

# Contextual Novelty Modulates the Neural Dynamics of Reward Anticipation

Nico Bunzeck,<sup>1\*</sup> Marc Guitart-Masip,<sup>2\*</sup> Ray J. Dolan,<sup>3</sup> and Emrah Düzel<sup>2,4</sup>

<sup>1</sup>Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany, <sup>2</sup>Institute of Cognitive Neuroscience, University College London (UCL), London WC1N 3AR, United Kingdom, <sup>3</sup>Wellcome Trust Centre for Neuroimaging at UCL, London WC1N 3BG, United Kingdom, and <sup>4</sup>Institute of Cognitive Neurology and Dementia Research, Otto von Guericke University, 39120 Magdeburg, Germany

We investigated how rapidly the reward-predicting properties of visual cues are signaled in the human brain and the extent these reward prediction signals are contextually modifiable. In a magnetoencephalography study, we presented participants with fractal visual cues that predicted monetary rewards with different probabilities. These cues were presented in the temporal context of a preceding novel or familiar image of a natural scene. Starting at ~100 ms after cue onset, reward probability was signaled in the event-related fields (ERFs) over temporo-occipital sensors and in the power of theta (5–8 Hz) and beta (20–30 Hz) band oscillations over frontal sensors. While theta decreased with reward probability beta power showed the opposite effect. Thus, in humans anticipatory reward responses are generated rapidly, within 100 ms after the onset of reward-predicting cues, which is similar to the timing established in non-human primates. Contextual novelty enhanced the reward anticipation responses in both ERFs and in beta oscillations starting at ~100 ms after cue onset. This very early context effect is compatible with a physiological model that invokes the mediation of a hippocampal-VTA loop according to which novelty modulates neural response properties within the reward circuitry. We conclude that the neural processing of cues that predict future rewards is temporally highly efficient and contextually modifiable.

## Introduction

Recordings in non-human primates show that midbrain dopamine neurons (Tobler et al., 2005), neurons in the prefrontal cortex (Watanabe, 1996), basal ganglia (Kawagoe et al., 1998), and parietal cortex (Platt and Glimcher, 1999) respond to cues that predict rewards with a rapid onset latency of ~100 ms and endure for ~200 ms. The early onset of these signals is thought to reflect the evolutionary need of a learning system that rapidly signals learned reward-associations to efficiently guide behavior and decision making (Schultz, 2007).

In humans, electrophysiological studies have focused on the neural timing of processing reward outcomes. For instance, event-related potentials (ERPs) revealed that reward size, valence and probability are signaled at ~200–300 ms after reward feedback (Yeung and Sanfey, 2004; Cohen et al., 2007; San Martín et al., 2010). Furthermore, time-frequency (TF) analysis of EEG-recordings revealed that theta (4–8 Hz) power is sensitive to monetary losses and outcome probability while beta power (~20–30 Hz) is sensitive to monetary gains (Cohen et al., 2007;

Marco-Pallares et al., 2008). Although these, and other, human studies (Aberg and Nilsson, 2001; Holroyd and Coles, 2002; Nieuwenhuis et al., 2004; Christie and Tata, 2009; Wu and Zhou, 2009; Zaghoul et al., 2009; Cavanagh et al., 2010; Donamayor et al., 2011) demonstrate a close relationship between different features of reward outcome and specific neural processes at ~200–300 ms, the neural timing of reward anticipation in healthy humans remains unexplored.

We have recently shown that reward-predicting cues elicit stronger fMRI signals in the striatum in a context of novelty, that is when preceded by photographs of novel (compared with familiar) environments (Guitart-Masip et al., 2010). It is yet unclear whether such contextual effects are embedded into the computation of reward prediction responses per se and hence impact on neural responses to these cues at a very early point in time. According to this possibility, the contextually modulated version of reward anticipation responses would replace the neural responses that represent the originally learned cue-outcome contingencies. Such a possibility is physiologically plausible because hippocampally generated novelty signals can increase the number of neurons in the substantia nigra/ventral tegmental area (SN/VTA) that effectively respond to reward-predicting stimuli (Goto and Grace, 2008) via a polysynaptic pathway termed the hippocampal-VTA loop (Lisman and Grace, 2005). Alternatively, contextual effects may be incorporated into reward prediction responses at a later point in time, hence modifying reward anticipation after the initially learned cue-outcome contingencies have been signaled.

Here, we used magnetoencephalography (MEG) in combination with a reward anticipation paradigm adapted from a previ-

Received Jan. 26, 2011; revised July 5, 2011; accepted July 15, 2011.

Author contributions: N.B., M.G.-M., and E.D. designed research; N.B. and M.G.-M. performed research; N.B. analyzed data; N.B., M.G.-M., R.J.D., and E.D. wrote the paper.

This work was supported by a Wellcome Trust Project Grant (Grant 81259 to E.D. and R.J.D.; www.wellcome.ac.uk). R.J.D. is supported by a Wellcome Trust Programme Grant. M.G.-M. holds a Marie Curie Fellowship (www.mariecurie.org.uk). N.B. is supported by Hamburg state cluster of excellence (neuroapt!).

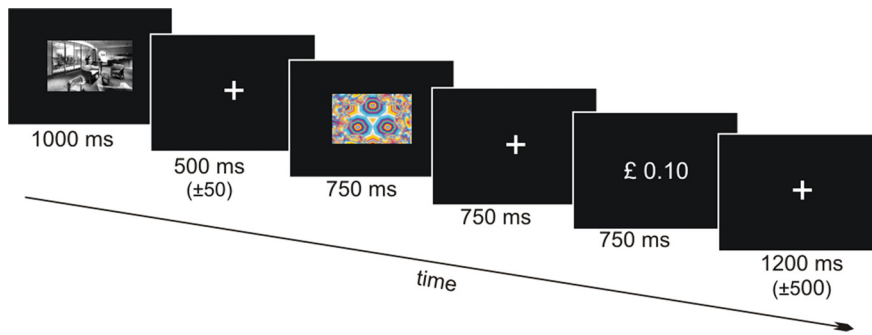
\*N.B. and M.G.-M. contributed equally to this work.

The authors declare no competing financial interests.

Correspondence should be addressed to Nico Bunzeck, Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany. E-mail: n.bunzeck@uke.de.

DOI:10.1523/JNEUROSCI.0461-11.2011

Copyright © 2011 the authors 0270-6474/11/3112816-07\$15.00/0



**Figure 1.** Experimental design (phase 3). Note that subjects underwent a familiarization and a conditioning phase inside the MEG during which no data were acquired.

ous fMRI study (Guitart-Masip et al., 2010). We hypothesized that reward anticipation would be reflected in the event-related field (ERF) and theta and beta band oscillations at  $\sim 100$ – $300$  ms as a function of reward probability (Cohen et al., 2007; Marco-Pallares et al., 2008). Furthermore, we hypothesized that contextual novelty changes these early reward anticipation signals (Kakade and Dayan, 2002; Lisman and Grace, 2005; Guitart-Masip et al., 2010).

## Materials and Methods

Sixteen adult subjects participated (nine female and seven male; age range 19–28 years; mean 21.6 years, SD = 2.5 years). All were healthy, right-handed, and had normal or corrected-to-normal visual acuity. None of the participants reported a history of neurological, psychiatric or medical disorders or any current medical problems. Subjects gave written informed consent according to the approval of the local ethics committee (University College London, UK).

The task was divided into three phases. In phase 1, subjects were familiarized with a set of 10 images (five indoor, five outdoor). Each image was presented 10 times for 1000 ms with an interstimulus interval of  $1750 \pm 500$  ms. Subjects indicated the indoor/outdoor status using their right index and middle finger. In phase 2, three fractal images were paired, under different probabilities (0, 0.4, and 0.8) with a monetary reward of 10 pence in a conditioning session. Each fractal image was presented 40 times. On each trial, one of three fractal images was presented on the screen for 750 ms and subjects indicated the detection of the stimulus with a button press. The probabilistic outcome (10 or 0 pence) was presented as a number on the screen 750 ms later for another 750 ms and subjects indicated whether they won any money or not using their index and middle finger. The intertrial interval was  $1750 \pm 500$  ms. Finally, in a test phase (phase 3), the effect of contextual novelty on reward-related responses was determined in four 11 min sessions (Fig. 1). Here, a scene image was presented for 1000 ms and subjects indicated the indoor/outdoor status using their right index and middle finger. The image was either from the familiarized set of pictures from phase 1 (referred to as “familiar images”) or from another set of pictures that had never been presented (referred to as “novel images”).

In total, 240 novel and 240 old images were presented to each subject. The image was followed by a fixation period of 500 ms ( $\pm 50$  ms) during which the subjects could indicate the indoor/outdoor status of the scene image if they have not done so during the picture presentation period. Subsequently, one of the three fractal images from phase 2 (referred to as reward predictive cue) was presented for 750 ms (here, subjects were instructed not to respond). As in the second phase, the probabilistic outcome (10 or 0 pence) was presented 750 ms later for another 750 ms and subjects indicated whether they won money or not using their index and middle finger. Responses could be made while the outcome was displayed on the screen and during the subsequent fixation period which lasted  $1250 \pm 500$  ms. During each session, each fractal image was presented 20 times following a novel picture and 20 times following a familiar picture, resulting in 120 trials per session. The presentation order of the six trial types

was fully randomized. All three experimental phases were performed inside the MEG scanner but data were acquired only during the test phase (phase 3). Subjects were instructed to respond as quickly and as correctly as possible and that they would be paid their earnings up to £20. Participants were told that 10 pence would be subtracted for each incorrect response—these trials were excluded from the analysis. Total earnings were displayed on the screen only at the end of the fourth block.

All images were gray-scaled and normalized to a mean gray-value of 127 and a SD of 75. None of the scenes depicted human beings or human body parts (including faces) in the foreground.

MEG recordings were made in a magnetically shielded room by using a 275-channel CTF system with SQUID-based axial gradiometers (VSM MedTech Ltd.) and second-order gradients. Neuromagnetic signals were digitized continuously at a sampling rate of 600 Hz and behavioral responses were made via a MEG-compatible response pad. Data were low-pass filtered at 120 Hz during acquisition. Data were analyzed with SPM8 (Wellcome Trust Centre for Neuroimaging, University College London, UK) and MATLAB (The MathWorks). One distinct and major strength of SPM for M/EEG is that voxel-based images are used to perform statistical analyses. To account for the problem of multiple comparisons that arises out of this approach the second-level analyses are based on general linear models and random field theory (Worsley et al., 1996, 2004; Kiebel and Friston, 2004a,b). There have been a number of experimental publications using this approach (Furl et al., 2007; Weil et al., 2007; Henson et al., 2008; Bunzeck et al., 2009).

For the ERF analysis data were extracted from 100 ms before to 1000 ms after stimulus onset (i.e., epoching) and baseline-corrected relative to the 100 ms before stimulus onset. Epoched data were down-sampled at 150 Hz and low-pass filtered at 20 Hz. Before averaging trials for each condition (pictures: one condition for novel and one for familiar images [2 conditions]; cues: one condition for each probability following either novel or familiar images [6 conditions]; outcome: one condition for each probability and novelty condition [6 conditions]) simple thresholding was applied to remove artifact-containing trials with signals exceeding 3500 fT. Only trials with correct behavioral responses to both picture and outcome were used for averaging. Epoched data were converted into nifti format. This produced a 3D image of channel space  $\times$  time (Kilner and Friston, 2010). The 2D channel space was created by projecting the sensor locations onto a plane followed by a linear interpolation to a  $64 \times 64$  pixel grid (pixel size =  $2.12 \times 2.69$  mm). The time dimension consisted of 166 6.67 ms samples per epoch. Finally, these images were smoothed using a Gaussian kernel [full-width half-maximum (FWHM)] of  $\text{FWHM} = 5 \times 5 \times 30$ . Smoothing is necessary to accommodate the spatial and/or temporal variance between subjects and it leads to a better conformity regarding random field theory (Litvak et al., 2011).

TF data were epoched from 450 ms before to 1000 ms after stimulus onset; baseline corrected from 450 ms before stimulus onset; down-sampled at 250 Hz; low-pass filtered at 50 Hz and thresholded at 3500 fT. Oscillatory activity in the MEG signal was quantified by continuous Morlet wavelet transformation (factor 7). The wavelet decomposition was applied to each trial, sensor and subject across the frequency range 4–45 Hz. This was followed by averaging across all trials of the same condition and a rescaling of the TF spectrogram by subtracting the power in the baseline ( $p_b$ ) from the power of the trial ( $p$ ). It should be noted that the reported results in the beta frequency range (where we observed our main results; see results section) could be replicated using two other methods for baseline correction [ $\text{LogR} = (\log_{10}(p/p_b))$ , and  $\text{Rel} = 100 \times (p - p_b)/p_b$ ]. Subsequently, the TF data were converted into nifti format for each of the three frequency ranges of interest (theta: 5–8 Hz; alpha: 8–12 Hz; beta: 20–30 Hz) separately. This produced a 3D image of channel space  $\times$  time (averaged across 5–8 Hz, 8–12 Hz and 20–30 Hz, separately) (Kilner and Friston, 2010). Similar to the ERP

analysis, the 2D channel space was created by projecting the sensor locations onto a plane followed by a linear interpolation to a  $64 \times 64$  pixel grid (pixel size =  $2.12 \times 2.69$  mm). The time dimension consisted of 363 4 ms samples per epoch. Finally, these images were smoothed using a Gaussian kernel of  $\text{FWHM} = 5 \times 5 \times 30$ .

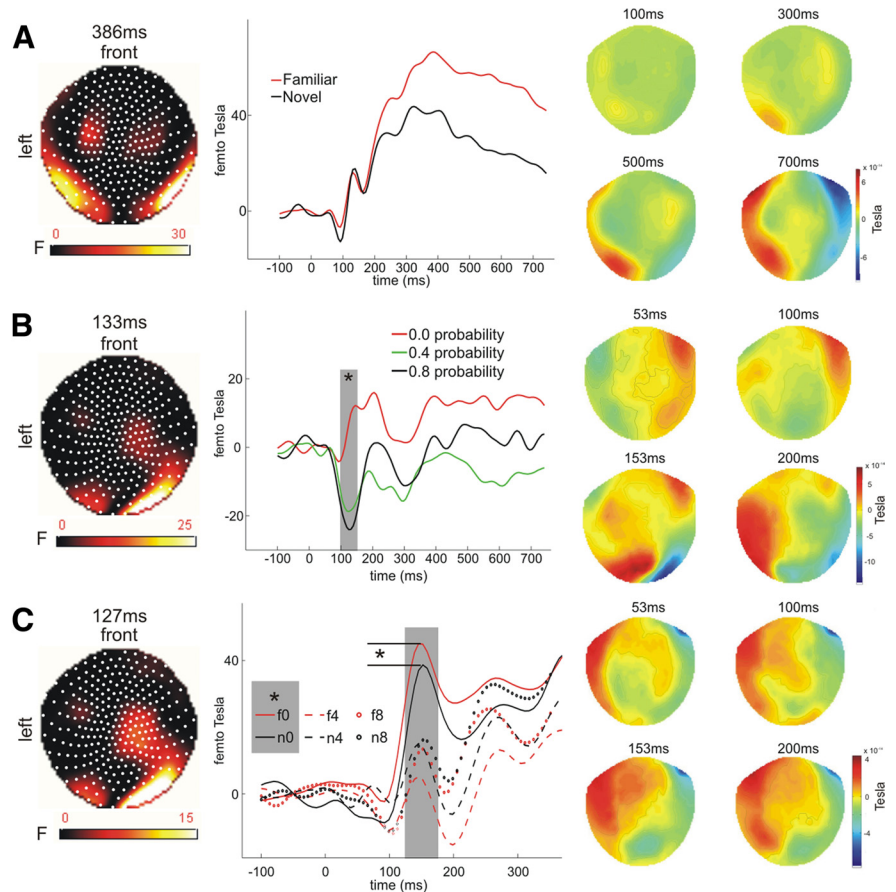
ERFs and TF data (for each frequency range) for pictures and cues were analyzed separately using random-effects  $2 \times 1$  (scene pictures) or  $3 \times 2$  (cues) ANOVAs. The ANOVAs for pictures comprised the factor novelty (new, familiar) and the ANOVAs for cues comprised the factors novelty (new, familiar) and probability (0.0, 0.4, 0.8). While the first set of ANOVAs were designed to test ERP and oscillatory effects of novelty associated with the pictures the second ANOVA allowed us to test cue effects of reward probability as well as the influence of novelty on cue probability. Statistical analysis for the TF data was limited from  $-250$  ms before until 800 ms after stimulus onset to avoid edge effects induced by Morlet wavelet transformation. All reported statistical parametric maps were thresholded at  $p < 0.001$  (uncorrected) unless stated otherwise.

Further detailed information on the methods of SPM8 for EEG and MEG data analysis can be found in the work by Litvak et al. (2011).

## Results

Behaviorally, subjects responded quickly and with high accuracy. In phase 3 (see Materials and Methods), indoor/outdoor discrimination was faster for familiar compared with novel images [mean reaction time (RT); familiar images: RT = 651 ms, SD = 114 ms; novel images: RT = 703 ms, SD = 119 ms;  $p < 0.001$ ]. There was no RT difference for the win/no-win discrimination at outcome (mean RT win: RT = 562 ms, SD = 92 ms; no-win = 562 ms, SD = 95 ms;  $p > 0.9$ ) and similarly, no difference in RT between outcomes following familiar or novel images (RT outcome after familiar images: RT = 561 ms, SD = 92 ms; RT outcome after novel images: RT = 561 ms, SD = 96 ms;  $p > 0.44$ ). On average, subjects responded correctly in 71% of all trials to both the indoor/outdoor discrimination as well as the win/no-win decision (hit-rate = 0.71; SD = 0.1). In phase 3, subjects were not required to make any button presses in response to reward-predicting cues. Thus, behaviorally no novelty-bonus effects could be reported. During an initial conditioning phase (phase 1, see Materials and Methods) there were no RT differences for the three different fractal images possibly due to ceiling effects (0.8-probability: RT = 401 ms, SD = 84 ms; 0.4-probability: RT = 405 ms, SD = 86 ms; 0-probability: RT = 398 ms, SD = 81 ms;  $F > 0.4$ ,  $p > 0.6$ ).

MEG results for both the ERF and TF analyses are reported from 0 to 800 ms after stimulus onset at an uncorrected level of  $p < 0.001$  (unless stated otherwise) followed by familywise error (FWE) correction for multiple comparisons. Motivated by previous studies (see Introduction) the time window for multiple comparisons for the cue-effects ranged from 100 to 300 ms. Since novelty effects have been reported for a much wider time window

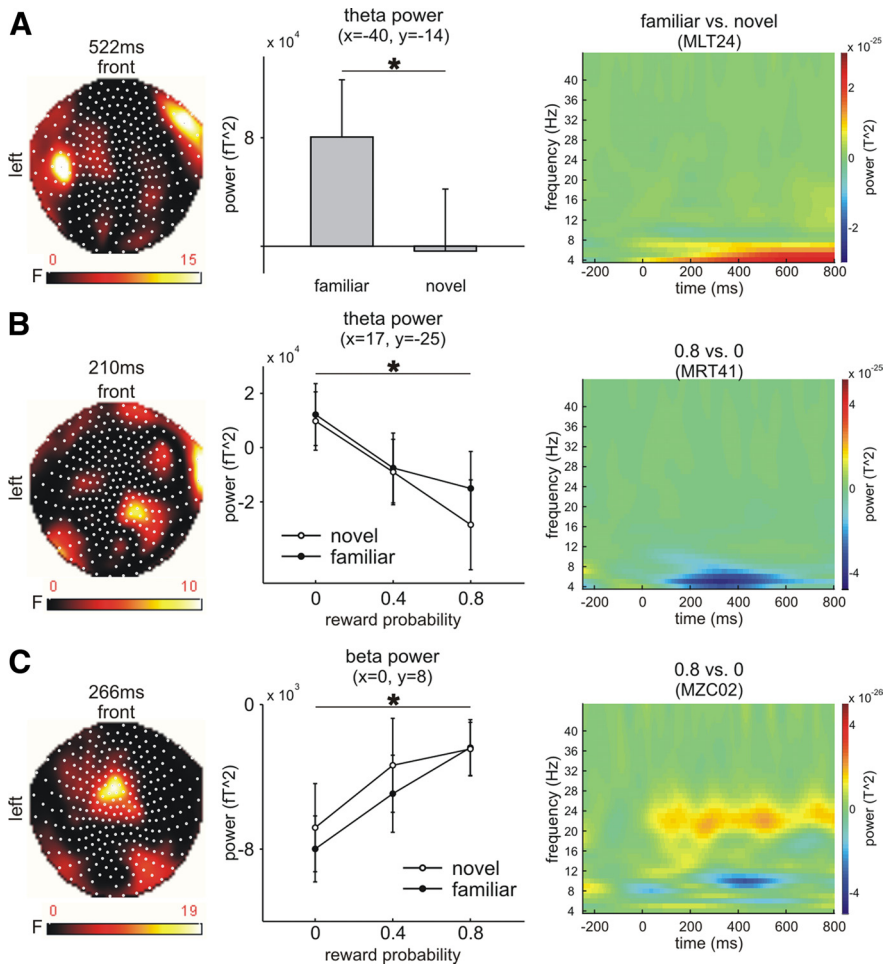


**Figure 2.** Event-related magnetic fields. **A**, Novelty effects emerged at  $\sim 200$  ms after stimulus onset over bilateral temporal sensors and peaked at  $\sim 400$  ms ( $x = 47$ ,  $y = -68$ ; nearest channel: MRT57). **B**, Reward probability was represented over right occipitotemporal sensors between 100 and 150 ms and peaked at 133 ms ( $x = 17$ ,  $y = -95$ ; nearest channel: MR051). **C**, A novelty bonus effect was observed for the low reward probability condition peaking at 127 ms over occipitotemporal sensors ( $x = 42$ ,  $y = -76$ ; nearest channel: MR053). Left columns show statistical parametric maps of the  $F$ -statistic. Middle columns show the time course of the effects as extracted from voxel space (on which the statistical analyses are based) (statistically significant differences,  $*p < 0.05$ ), and right columns show the topographical maps of the MEG data before 3D conversion (see Materials and Methods).

(Rugg and Curran, 2007; Bunzeck et al., 2009) the range for multiple comparisons has been chosen from 85 to 800 ms. It should be noted that FWE correction did not involve any a priori restrictions in the  $x$ ,  $y$  dimension (space) but the  $z$  dimension (time).

In a first ERF analysis we investigated the neural effects associated with initial scene/novelty processing (simple  $1 \times 2$  ANOVA with the levels old/new) followed by a  $3 \times 2$  ANOVA on subsequent cue processing. The  $F$ -contrast on picture novelty revealed main effects of novelty over several sensors. The strongest novelty effects were observed over bilateral temporal regions peaking at 387 ms with an onset latencies at  $\sim 200$  ms after picture onset and a duration until the end of the analysis window (1000 ms) (Fig. 2A). This effect survived FWE correction for multiple comparisons (using the time window of 85–800 ms and no spatial restrictions). Novelty effects were also observed over left frontal (peak at 767 ms) and right central sensors (peak at 333 ms). These results are in line with previous electrophysiological studies on recognition memory (Tsivilis et al., 2001; Ranganath and Rainer, 2003; Düzel et al., 2004; Gonsalves et al., 2005; Rugg and Curran, 2007; Bunzeck et al., 2009).

The  $3 \times 2$  ANOVA ( $F$ -contrast) on cue activation (note the temporal separation between the picture and the cue) revealed an early main effect of reward probability over right occipitotempo-



**Figure 3.** Results of the time-frequency analyses. **A**, Theta power increased over right frontal and left temporal sensors ( $x = -40, y = -14$ ; nearest channel: MLT24) for familiar stimuli with a peak at 522 ms. **B**, Theta power over right temporal and centroparietal sensors ( $x = 17, y = -25$ ; nearest channel: MRT41) decreased as a function of reward probability. **C**, Beta power increased with reward probability over frontocentral sensors ( $x = 0, y = 8$ ; nearest channel: MZC02). Left columns show statistical parametric maps of the  $F$ -statistics, middle columns show the effects for each condition at the peak time point as extracted from voxel space, and right columns show the time-frequency plots as extracted from the nearest channel (before conversion to voxel space; see Materials and Methods). Error bars denote 1 SEM (statistically significant differences,  $*p < 0.05$ ).

ral sensors (duration  $\sim 100$ – $200$  ms; peak at 133 ms after cue onset; Fig. 2B) and a later effect over central sensors (peak at 413 ms). The early effect survived FWE correction for multiple comparisons ( $p < 0.05$ ) using our a priori defined time window of interest: 100–300 ms. As mentioned above, this analysis did not involve any a priori restriction in the  $x, y$  dimension (space) but the  $z$  dimension (time).

In a planned *post hoc* comparison, we tested for sensors and time points that exhibited an effect of reward-probability (0 vs 0.8 probability) as well as a difference between cues that followed novel vs familiar images (for each probability separately; using implicit masking). For the 0-probability condition this contrast revealed a statistically significant effect over right temporal sensors peaking at 127 ms (65 voxels; Fig. 2C) and right occipital sensors (peak at 207 ms, 31 voxels; FWE  $< 0.05$  using the time window 100–300 ms). There were no novelty-bonus effects for either the 0.4 or 0.8 probability condition at our initial statistical threshold (0.001). However, at a lower threshold of  $p < 0.005$  we observed effects for both conditions over occipital (medium; peak at 207 ms) and central sensors (high; peak at 187 ms). Together, the most robust “novelty bonus” effect was observed for a 0-probability

condition at 127 ms after stimulus onset over temporal sensors (Fig. 2C). For both the medium and high probability condition, novelty bonus effects were also observed at  $\sim 200$  ms after stimulus onset but only at a lower statistical threshold.

We next assessed oscillatory activity in a series of ANOVAs in relation to presentation of pictures ( $1 \times 2$  ANOVAs with levels novel, familiar) and cues ( $3 \times 2$  ANOVAs with factors probability [0, 0.4, 0.8] and novelty [novel, familiar]). We focused our analyses on the theta (5–8 Hz), beta (20–30 Hz) and the alpha (8–12 Hz) bands. For pictures, we observed significantly enhanced theta oscillations for familiar compared with novel images over right frontolateral (peak at 798 and 522 ms; Fig. 3A) and left temporal sensors (peak 198 ms). Although this effect did not survive FWE-correction using the time window of 85–800 ms, it is compatible with a wide range of previous studies that link theta oscillation with regulating the neural dynamics of recognition memory (Buzsáki, 2002; Axmacher et al., 2006; Hasselmo, 2006; Klimesch et al., 2008; Düzel et al., 2010). There were no significant novelty effects in the alpha or beta band.

For cues, we observed a significant main effect of cue probability in the theta band over right centroparietal sensors (peak 634 ms, 854 voxels) and right temporal sensors (peak 210 ms, 388 voxels). At both sides theta decreased as a function of reward probability (see Fig. 3B for the right temporal effect) and the  $t$ -contrast of the early temporal effect survived FWE correction ( $p < 0.05$ ; time window: 100–300 ms). In the beta band we observed a main effect of reward-probability over frontocentral sensors (peak 266 ms, 982 voxels) with increases in beta power as a function of reward-probability [Fig. 3C; FWE correction (using the time window 100–300 ms):  $p < 0.05$ ]. There was no statistically significant main effect of reward probability in the alpha band.

Finally, in the beta frequency range we observed a reliable novelty-bonus effect in the medium (0.4) and high probability (0.8) condition (Fig. 4B). Both probabilities exhibited the novelty bonus effect over frontocentral sensors [medium: peak at 142 ms, 118 voxels; high: peak 142 ms, 275 voxels (FWE correction approached significance:  $p = 0.16$ ; using the time window 100–300 ms)]. For the low reward probability condition we also observed a novelty bonus over central sensors (peak at 254 ms, 148 voxels; Fig. 4A) and right frontal (peak 426 ms, 204 voxels) sensors but only at a lower statistical threshold of  $p = 0.005$ . There was no contextual effect of novelty on reward-predicting cues for theta and alpha frequencies.

### Discussion

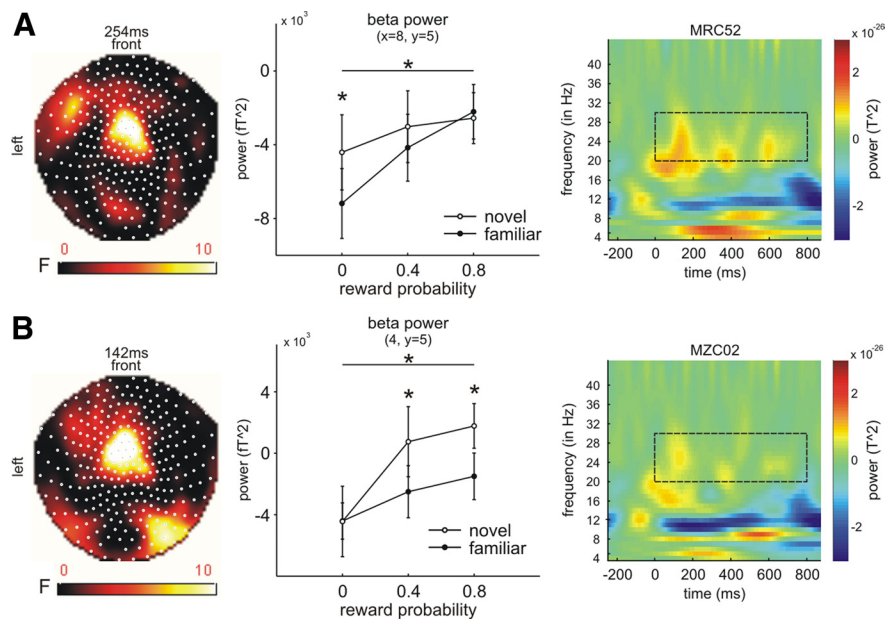
We used a reward anticipation paradigm with abstract cues predicting three different reward probabilities, presented in the con-

text of either a novel or a familiar image. Reward anticipation was represented in the ERFs and theta (5–8 Hz) and beta (20–30 Hz) oscillatory power with an onset latency of  $\sim 100$  ms. Over right temporal sensors higher reward probabilities led to more negative deflections in the ERFs (Fig. 2A). Furthermore, with higher reward probability, theta power decreased over right temporal and centroparietal sensors (Fig. 3B) whereas beta power increased over frontocentral sensors (Fig. 3C).

As hypothesized (Kakade and Dayan, 2002; Lisman and Grace, 2005) contextual novelty enhanced the neural responses to reward-predicting cues. This enhancement was visible in the ERFs and beta oscillatory power at a very early time window ( $\sim 100$  ms) after cue onset. These data indicate that reward-predicting cues are processed as rapidly in humans as in non-human primates and that contextual novelty modulates the neural representation of learned cue-reward probabilities per se rather than influencing later stages of reward anticipation.

The finding of rapid ( $\sim 100$  ms) neural reward anticipation signals in the human brain is compatible with animal studies that show neural reward anticipation with an onset latency of  $\sim 100$  ms in a series of brain regions (Schultz, 2007). In non-human primates neural reward anticipation at  $\sim 100$  ms after cue-onset has been demonstrated in the dopaminergic midbrain (Tobler et al., 2005), prefrontal cortex (Watanabe, 1996), basal ganglia (Kawagoe et al., 1998), and parietal cortex (Platt and Glimcher, 1999). In rats, even the primary visual cortex signals reward anticipation on a subsecond time scale (Shuler and Bear, 2006). The present data are the first demonstration in healthy human subjects that neural reward anticipation signals arise at a similar time scale as observed in animals. Moreover, they comply with the idea that such early neural reward responses may guide behavior and decision making.

Electrophysiological studies in experimental animals, and fMRI studies in humans, have identified neural signals to predictions of upcoming rewards in both cortical and subcortical brain areas (for review, see O'Doherty, 2004; Schultz, 2004; Knutson and Cooper, 2005). Apart from the well known mesolimbic areas, a recent study in humans has shown that BOLD signals in the middle occipital gyrus correlate with subjective value in a probabilistic context (Peters and Buchel, 2009). Similarly, in our previous experiment with fMRI we observed a cluster of activation within the superior occipital gyrus which expressed anticipated reward probability (Guitart-Masip et al., 2010). Therefore, these occipital areas would appear to be good candidates as neural generators for the effects that we observed in our ERFs over occipital/temporal sensors. Moreover, these effects add more evidence to the suggestion that visual areas are one of the initial regions of a widespread network responsive to signals of forthcoming rewards (Shuler and Bear, 2006).



**Figure 4.** Novelty bonus effects in the beta power. **A, B**, Over central sensors (**A**:  $x = 8, y = 5$ ; nearest channel: MRC52; **B**:  $x = 4, y = 5$ ; nearest channel: MZC02), beta power increased as a function of reward probability, and this effect was enhanced in the context of novelty for all reward probabilities (low, **A**; medium and high, **B**). Left columns show statistical parametric maps of the  $F$ -statistics, middle columns show the effects for each condition at the peak time point as extracted from voxel space, and right columns show the time-frequency plots as extracted from the nearest channel (before conversion to voxel space; see Materials and Methods). For **A**, right: p0 novel vs p0 familiar; for **B**, right: (p4 and p8) novel vs (p4 and p8) familiar. The dashed box indicates the beta frequency range (20–30 Hz) from 0 to 800 ms. It should be noted that the time-frequency data (right columns) were not rescaled by any logarithmic function (see Materials and Methods).

In physiological terms, beta and theta frequency band oscillations may allow for binding (Buzsáki and Draguhn, 2004) the distributed neural assemblies that jointly signal reward anticipation in response to a symbolic cue. Support for this notion comes from recent observations that identified theta and beta oscillatory activity as a signature of communication between brain regions during reward anticipation in animals (van Wingerden et al., 2010) or reward obtainment in humans using both EEG and MEG (Marco-Pallares et al., 2008; Donamayor et al., 2011). In our paradigm, theta power decreased with higher reward probability while beta showed the opposite pattern and increased. A link between beta oscillation in a reward learning task (gambling) and dopaminergic neuromodulation was reported in a recent EEG study (Marco-Pallares et al., 2009) where increased beta oscillation following monetary gains was modulated by genetic variability in the dopamine system. More precisely, beta power for gains was enhanced for carriers of the catechol-O-methyltransferase Val/Val allele compared with homozygous Met/Met carriers. On the basis of a recent theoretical approach (Bilder et al., 2004), it has been suggested that the carriers of the Val/Val allele show higher phasic dopamine responses (Marco-Pallares et al., 2009). Therefore, it is conceivable that our observed neural effects in the beta frequency range may (at least in part) reflect dopaminergic effects of neural reward anticipation.

Conceptually, the ability to learn to anticipate reward differs from processing reward at the time point of its presentation (outcome). While both reward anticipation and outcome processing can lead to increased vigor (Tobler et al., 2005; Pessiglione et al., 2007) the neural substrates that underlie reward anticipation and outcome processing can differ as revealed, for instance, by fMRI (O'Doherty, 2004). Our results add to the current electrophysio-

logical literature by showing parallels between reward anticipation and outcome: where both beta and theta oscillations seem to play a critical role. But they also show that the neural timing of anticipation and outcome processing differs. While signals of neural reward anticipation emerged already at  $\sim 100$  ms after cue onset (Fig. 2*B*) earlier work in humans reported neural reward signals at outcome at  $\sim 200$ – $300$  ms (Yeung and Sanfey, 2004; Cohen et al., 2007; San Martín et al., 2010). This temporal difference further suggests that different neural substrates can underlie anticipation and outcome processing of reward.

As expected, the context of novelty modulated neural reward anticipation responses. This finding is congruent with a recent fMRI study with the same paradigm in which we showed that contextual novelty enhances striatal reward responses (Guitart-Masip et al., 2010). However, due to the slow time resolution of fMRI in this previous study, we were unable to disambiguate anticipation and outcome phase of reward processing. Our results here clearly demonstrate that contextual novelty can indeed enhance neural reward anticipation signals. Importantly, we observed that the effect of contextual novelty on cue processing was apparent as early as  $\sim 100$  ms. This early onset is well compatible with a physiological model of novelty processing involving a functional loop between the hippocampus and dopaminergic SN/VTA (Lisman and Grace, 2005; Bunzeck and Düzel, 2006; Bunzeck et al., 2007; Düzel et al., 2009). According to this model, hippocampal novelty signals can increase the number of tonically activated dopamine SN/VTA neurons. Because only tonically active dopamine neurons can be excited into phasic firing (Goto and Grace, 2008), SN/VTA responses to reward-predicting cues can be expected to be stronger in the context of novelty. Hence, our findings indicate that contextual novelty replaces neural responses that represent the initially learned cue-outcome contingencies. The alternative finding would have been that contextual effects were incorporated into reward prediction responses at a later point in time, hence modifying reward anticipation after the initially learned cue-outcome contingencies have been signaled.

It is important to note that the conclusion regarding the early physiological effects of novelty on reward cue processing rest solely on the timing of our observed MEG responses and does not depend on an ability to establish a firm link between MEG responses and subcortical dopaminergic circuitry. However, our observation that reward cue responses are prominent in the theta and beta frequency range does make contact with previous intracranial recordings in the human nucleus accumbens showing increased theta oscillatory power (4–8 Hz) following losses (at outcome) in a gambling task (Cohen et al., 2009*b,c*). In the same task, stimulus locked beta and gamma power (20–80 Hz) increased to both reward and loss feedback and its synchronized activity with simultaneous alpha (8–12 Hz) dissociated between feedback conditions (Cohen et al., 2009*c*). Furthermore, intracranially recorded nucleus accumbens activity during cue and outcome processing correlated with activity of (scalp) surface electrodes near the vertex (Cohen et al., 2009*a*). MEG signals could relate to the reward-related oscillatory dynamics in the nucleus accumbens either because these are driven by cortical (e.g., medial prefrontal) input (Cohen et al., 2011) which can be detected by MEG or by virtue of direct back-projections.

It should be noted that the ERF analysis, time-frequency analysis, and our previous fMRI study (Guitart-Masip et al., 2010) revealed slightly divergent but complementary results. In particular, the ERF results showed a robust enhancement of neural reward anticipation in the context of novelty only for the low reward probability but not the higher reward probabilities (for

0.4 and 0.8 only at lower statistical threshold) which is in line with the effects observed in the ventral striatum (Guitart-Masip et al., 2010). The time-frequency analysis, on the other hand, showed robust contextual novelty effects for the high reward probabilities (0.4, 0.8) and for the low (0) reward probability only at a lower statistical threshold. This pattern is similar to that observed in the SN/VTA (Guitart-Masip et al., 2010). Our ERF and time-frequency data, as well as our previous fMRI results, support the suggestion that novelty acts as an exploration bonus for rewards by increasing expected reward probabilities (Kakade and Dayan, 2002). Furthermore, the MEG and fMRI data together demonstrate that novelty can influence a wide range of learned probability values. It remains to be determined, however, why ERF, time-frequency and BOLD-fMRI are sensitive to different probability values of the novelty bonus. Certainly, this differential sensitivity is compatible with the suggestion that ERFs, time-frequency data and fMRI are complementary (Cohen et al., 2007).

In this study, we have conceptualized novelty as the differential effects exerted by novel and preexposed items (Henson, 2003). fMRI studies have suggested that neural differences between novel and preexposed items entail two types of effects (e.g., Johnson et al., 2008). One type of effect linearly relates to the number of preexposures and hence is likely to reflect a degree of stimulus familiarity. The other effect categorically distinguishes novel items from preexposed ones and hence reflects novelty. Our experimental design was not suited to distinguish between these two types of effect and hence our results are neutral with respect to the question whether a familiarity type of response could exert a novelty bonus.

To summarize, we show that anticipatory reward responses are generated as rapidly in humans as in non-human primates, namely within  $\sim 100$  ms after the onset of reward-predicting cues. Event-related fields and beta oscillations appear to signal complementary aspects of reward anticipation and may reflect activity of different parts of the reward circuitry. Furthermore, these very rapid neural responses to reward-predicting stimuli are enhanced in the context of novelty compatible with the possibility that the primary response properties of reward circuitry can be changed such that originally learned reward probabilities are replaced by contextually altered versions. Hence, the neural processing of cues that predict future rewards is temporally highly efficient and contextually modifiable.

## References

- Aberg C, Nilsson LG (2001) Facilitation of source discrimination in the novelty effect. *Scand J Psychol* 42:349–357.
- Axmacher N, Mormann F, Fernández G, Elger CE, Fell J (2006) Memory formation by neuronal synchronization. *Brain Res Rev* 52:170–182.
- Bilder RM, Volavka J, Lachman HM, Grace AA (2004) The catechol-O-methyltransferase polymorphism: relations to the tonic-phasic dopamine hypothesis and neuropsychiatric phenotypes. *Neuropsychopharmacology* 29:1943–1961.
- Bunzeck N, Düzel E (2006) Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron* 51:369–379.
- Bunzeck N, Schütze H, Stallforth S, Kaufmann J, Düzel S, Heinze HJ, Düzel E (2007) Mesolimbic novelty processing in older adults. *Cereb Cortex* 17:2940–2948.
- Bunzeck N, Doeller CF, Fuentemilla L, Dolan RJ, Düzel E (2009) Reward motivation accelerates the onset of neural novelty signals in humans to 85 milliseconds. *Curr Biol* 19:1294–1300.
- Buzsáki G (2002) Theta oscillations in the hippocampus. *Neuron* 33:325–340.
- Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 304:1926–1929.
- Cavanagh JF, Frank MJ, Klein TJ, Allen JJ (2010) Frontal theta links predic-

- tion errors to behavioral adaptation in reinforcement learning. *Neuroimage* 49:3198–3209.
- Christie GJ, Tata MS (2009) Right frontal cortex generates reward-related theta-band oscillatory activity. *Neuroimage* 48:415–422.
- Cohen MX, Elger CE, Ranganath C (2007) Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage* 35:968–978.
- Cohen MX, Axmacher N, Lenartz D, Elger CE, Sturm V, Schlaepfer TE (2009a) Neuroelectric signatures of reward learning and decision-making in the human nucleus accumbens. *Neuropsychopharmacology* 34:1649–1658.
- Cohen MX, Axmacher N, Lenartz D, Elger CE, Sturm V, Schlaepfer TE (2009b) Good vibrations: cross-frequency coupling in the human nucleus accumbens during reward processing. *J Cogn Neurosci* 21:875–889.
- Cohen MX, Axmacher N, Lenartz D, Elger CE, Sturm V, Schlaepfer TE (2009c) Nuclei accumbens phase synchrony predicts decision-making reversals following negative feedback. *J Neurosci* 29:7591–7598.
- Cohen MX, Bour L, Mantione M, Figeé M, Vink M, Tijssen MA, Rootselaar AF, Munckhof PV, Richard Schuurman P, Denys D (2011) Top-down-directed synchrony from medial frontal cortex to nucleus accumbens during reward anticipation. *Hum Brain Mapp*. Advance online publication. Retrieved May 27, 2011. doi:10.1002/hbm.21195.
- Donamayor N, Marco-Pallares J, Heldmann M, Schoenfeld MA, Munte TF (2011) Temporal dynamics of reward processing revealed by magnetoencephalography. *Hum Brain Mapp*. Advance online publication. Retrieved February 28, 2011. doi:10.1002/hbm.21184.
- Düzel E, Habib R, Guderian S, Heinze HJ (2004) Four types of novelty-familiarity responses in associative recognition memory of humans. *Eur J Neurosci* 19:1408–1416.
- Düzel E, Bunzeck N, Guitart-Masip M, Wittmann B, Schott BH, Tobler PN (2009) Functional imaging of the human dopaminergic midbrain. *Trends Neurosci* 32:321–328.
- Düzel E, Penny WD, Burgess N (2010) Brain oscillations and memory. *Curr Opin Neurobiol* 20:143–149.
- Furl N, van Rijsbergen NJ, Treves A, Friston KJ, Dolan RJ (2007) Experience-dependent coding of facial expression in superior temporal sulcus. *Proc Natl Acad Sci U S A* 104:13485–13489.
- Gonsalves BD, Kahn I, Curran T, Norman KA, Wagner AD (2005) Memory strength and repetition suppression: multimodal imaging of medial temporal cortical contributions to recognition. *Neuron* 47:751–761.
- Goto Y, Grace AA (2008) Limbic and cortical information processing in the nucleus accumbens. *Trends Neurosci* 31:552–558.
- Guitart-Masip M, Bunzeck N, Stephan KE, Dolan RJ, Düzel E (2010) Contextual novelty changes reward representations in the striatum. *J Neurosci* 30:1721–1726.
- Hasselmo ME (2006) The role of acetylcholine in learning and memory. *Curr Opin Neurobiol* 16:710–715.
- Henson RN (2003) Neuroimaging studies of priming. *Prog Neurobiol* 70:53–81.
- Henson RN, Mouchlianitis E, Matthews WJ, Kouider S (2008) Electrophysiological correlates of masked face priming. *Neuroimage* 40:884–895.
- Holroyd CB, Coles MG (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 109:679–709.
- Johnson JD, Muftuler LT, Rugg MD (2008) Multiple repetitions reveal functionally and anatomically distinct patterns of hippocampal activity during continuous recognition memory. *Hippocampus* 18:975–980.
- Kakade S, Dayan P (2002) Dopamine: generalization and bonuses. *Neural Netw* 15:549–559.
- Kawagoe R, Takikawa Y, Hikosaka O (1998) Expectation of reward modulates cognitive signals in the basal ganglia. *Nat Neurosci* 1:411–416.
- Kiebel SJ, Friston KJ (2004a) Statistical parametric mapping for event-related potentials: I. Generic considerations. *Neuroimage* 22:492–502.
- Kiebel SJ, Friston KJ (2004b) Statistical parametric mapping for event-related potentials: II. A hierarchical temporal model. *Neuroimage* 22:503–520.
- Kilner J, Friston K (2010) Topological inference for EEG and MEG. *Ann Appl Statistics* 4:1272–1290.
- Klimesch W, Freunberger R, Sauseng P, Gruber W (2008) A short review of slow phase synchronization and memory: evidence for control processes in different memory systems? *Brain Res* 1235:31–44.
- Knutson B, Cooper JC (2005) Functional magnetic resonance imaging of reward prediction. *Curr Opin Neurol* 18:411–417.
- Lisman JE, Grace AA (2005) The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46:703–713.
- Litvak V, Mattout J, Kiebel S, Phillips C, Henson R, Kilner J, Barnes G, Oostenveld R, Daunizeau J, Flandin G, Penny W, Friston K (2011) EEG and MEG data analysis in SPM8. *Comput Intell Neurosci* 2011:852961.
- Marco-Pallares J, Cucurell D, Cunillera T, García R, Andrés-Pueyo A, Münte TF, Rodríguez-Fornells A (2008) Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia* 46:241–248.
- Marco-Pallares J, Cucurell D, Cunillera T, Krämer UM, Càmarà E, Nager W, Bauer P, Schüle R, Schöls L, Münte TF, Rodríguez-Fornells A (2009) Genetic variability in the dopamine system (dopamine receptor D4, catechol-O-methyltransferase) modulates neurophysiological responses to gains and losses. *Biol Psychiatry* 66:154–161.
- Nieuwenhuis S, Yeung N, Holroyd CB, Schurger A, Cohen JD (2004) Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. *Cereb Cortex* 14:741–747.
- O'Doherty JP (2004) Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol* 14:769–776.
- Pessiglione M, Schmidt L, Draganski B, Kalisch R, Lau H, Dolan RJ, Frith CD (2007) How the brain translates money into force: a neuroimaging study of subliminal motivation. *Science* 316:904–906.
- Peters J, Büchel C (2009) Overlapping and distinct neural systems code for subjective value during intertemporal and risky decision making. *J Neurosci* 29:15727–15734.
- Platt ML, Glimcher PW (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400:233–238.
- Ranganath C, Rainer G (2003) Neural mechanisms for detecting and remembering novel events. *Nat Rev Neurosci* 4:193–202.
- Rugg MD, Curran T (2007) Event-related potentials and recognition memory. *Trends Cogn Sci* 11:251–257.
- San Martín R, Manes F, Hurtado E, Isla P, Ibañez A (2010) Size and probability of rewards modulate the feedback error-related negativity associated with wins but not losses in a monetarily rewarded gambling task. *Neuroimage* 51:1194–1204.
- Schultz W (2004) Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Curr Opin Neurobiol* 14:139–147.
- Schultz W (2007) Multiple dopamine functions at different time courses. *Annu Rev Neurosci* 30:259–288.
- Shuler MG, Bear MF (2006) Reward timing in the primary visual cortex. *Science* 311:1606–1609.
- Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons. *Science* 307:1642–1645.
- Tsivivilis D, Otten LJ, Rugg MD (2001) Context effects on the neural correlates of recognition memory: an electrophysiological study. *Neuron* 31:497–505.
- van Wingerden M, Vinck M, Lankelma J, Pennartz CM (2010) Theta-band phase locking of orbitofrontal neurons during reward expectancy. *J Neurosci* 30:7078–7087.
- Watanabe M (1996) Reward expectancy in primate prefrontal neurons. *Nature* 382:629–632.
- Weil RS, Kilner JM, Haynes JD, Rees G (2007) Neural correlates of perceptual filling-in of an artificial scotoma in humans. *Proc Natl Acad Sci U S A* 104:5211–5216.
- Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC (1996) A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp* 4:58–73.
- Worsley KJ, Taylor JE, Tomaiuolo F, Lerch J (2004) Unified univariate and multivariate random field theory. *Neuroimage* 23 [Suppl 1]:S189–S195.
- Wu Y, Zhou X (2009) The P300 and reward value, magnitude, and expectancy in outcome evaluation. *Brain Res* 1286:114–122.
- Yeung N, Sanfey AG (2004) Independent coding of reward magnitude and valence in the human brain. *J Neurosci* 24:6258–6264.
- Zaghloul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, Baltuch GH, Kahana MJ (2009) Human substantia nigra neurons encode unexpected financial rewards. *Science* 323:1496–1499.