETHs pertaining to trial types with different reward probabilities.

Given the proportion of cells that not only shows a significant firing response during trials but also a modulation by reward probability (determined at $p < 0.05$), we used the following method to assess whether this proportion of the total number of task-related cells tested ($n = 78$) (see Results) is attributable to chance or not. When determining what the probability is that the true proportion (π) of modulated neurons is actually 5% or smaller, given a proportion P found in the sample, the confidence interval for proportions is determined by the following:

$$\pi = P \pm Z \sqrt{\frac{P(1-P)}{n}}\qquad (1)$$

where Z is the Z-score, and n is the size of the sample tested (Sokal and Rohlf, 1995).

probabilities in the decoding phase, based on the activity patterns found in the encoding phase. Hence, the ensemble code for reward probability is made up of the different firing rates of all recorded cells combined in the encoding and decoding phase in relation to reward probability. Note that, besides mean firing rate per trial phase, other aspects of firing patterns, such as related to spike timing, may in general make additional contributions to ensemble coding (cf. Narayanan et al., 2005).

Template matching was used as described previously by Lehky and Sejnowski (1990) and Zhang et al. (1998). The similarity ("matching") between the two vectors containing the spike count in the defined time window for the encoding and decoding block was calculated by computing the cosine of the angle between them. A value of 1 represents an exact similarity between the two vectors and −1 is the exact opposite, whereas 0 (i.e., orthogonal) indicates no similarity between the two vectors. First, the inner product of **x** and **y** was calculated as follows:

$$\sum_{i=1}^{N} x_i y_i \qquad (3)$$

where x_i and y_i indicate the average firing rate of neuron i from a total of N cells within the specified time window for the encoding and decoding block, respectively.

The cosine value was calculated as follows:

$$\cos \theta = \frac{\sum_{i=1}^{N} x_i y_i}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \qquad (4)$$

with the denominator representing the product of the absolute vector lengths. Whenever the decoding spike vector belonging to a particular reward probability provided the highest cosine value with respect to the encoding vector, then that particular probability was selected as the reconstructed likelihood.

For our standard analysis, the initial three-quarters of the trials within a session was used for decoding, whereas the final one-quarter of the trials was used for encoding. A standard time window was used for the various trial phases for which reward probability was reconstructed, corresponding to the duration of that particular phase within the trial. The decoding time frame used for the period in which the animal moved from the odor sampling port to the food trough (the movement period) was 0.7 s, and the time frame for the waiting period at the food trough was 1.5 s. For the reward phase, the decoding time frame was 5 s, unless otherwise noted.

The decoding score was expressed as a function of time and of the size of the "reconstruction ensemble," i.e., the group of neurons that was subsampled from the entire population and used for the calculations. The maximum size of the reconstruction ensemble was 27, which corresponds to the lowest amount of cells recorded in the sessions used for this analysis. Thus, all ensembles used for our population coding study contained at least 27 units. Calculations were made for each recording session separately, after which decoding scores were averaged across sessions. For the assessment of decoding as a function of size of the reconstruction ensemble, the decoding score was calculated 100 times for each group size, each time with neurons randomly picked from the population recorded in that particular session. The decoding curves were analyzed further by applying linear regression analysis ($p < 0.05$) and a one-way ANOVA test with, if appropriate, a Bonferroni's correction ($p < 0.05$). Besides template matching, we applied Bayesian reconstruction as a method to study population coding (Lehky and Sejnowski, 1990; Zhang et al., 1998). The decoding performance obtained with this method, however, was generally similar or slightly lower than for template matching, and therefore these results will not be discussed here.

Results

Behavior
For the analysis, we used data from 19 recording sessions, obtained from four rats. Animals performed on average 32 trials for

each of the three highest reward probabilities ($p = 100\%$, 32.2 ± 2.2; $p = 75\%$, 32.6 ± 2.9; $p = 50\%$, 32.3 ± 2.1). For probability $p = 0\%$, animals performed significantly fewer trials compared with the other three probabilities, on average 9.0 ± 2.3 (paired sampled t test, for all three comparisons, $p < 0.001$; note that each odor–probability coupling was novel at the beginning of each session). Movement time (the interval between nose retraction from the odor port and nose entry into the food trough) showed no significant difference between the probabilities $p = 100$, 75, and 50% (respectively, 0.69 ± 0.01, 0.68 ± 0.01, and 0.72 ± 0.01 s), but the movement time for each of these reward probabilities was significantly shorter than for the null probability (1.15 ± 0.05 s). Furthermore, examination of the overall response time (the duration of the behavioral sequence starting with odor sampling and ending with the nose poke in the food trough) revealed that animals responded significantly faster on $p = 100$ and 75% trials compared with $p = 50\%$; no significant difference was found between $p = 100\%$ and $p = 75\%$ ($p = 100\%$, 2.43 ± 0.03 s; $p = 75\%$, 2.39 ± 0.03 s; and $p = 50\%$, 2.61 ± 0.04 s). Thus, learning within this task was evident from the shorter overall response time for the two highest reward probabilities, as well as from the lower amount of trials and slow responding for the $p = 0\%$ reward condition. The result that no differences in the movement time were observed except between the $p = 100$, 75, and 50% conditions versus the 0% condition is similar to the behavioral result from our recent study examining coding of expected reward magnitude, also based on olfactory cues (van Duuren et al., 2007). Also here, it was argued that moving to the reward site to obtain reward is a relatively fast, stereotyped behavior compared with sampling odors containing predictive information.

Electrophysiology
Single units: modulation of firing rate by reward probability
During the 19 recording sessions, a total of 541 single units was recorded in the OFC, with a firing rate of 1.30 ± 0.07 spikes/s (mean ± SEM). Of these 541 units, 136 (25%) showed 184 statistically significant responses during the task, which implies that a considerable proportion of cells exhibited more than one correlate. Task-related firing-rate modulations were observed in neurons that responded during sampling of odors ($n = 38$) (Fig. 2A), during the behavioral period in which animals moved from the odor sampling port toward the food trough ($n = 25$) (Fig. 2B), during the waiting period at the reward site ($n = 53$) (Fig. 2C), and after pellet delivery ($n = 69$) (Fig. 2D) (cf. Schoenbaum et al., 1998; Ramus and Eichenbaum, 2000). Notably, these firing-rate modulations were transient in nature and primarily restricted to one trial phase (except in cells with dual correlates). For instance, no odor-response cells were found that continued being active throughout the movement or waiting period, which argues against a working memory-like maintenance of odor information throughout the trial.

Examination of coding of expected reward probability focused on the two task periods in which predictive information was expected to be found, i.e., the movement and waiting period (Schoenbaum et al., 1998; van Duuren et al., 2007). Given that the mathematical concept of probability involves the very notion of expectation, we will use the adjective "expected" here as referring to the subjective state of the animal anticipating a trial outcome based on the preceding odor cue. Reward-predictive coding may also occur during odor sampling, but this task phase was not analyzed here because odor identity could not be dissociated from reward probability. During the waiting and movement pe-
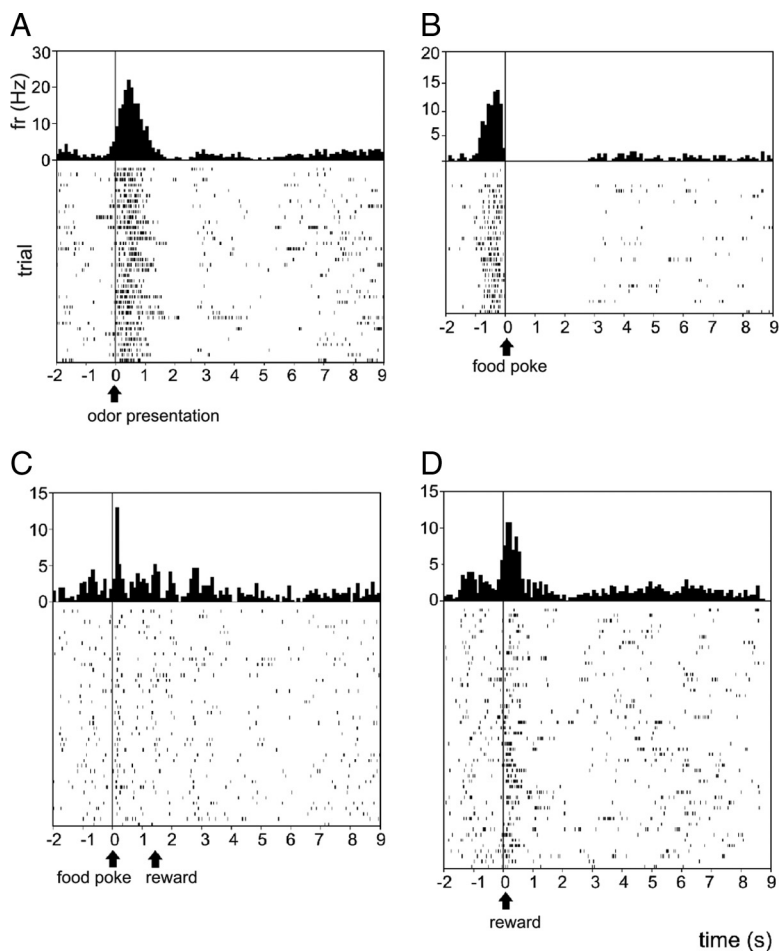
Figure 2. Overview of behavioral correlates of neural activity changes observed during task performance. Perievent time histograms and raster plots showing examples of the observed task-related behavioral correlates. Examples from four different units recorded in four different sessions demonstrating correlates related to the following: *A*, odor sampling (synchronized on onset of odor presentation during $p = 100\%$ trials); *B*, movement activity preceding nose entry into the food trough (synchronized on entry of the food trough during $p = 50\%$ trials); *C*, waiting period of 1.5 s with nose in the food trough, synchronized on onset of waiting; and *D*, pellet delivery (both during $p = 100\%$ condition). These as well as the following histograms (Figs. 3, 4) are presented with a bin size of 100 ms. In all raster plots, individual consecutive trials are represented as horizontal lines, with the first trial at the top row. Horizontal calibration denotes time (seconds), and vertical calibration denotes firing rate (fr) (hertz).

riods, a total number of 78 neurons demonstrated task-related activations (movement period, $n = 25$, 32%; waiting period, $n = 53$, 68%). Modulation of this activity by the probability of reward was found in 17 (22%) of these neurons. The proportion of probability-modulated cells was higher for the waiting period compared with the movement phase (waiting, 14 of 53 cells, 26%; movement, 3 of 25 cells, 12%). To assess whether the proportion of significantly probability-modulated neurons deviates from chance level, we computed confidence intervals for proportions (see Materials and Methods) and found that, given an overall 22% fraction on the total number of neurons subjected to testing ($n = 78$), the probability that the true proportion of probability-modulated neurons is actually 5% or smaller amounts to $p < 1.10^{-4}$ ($Z = 3.62$).

Figure 3, *A* and *B*, illustrates the response profiles of these neurons: activity was found to either increase or decrease with decreasing probability, or neurons displayed a peak or valley in firing rate at 75 or 50% probability. Thus, probability-related firing-rate modulations of OFC neurons are not always expressed as monotonic increments or decrements as a function of probability but may also reach peaks or troughs for intermediate values

of the tuning curve. All of these varieties of tuning may support the information-processing capacity of the OFC network using probability prospects to guide behavior. The response patterns during the movement phase could be identified as being task-related because in none of the 25 movement-related cells were they found when the same type of behaviors was displayed during the intertrial intervals (Fig. 3*C*).

Single units: firing-rate modulation by reward probability versus uncertainty
Although reward probability and uncertainty are considered fundamentally different outcome parameters (Fiorillo et al., 2003; Dreher et al., 2006; Tobler et al., 2007), they are intimately linked in the sense that with the two extreme probabilities ($p = 0$ and 100%) uncertainty is absent, whereas at intermediate probabilities uncertainty increases, being maximal in the $p = 50\%$ condition. Here, "probability" is defined as the numerical value quantifying the chance that a specified outcome of several possible outcomes will occur as a consequence of a course of events that is unpredictable to the animal. In contrast, "uncertainty" refers to the width of the probability distribution of outcomes and can to a first degree be expressed as the variance of this distribution. Whereas the expected reward value is considered a linear function of reward probability, uncertainty follows an inverted U-shaped function of probability, being minimal at $p = 0$ or 1 and maximal at $p = 0.5$ (cf. Schultz et al. 2008). If single OFC neurons predominantly code uncertainty and not probability, more OFC neurons will be found that discriminate in their pre-reward mean firing rate between trials with certain versus uncertain outcomes than OFC neurons discriminating on the basis of probability value. Of the 78 cells showing task-related activation in pre-reward periods, only one neuron significantly discriminated on the basis of the certain versus uncertain contrast. (Namely, the mean firing rate across a first pair of 100 and 0% trials was computed and plotted against the mean rate for temporally proximal trials with 50 and 75% probability, and so on for subsequent pairs of temporally proximal trials. Next, a sign test was applied across all plotted points, $p < 0.05$; one cell was found for movement period.) When, however, the contrast based on probability value was applied (namely 100 and 75% versus 0 and 50% reward probability), 17 significantly discriminating neurons were found (10 and 7 for the movement and waiting periods, respectively). The ratios applying to these two contrasts (1 of 78 vs 17 of 78) were significantly different at $p < 3.10^{-4}$ (ratio test). In conclusion, these results reveal that single OFC neurons preferentially discriminate on the basis of probability value, not uncertainty.

Single units: activity after reward delivery
For the reward delivery period, neural activity was examined for rewarded or unrewarded trials. During this period, 69 significant

neuronal responses (39%) were observed, which could be divided into three subgroups. The first subgroup consisted of 32 neurons (47%) that specifically responded during rewarded but not unrewarded trials [Fig. 4A: 25 (78%) and 7 (22%) neurons showed an increase or decrease in activity, respectively]. The second subgroup consisted of 30 neurons (44%), all showing a significant increase in firing during both rewarded and unrewarded trials (Fig. 4B). Within this group, 19 neurons (63%) demonstrated differential firing activity toward both the rewarded and unrewarded condition: responses were either larger for the rewarded condition (7 neurons, 37%) or unrewarded condition (12 neurons, 63%; Mann–Whitney U test, $p < 0.05$). For the remaining 11 neurons (37%), no difference in firing activity between the two conditions was found. The third group consisted of seven neurons (9%) that increased their firing activity during unrewarded but not rewarded trials (Fig. 4C). To examine whether this type of response occurs specifically in the context of task performance, we tested whether it also occurred during nose poking in the food trough during ITIs. Because no odor cues were provided and nose poking is presumably habit driven, a clear reward expectation may well be absent or at least less pronounced during the ITI period. Indeed, none of these neurons significantly increased their firing rate in the absence of reward in the intertrial interval (Fig. 4C). This indicates that the activity of these neurons does not reflect motor behavior associated with visiting and departing from the reward site but most likely reflects the omission of reward within the task context. Whether this phenomenon signifies an error in reward prediction or is related to attention and the saliency of an omission event is an issue that must await additional investigation.

*Variability of the representation of reward probability*

To examine the extent to which firing-rate variability throughout the various trial phases and across the recorded population is attributable to reward probability (50, 75, or 100%), we calculated a measure labeled "parameter variability" (V_{par}; alternatively called "parameter sparseness") and contrasted this to a different measure capturing the overall variability in firing rates across the population (V_{pop}, population variability), regardless of the influence of reward probability. The time windows used for this calculation were 0.7 s for the movement period and 1.5 s for both the waiting and reward delivery period. The mean V_{par}, which is indicative of the response variability of the individual neurons associated with variations in
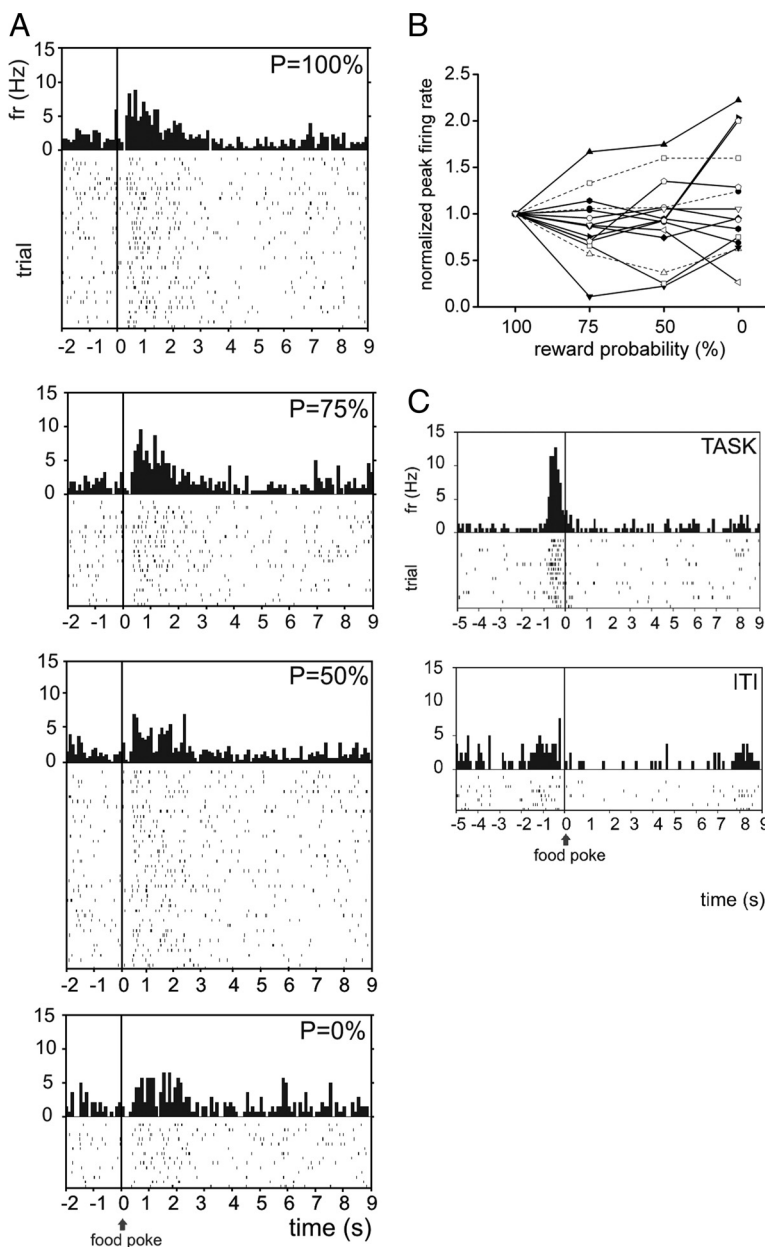
Figure 3. Differential firing in relation to reward probability. **A**, Example of a unit showing differential firing toward different reward probabilities during the waiting period. Activity is synchronized on nose entry into the food trough. Activity of this neuron decreases with decreasing reward probability: all four conditions differ significantly from each other except the $p = 100$ and 75% conditions. Fr, Firing rate. **B**, Overview of the significantly different firing rate profiles toward reward probability found during both the movement (dashed lines) and waiting (solid lines) periods. Different units are represented by different symbols. On the horizontal scale, reward probability (percentage) is plotted; the vertical scale displays the peak firing rate of individual units in association with different reward probabilities, normalized to the peak in the $p = 100\%$ condition. Note that only very few neurons exhibit distinct firing-rate peaks or valleys under maximal uncertainty ($p = 50\%$). **C**, Activity of a neuron showing activity during the movement period as part of task performance ($p = 75\%$ condition) and during the same behavior in the ITI.

reward probability, was 0.29, 0.22, and 0.27 for the movement, waiting, and reward delivery period, respectively (Fig. 5A–C). The mean V_{pop}, expressing the variability in firing rate across the population regardless of reward probability, was 0.76, 0.71, and 0.71 for the movement, waiting, and reward delivery phase, respectively (Fig. 5D–F). These results demonstrate that the overall population is marked by a high variability in firing rate, but the firing rates of individual OFC neurons are modulated by reward probability to a generally modest degree. However, in all three
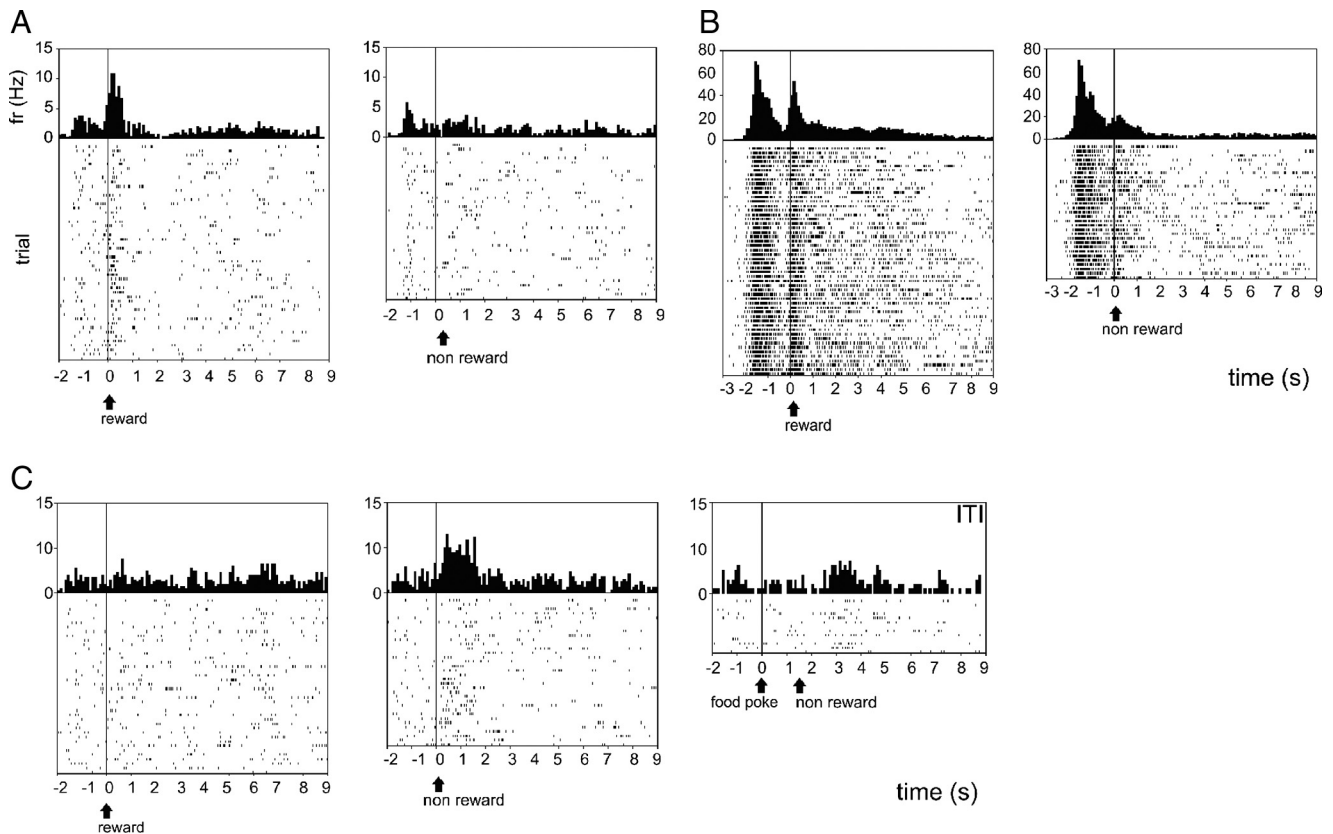
Figure 4. Differential firing after reward delivery during the rewarded and unrewarded condition. Activity in rewarded trials is synchronized on pellet delivery, in unrewarded trials at the same time point as pellet delivery in rewarded trials. *A*, Example of a unit demonstrating an increase in firing activity solely during the rewarded trials. *B*, Activity of a different unit showing a double correlate: both during the reward and waiting period (starting at −1.5 s), this neuron showed a significant increase in firing activity. During the reward phase, the increase in firing rate during rewarded trials was significantly larger compared with unrewarded trials. No difference was found between the two conditions in the waiting period. *C*, This unit increased firing specifically during unrewarded trials during the task. Activity did not significantly change during the ITI.

trial periods, a small subset of neurons was present that showed a very high degree of modulation by reward probability (parameter variability, range of 0.9–1.0) (Fig. 5*A–C*).

Population coding of expected reward probability
The results described above indicate that activity of a restricted subset of orbitofrontal neurons reflects expectancy of reward coming up with a specific probability, which is generally considered to be tightly linked to expected value (Kalenscher and Pennartz, 2008). We next asked whether not only single neurons but also the whole population of recorded neurons codes information regarding this reward parameter. Answering this question may shed light on how target regions of the OFC may read out information represented at the population level. Although the analyses above suggest a probability representation at the single-cell level, ensembles may not show a robust representation when, for instance, noise or other types of response variability obscure single-cell contributions. To this end, we decoded reward probability from the population activity for the three trial phases under examination, using template matching as reconstructing algorithm. For this analysis, the eight sessions with the largest amount of simultaneously recorded cells were used, with a minimum of 27 cells per session; the total number of neurons recorded across the eight sessions was 338 (cell count per session was 37, 51, 52, 27, 59, 39, 42, and 31, recorded from 2 rats). It should be noted that the study of probability coding in the reward delivery period primarily subserves the purpose of comparison with the motor and waiting periods. If an orbitofrontal neuron

fires more vigorously on rewarded versus non-rewarded trials (which is often the case as indicated by the single-unit data and is likely to depend on direct sensorimotor feedback correlated to food ingestion, taste, etc.), its accumulated spike counts will naturally come to correlate with reward probability because often-rewarded trial types will elicit more spikes than rarely rewarded types, whereas no specific coding of probability can be said to exist. On this account, we predicted that probability coding will be more accurate during the reward period than during the anticipatory movement and waiting periods.

The probability of reward (50, 75, or 100%) could be reconstructed from ensemble activity during all three trial periods, with a percentage correct significantly above the one-third chance level (one-way ANOVA, $p < 0.001$ in all three cases; the $p = 0\%$ condition was not part of our standard reconstruction analysis because of the lower amount of trials, but see below). Plotting the decoding score as a function of reconstruction ensemble size showed that, for all trial phases, performance improved with an increasing amount of cells, with the slope of the decoding curve being significantly positive (linear regression; in all cases, $p < 0.001$). The highest decoding scores obtained within these periods were 44% for the movement period (at $n = 27$ cells) and 48% ($n = 27$) and 44% ($n = 23$) for the waiting and reward period, respectively (Fig. 6). That the optimal decoding performance for the reward period was comparable with that of the two anticipatory periods is remarkable and stands in contrast to the prediction that population coding would be more accurate for
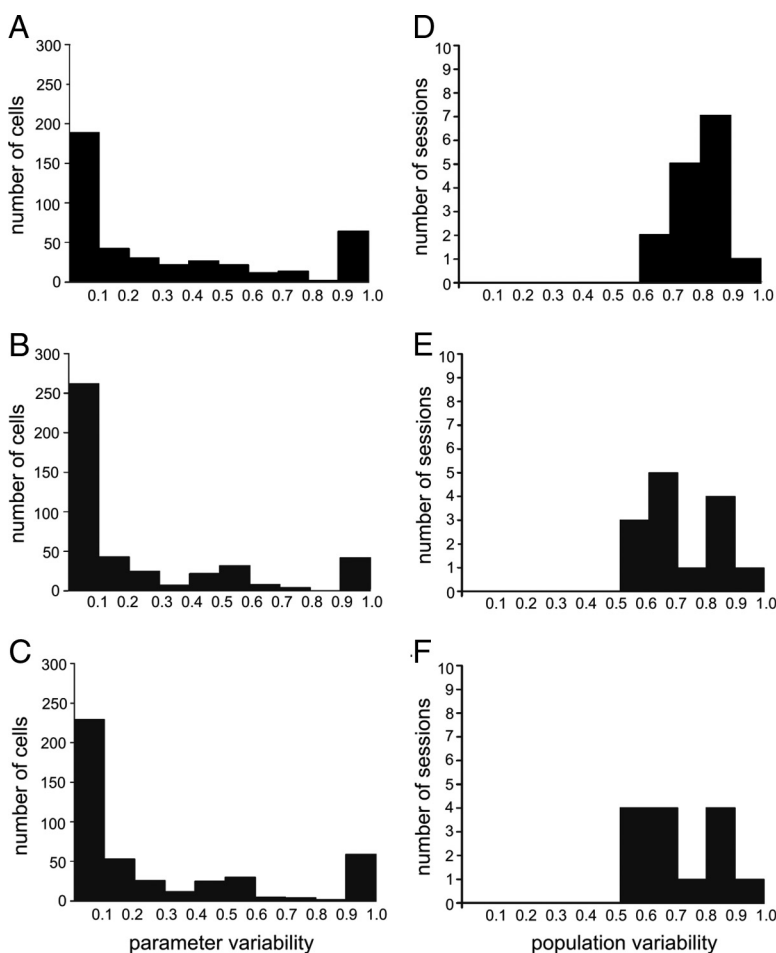
this period because of the sensorimotor feedback the rat receives during reward consumption. Whereas the curves for the movement and waiting periods both showed a gradually rising decoding success when ensemble size increased, the curve for the reward period rose more steeply at low cell counts, after which decoding success saturated around ensembles sizes of ~8 and higher. This difference suggests a higher redundancy of coding in the reward period compared with the anticipatory phases. This is supported by the finding that removal of cells displaying the largest variability in their response toward probability (i.e., parameter variability between 0.9 and 1.0) (Fig. 5) resulted in decoding scores that did not differ significantly from the decoding curves obtained using the entire population.

Because we introduced new odors to the rat in every novel recording session, it can be argued that significant population coding of probability should survive the removal of early-learning trials, because the rat will need this initial period to acquire knowledge about odor–probability associations. Although it is unknown a priori at which time point in the session neural representations of reward probability may begin to surface, we recalculated decoding scores when the first nine trials were left out of the reconstruction procedure, corresponding to the initial period in which rats on average kept on generating go responses to the odor cue signaling 0% reward probability. Using the last quarter of the trials for encoding and the remaining trials for decoding, significant decoding was found ($p < 0.001$ for all periods). Similar results were found when, after removal of the first nine trials, a quarter of the trials was randomly selected for encoding and the remaining trials were used for decoding ($p < 0.001$ for all periods).

Figure 5. Distribution of population and parameter variability for the movement period (**A**, **D**), the waiting period (**B**, **E**), and the reward period (**C**, **F**). On average, parameter-related variability covered a broader range of values than population variability, spanning the whole range from 0.0 to 1.0 (average of 0.29, 0.22, and 0.27 for the movement, waiting, and reward periods, respectively). Variation in population variability was less, with values ranging from 0.5 to 1.0 (average of 0.76, 0.71, and 0.71 for the movement, waiting, and reward periods, respectively).

Contribution of individual cells to coding of reward probability
Until here, the success of probability reconstruction from population activity was computed as a decoding score averaged across 100 groups of randomly selected neurons. This, however, does not provide insight in the contribution of individual neurons to an ensemble code for reward probability. To acquire more insight into the redundancy versus sparsity of coding, we calculated for all neurons the difference in the percentage of decoded information when a specific cell was added to a group of five neurons randomly selected from the same session. For each cell, this calculation was done 100 times, each time with a new randomly selected group of five additional neurons. Apart from the consideration that single cells may contribute reasonably to coding by such a relatively small group (Fig. 6), this size was chosen arbitrarily.

Because for all three trial phases this analysis yielded similar results, we only provide the data for the waiting period. During this period, 25% of the cells ($n = 85$) made a minimal contribution to the decoding success (between −0.5 and 0.5%), 40% ($n = 135$) made a

positive contribution ($>0.5\%$) (average ± SEM, +5.5 ± 0.4%), whereas 35% ($n = 118$) made a negative contribution ($<0.5\%$) to the reconstruction, with an average of −5.2 ± 0.5% (i.e., addition of these cells led to a decrease in correct decoding). The average positive contribution did not differ significantly from the negative one, as examined with an unpaired t test. This lack of significance agrees with the absence of a net positive slope in the reconstruction curve at an ensemble size of 5 (Fig. 6B). We also examined the percentage of cells showing an extremely strong contribution (more than 15% or less than −15%). This showed that only 2% of the cells ($n = 7$) made such extreme (positive or negative) contributions, with an average contribution of + 20 ± 1.4 and −20 ± 2.0%, respectively. Additional inspection of the distributions of single-cell contributions confirmed that there was no particular subset of cells contributing especially to the coding of reward probability and that the positive and negative contributions were nearly symmetrically distributed around zero, which is altogether consistent with a distributed representation across populations that contain cells making highly variable contributions.

Population coding of reward probability versus reward uncertainty
As already mentioned, reward probability and uncertainty are linked in the sense that uncertainty is absent at probabilities of 0 and 100% and maximal at a probability of 50%. To examine
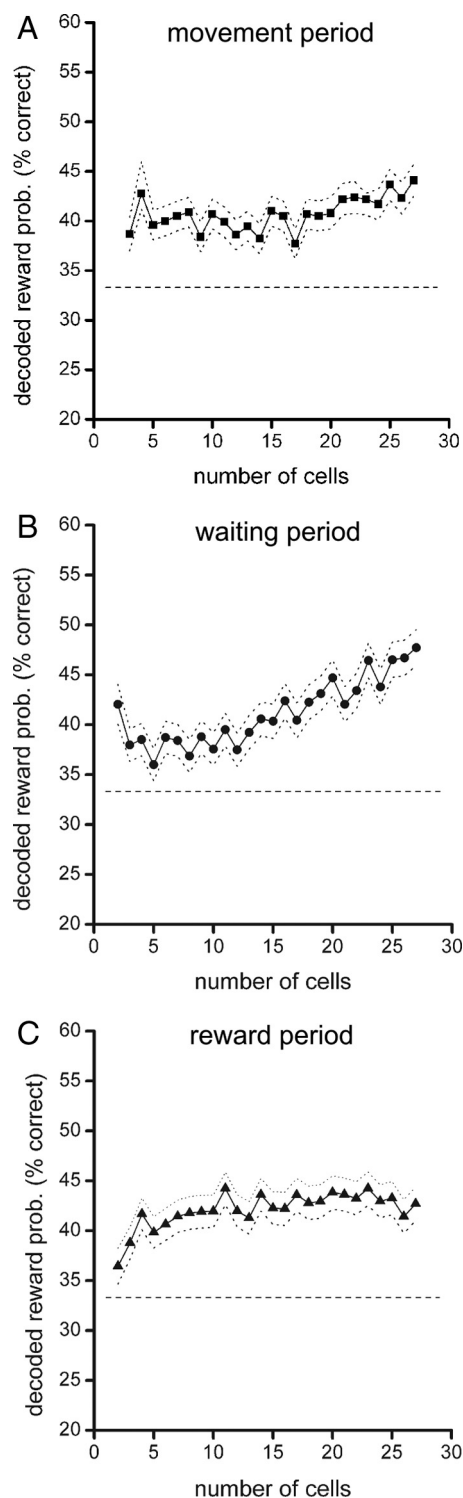
Figure 6. Decoding of reward probability with an increasing number of neurons for the movement (**A**), waiting (**B**), and reward (**C**) periods. Decoding time windows used were 0.7, 1.5, and 5 s, respectively. The horizontal axis indicates the size of the reconstruction ensemble, and the vertical axis indicates the percentage of trials in which reward probability was correctly decoded. The horizontal dashed line indicates chance level (33.3%), and dotted lines flanking the curves represent the 95% confidence interval (2 times the SE of proportion).

whether the observed population activity during the movement and waiting periods reflects reward probability or uncertainty, we examined probability reconstruction success by using unrewarded trials ($p = 0\%$) for encoding. A first hypothesis holds that, whenever reward probability is coded by OFC ensembles, encoding by unrewarded trials and decoding by $p = 50\%$ trials should yield decoding above chance, because the global difference in reward probability is smaller for these two trial types than $p = 0\%$ versus $p = 75\%$ and $p = 100\%$ trials. In case OFC would code uncertainty, however, one expects that encoding by unrewarded trials and decoding by $p = 100\%$ trials result in a decoding score above chance, because these two reward conditions are more alike in terms of uncertainty than $p = 0\%$ versus $p = 75\%$ and $p = 50\%$ trials. A third hypothesis holds that, in this procedure, decoding scores for $p = 50, 75$, and 100% should be random (33.3% success) because the $p = 0\%$ condition is set within a different trial type (as signaled by a distinct odor), without carrying over any quantitative information about reward probability to other trial types.

Given encoding by $p = 0\%$ trials, decoding for the movement and waiting periods was similar in that $p = 100\%$ trials yielded significant below-chance scores, whereas $p = 50\%$ trials were significantly above chance (Fig. 7) (one-way ANOVA, $p < 0.001$). For the $p = 75\%$ condition, decoding was at a random level in the movement period but gradually decreased below this level in the waiting period. When decoding scores for these two trial periods were averaged across all reward probabilities, performance was around chance level (data not shown). In agreement with the single-unit data, these results indicate that the observed variations in population activity during these trial phases are attributable to reward probability rather than uncertainty.

Discussion

The OFC has been strongly implicated in decision-making under uncertain conditions (Bechara et al., 1996, 1997). Here we examined to our knowledge for the first time whether and how single-cell and population activity within OFC is affected by the probability of future reward. The behavioral results showed animals to respond differentially depending on reward probability. During the periods in which predictive information coding could be studied (i.e., movement and waiting period), 22% of the cells demonstrating expectancy-related activity showed firing-rate changes differentiating between probabilities. This discriminatory activity was represented at the population level as well: predicted reward probability could be reconstructed from ensemble activity significantly above chance level for both phases. Although the overall decoding scores may seem rather low, the task required the animals to learn novel odor–probability associations each session. To estimate reward probability accurately, animals will need to accumulate experience across many trials. Moreover, there was no task requirement necessitating the animal to discriminate expected probabilities, because the chance of obtaining reward was not contingent on the speed of go responses.

In addition to the waiting period, expectancy-related activity was also found during the movement period. That this activity is likely reward related and not primarily determined by motor activity is supported by the absence of a significant neural response when the same behavioral sequence was executed during the intertrial interval (all 25 cells showed this difference) (Fig. 3C). Given the finding that our sample of cells showed transient firing responses during task periods but no type of activity change that arose during odor sampling and persisted throughout the trial until reward delivery, the question comes up how reward-predictive information may be neurally maintained during task performance. Speculatively, one possibility holds that cells in another brain area provide a working-memory-like buffer that is

active during task performance (cf. Mulder et al., 2003), whereas another scheme holds that ensembles of OFC neurons, active at an early trial stage, transfer their outcome-related information to other OFC ensembles active at later trial stages (cf. Baeg et al., 2003).

The probability of reward could be reconstructed from population activity in the reward period with a performance comparable with the other two trial phases. However, when instead of reward probability the availability of reward was reconstructed from population activity, the decoding score went up to 89% (at $n =$ 27 cells; data not shown). This indicates that, during the reward period, the presentation of a reward is coded more reliably than the overall reward probability. This need not be surprising given that neural activity during this period may be determined by processes other than "tracking" actual reward probability, for example, taste sensations or ingestion, which are closely related to processes of reward appraisal.

Variability and distribution of the representation of reward probability

That single units showed differential firing toward varying reward probability leaves unanswered the question whether probability of reward is represented in a sparse or redundant manner within OFC, i.e., by a few highly specifically tuned cells or in a broadly distributed way. Therefore, we first examined the firing-rate variability attributable to probability and found that this measure assumed relatively modest values compared with the overall variability in mean firing rate across the population. Second, except for the reward period, decoding scores depended only weakly on the size of the reconstruction ensemble (Fig. 6), and removal of cells displaying the largest variability in their response toward probability (i.e., parameter sparseness between 0.9 and 1.0) (Fig. 5) resulted in a decoding score for both periods that did not differ significantly from the decoding curves obtained using the entire population. Third, considering the widely dispersed single-cell contributions to the decoding score, with nearly symmetrical distribution of positive and negative values, these results indicate that reward probability is coded in a broadly distributed manner within OFC. However, because the decoding score did not rapidly saturate when cell count increased in pre-reward periods (Fig. 6, compare *A*, *B* with *C*), coding does not appear to be highly redundant in these phases, but instead cells make variable contributions to it. It is important to address whether and how a distributed representation of probability may be used by other brain structures targeted by the OFC to guide behavior and attention. How these structures integrate population signals into adaptive behavioral decision-making is essentially unknown, but it is of note that feedforward or recurrent
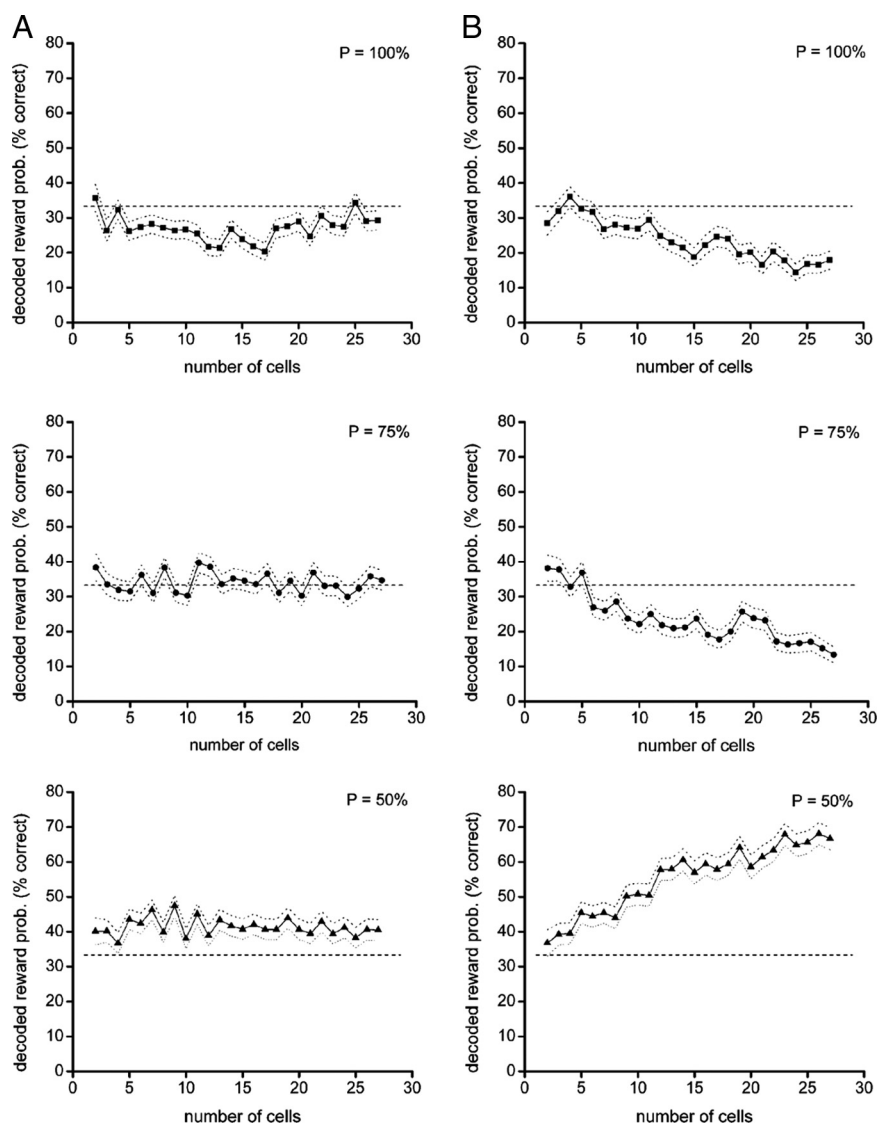
Figure 7. Decoding scores for the various reward probabilities with an increasing number of neurons during the movement period (*A*) and the waiting period (*B*) when spike vectors from unrewarded trials ($p = 0$%) were used for encoding. In both task phases, decoding for $p = 50$% trials was significantly above chance level (33.3%), whereas the score for $p = 100$% trials was below chance level. Decoding performance for $p = 75$% was either at chance level (movement period) or decreased below chance with increasing ensemble sizes (waiting period). When the decoding performance was averaged across all three probabilities, curves were at chance level for both trial phases (data not shown).

networks of units with broad tuning curves can extract sensory, motor, or motivational variables from a source population of noisy neurons (Zhang et al., 1998; Deneve et al., 1999). Such networks may be implemented in target structures of OFC such as higher associational cortical areas or striatopallidal circuits (Uylings et al., 2003; Voorn et al., 2004). Interesting in this respect is the result that nucleus accumbens lesions disrupt probabilistic discounting as well (Cardinal and Howes, 2005). Notably, network architectures capable of sustaining continuous attractors can read out population activity by a natural form of template matching (Wu and Amari, 2005). Prefrontal output also reaches mesencephalic dopamine cells (Phillipson, 1979; Uylings et al., 2003; Van De Werd and Uylings, 2008), potentially supporting the generation of phasic reward-prediction errors and of more tonic signals representing reward uncertainty (Schultz et al., 1997; Fiorillo et al., 2003).

Probability versus uncertainty coding

A hotly debated issue is whether OFC codes expected reward probability, uncertainty, or both. Functional magnetic resonance imaging (fMRI) findings in humans by Tobler et al. (2007) and Critchley et al. (2001) suggested that reward uncertainty rather than probability is coded within lateral orbital areas. Recently, Kepecs et al. (2008) recorded activity of single OFC neurons during an odor categorization task in which decision confidence was manipulated by presenting rats with odor mixtures and concluded that uncertainty, or decisional confidence, was coded in OFC. Both fMRI studies might be interpreted as contrasting with our finding that reward probability, but not uncertainty, is predominantly coded in OFC (Fig. 7). However, the blood oxygenation level-dependent (BOLD) signal as observed with fMRI is not considered to reflect the spike output of a particular brain area but rather the synaptic inputs from afferent structures and local intracortical processing (Logothetis et al., 2001). Hence, the limited spatiotemporal resolution of BOLD signals may explain why modulations of neural activity as observed in the current study may thus far have escaped detection with fMRI. The behavioral design of Kepecs et al. (2008) was very different from our design, because reward probability was not parametrically varied as a function of stimulus identity, and no stimulus predicting the null probability was included as highly certain but unrewarded outcome. Although Kepecs et al. (2008) found OFC cells responding in agreement with their model of choice confidence based on odor categorization, this confirmation was restricted to a subset of neurons. Although only very rarely found, we did identify 1 of 17 OFC neurons discriminating on the basis of uncertainty, so the results from the two studies may not contradict each other principally; uncertainty or probability coding may predominate depending on task requirements.

Additional implications

It is still unclear whether parameters related to expected reinforcement or utility are coded by a form of integrative neural activity subserving the role of a "common currency" within the OFC or elsewhere in the brain, i.e., whether neurons code a lumped measure of expected utility in which all relevant parameters (such as delay, magnitude, uncertainty) have been included (Tremblay and Schultz, 1999; Montague and Berns, 2002; Padoa-Schioppa and Assad, 2006; Kalenscher and Pennartz, 2008). As suggested by Roesch et al. (2006), coding of time-discounted rewards in rat OFC seems independent of coding of absolute reward value. In contrast, previous findings by Roesch and Olson (2004) in primate OFC indicated that neurons do code reward value in a common currency: single-unit activity elicited by visual cues associated with differently delayed or sized rewards was shown to covary with both parameters. As demonstrated here, the probability of future reward is population coded in a similar manner within OFC as is the case for reward magnitude (van Duuren et al., 2008). Both parameters are represented in a distributed manner by neurons that display a large diversity in parameter sensitivity. If independent parameter coding would turn out predominant for single neurons, it is still possible that larger OFC ensembles act as functional entities coding a common currency. The finding that both reward probability and magnitude exert modest modulatory effects on single cells and that parameter information appears to be represented in a spatiotemporally distributed form suggests that the ensemble level is at least as relevant for studying the common currency problem at the single-unit level.

References

Baeg EH, Kim YB, Huh K, Mook-Jung I, Kim HT, Jung MW (2003) Dynamics of population code for working memory in the prefrontal cortex. Neuron 40:177–188.

Bechara A, Tranel D, Damasio H, Damasio AR (1996) Failure to respond autonomically to anticipated future outcomes following damage to the prefrontal cortex. Cereb Cortex 6:215–225.

Bechara A, Damasio H, Tranel D, Damasio AR (1997) Deciding advantageously before knowing the advantageous strategy. Science 275:1293–1295.

Cardinal RN, Howes NJ (2005) Effects of lesions of the nucleus accumbens core on choice between small certain rewards and large uncertain rewards in rats. BMC Neurosci 6:37.

Critchley HD, Mathias CJ, Dolan RJ (2001) Neural activity in the human brain relating to uncertainty and arousal during anticipation. Neuron 29:537–545.

Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat Neurosci 8:1704–17011.

Deneve S, Latham PE, Pouget A (1999) Reading population codes: a neural implementation of ideal observers. Nat Neurosci 2:740–745.

Dreher JC, Kohn P, Berman KF (2006) Neural coding of distinct statistical properties of reward information in humans. Cereb Cortex 16:561–573.

Ernst M, Nelson EE, McClure EB, Monk CS, Munson S, Eshel N, Zarahn E, Leibenluft E, Zametkin A, Towbin K, Blair J, Charney D, Pine DS (2004) Choice selection and reward anticipation: an fMRI study. Neuropsychologia 42:1585–1597.

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

Gothard KM, Skaggs WE, Moore KM, McNaughton BL (1996) Binding of hippocampal CA1 neural activity to multiple reference frames in a landmark-based navigation task. J Neurosci 16:823–835.

Gray CM, Maldonado PE, Wilson M, McNaughton B (1995) Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. J Neurosci Methods 63:43–54.

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

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

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

Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 27:712–719.

Lehky SR, Sejnowski TJ (1990) Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity. J Neurosci 10:2281–2299.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. Nature 412:150–157.

Mobini S, Body S, Ho MY, Bradshaw CM, Szabadi E, Deakin JF, Anderson IM (2002) Effects of lesions of the orbitofrontal cortex on sensitivity to delayed and probabilistic reinforcement. Psychopharmacology 160:290–298.

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

Mulder AB, Nordquist RE, Orgüt O, Pennartz CM (2003) Learning-related changes in response patterns of prefrontal neurons during instrumental conditioning. Behav Brain Res 146:77–88.

Narayanan NS, Kimchi EY, Laubach M (2005) Redundancy and synergy of neuronal ensembles in rat motor cortex. J Neurosci 25:4207–4216.

O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001) Abstract reward and punishment representations in the human orbitofrontal cortex. Nat Neurosci 4:95–102.

O'Doherty J, Critchley H, Deichmann R, Dolan RJ (2003) Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. J Neurosci 23:7931–7939.

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

Pais-Vieira M, Lima D, Galhardo V (2007) Orbitofrontal cortex lesions disrupt risk assessment in a novel serial decision-making task for rats. Neuroscience 145:225–231.

Paxinos G, Watson C (2005) The rat brain in stereotaxic coordinates, Ed 5. Amsterdam: Elsevier.

Perez-Orive J, Mazor O, Turner GC, Cassenaer S, Wilson RI, Laurent G (2002) Oscillations and sparsening of odor representations in the mushroom body. Science 297:359–365.

Phillipson OT (1979) Afferent projections to the ventral tegmental area of Tsai and interfascicular nucleus: a horseradish peroxidase study in the rat. J Comp Neurol 187:117–143.

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

Ramus SJ, Eichenbaum H (2000) Neural correlates of olfactory recognition memory in the rat orbitofrontal cortex. J Neurosci 20:8199–8208.

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

Roesch MR, Taylor AR, Schoenbaum G (2006) Encoding of time-discounted reward sin orbitofrontal cortex is independent of value representation. Neuron 51:509–520.

Rogers RD, Owen AM, Middleton HC, Williams EJ, Pickard JD, Sahakian BJ, Robbins TW (1999) Choosing between small, likely rewards and large, unlikely rewards activates inferior and orbital prefrontal cortex. J Neurosci 19:9029–9038.

Rolls ET, Tovee MJ (1995) Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. J Neurophysiol 73:713–726.

Schoenbaum G, Chiba AA, Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. Nat Naurosci 1:155–159.

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

Schultz W, Preuschoff K, Camerer C, Hsu M, Fiorillo CD, Tobler PN, Bossaerts P (2008) Explicit neural signals reflecting reward uncertainty. Philos Trans R Soc Lond B Biol Sci 363:3801–3811.

Sokal RR, Rohlf FJ (1995) Biometry, Ed 3. New York: Freeman.

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

Tobler PN, O'Doherty JP, Dolan RJ, Schultz W (2007) Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. J Neurophysiol 97:1621–1632.

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

Uylings HB, Groenewegen HJ, Kolb B (2003) Do rats have a prefrontal cortex? Behav Brain Res 146:3–17.

Van De Werd HJ, Uylings HBM (2008) The rat orbital and agranular insular prefrontal cortical area: a cytoarchitectonic and chemoarchitectonic study. Brain Struct Funct 212:387–401.

van Duuren E, Escámez FA, Joosten RN, Visser R, Mulder AB, Pennartz CM (2007) Neural coding of reward magnitude in the orbitofrontal cortex of the rat during a five-odor discrimination task. Learn Mem 14:446–456.

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

Voorn P, Vanderschuren LJ, Groenewegen HJ, Robbins TW, Pennartz CM (2004) Putting a spin on the dorsal-ventral divide of the striatum. Trends Neurosci 27:468–474.

Wu S, Amari S (2005) Computing with continuous attractors: stability and online aspects. Neural Comput 17:2215–2239.

Zhang K, Ginzburg I, McNaughton BL, Sejnowski TJ (1998) Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. J Neurophysiol 79:1017–1044.