



HAL
open science

Reward Activates Stimulus-Specific and Task-Dependent Representations in Visual Association Cortices

Anne-Marike Schiffer, T Muller, Nick Yeung, Florian Waszak

► **To cite this version:**

Anne-Marike Schiffer, T Muller, Nick Yeung, Florian Waszak. Reward Activates Stimulus-Specific and Task-Dependent Representations in Visual Association Cortices. *Journal of Neuroscience*, 2014, 34 (47), pp.15610-15620. 10.1523/JNEUROSCI.1640-14.2014 . hal-02385067

HAL Id: hal-02385067

<https://hal.science/hal-02385067>

Submitted on 28 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reward Activates Stimulus-Specific and Task-Dependent Representations in Visual Association Cortices

Anne-Marike Schiffer,¹ Timothy Muller,¹ Nick Yeung,¹ and Florian Waszak^{2,3}

¹Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom, ²Université Paris Descartes, Sorbonne Paris Cité, 75270 Paris, France, and ³CNRS Laboratoire Psychologie de la Perception, Unité Mixte de Recherche 8242, 75006 Paris, France

Humans reliably learn which actions lead to rewards. One prominent question is how credit is assigned to environmental stimuli that are acted upon. Recent functional magnetic resonance imaging (fMRI) studies have provided evidence that representations of rewarded stimuli are activated upon reward delivery, providing possible eligibility traces for credit assignment. Our study sought evidence of postreward activation in sensory cortices satisfying two conditions of instrumental learning: postreward activity should reflect the stimulus category that preceded reward (stimulus specificity), and should occur only if the stimulus was acted on to obtain reward (task dependency). Our experiment implemented two tasks in the fMRI scanner. The first was a perceptual decision-making task on degraded face and house stimuli. Stimulus specificity was evident as rewards activated the sensory cortices associated with face versus house perception more strongly after face versus house decisions, respectively, particularly in the fusiform face area. Stimulus specificity was further evident in a psychophysiological interaction analysis wherein face-sensitive areas correlated with nucleus accumbens activity after face-decision rewards, whereas house-sensitive areas correlated with nucleus accumbens activity after house-decision rewards. The second task required participants to make an instructed response. The criterion of task dependency was fulfilled as rewards after face versus house responses activated the respective association cortices to a larger degree when faces and houses were relevant to the performed task. Our study is the first to show that postreward sensory cortex activity meets these two key criteria of credit assignment, and does so independently from bottom-up perceptual processing.

Key words: credit assignment; fMRI; reward-related learning; stimulus-specific postreward activation

Introduction

Humans learn how to act on stimuli to gain reward. Substantial effort has been devoted to understanding how reward delivery fosters associative learning (Rescorla and Wagner, 1972; Schultz, 2007). This research has revealed that reward-driven learning depends on midbrain dopamine neurons, which display a firing pattern resembling reward prediction error signals in models of reinforcement learning (Schultz et al., 1997; Waelti et al., 2001). However, whereas computational approaches provide solutions to the critical problem of credit assignment—determining which features are predictive of positive outcomes—little is known about how such eligibility traces (Sutton and Barto, 1990) are represented in the brain. In this study, we aimed to identify neu-

ral signatures of potential eligibility traces, i.e., stimulus representations that have two crucial properties. To guarantee precision of ensuing reward predictions, activated representations should be stimulus specific and task dependent. Stimulus specificity ensures that the precise environmental conditions that preceded reward will trigger its prediction. Task dependency warrants that environmental conditions are only associated with reward if they were used to perform the rewarded action.

Some recent studies have investigated related questions, focusing on the hypothesis that learning should depend on activation of stimulus representations at the time of reward delivery (Pleger et al., 2008, 2009; Weil et al., 2010; FitzGerald et al., 2012; Arsenault et al., 2013). However, whereas functional magnetic resonance imaging (fMRI) in animals has demonstrated stimulus-specific reward-related activity (Arsenault et al., 2013), corresponding evidence in human studies has not been consistently observed (Weil et al., 2010; FitzGerald et al., 2012). It therefore remains unclear whether reward-based activity in human sensory cortex is stimulus specific.

To investigate this question, we conceived a novel paradigm in which subjects performed a perceptual discrimination task, deciding whether degraded stimuli contained images of faces or houses. The analysis focused on trials in which, unbeknownst to participants, the stimulus was pure noise. This renders activation by reward independent of initial bottom-up activation as well as of potential category-specific reward expectations, while mini-

Received April 23, 2014; revised Aug. 12, 2014; accepted Aug. 17, 2014.

Author contributions: A.-M.S., N.Y., and F.W. designed research; A.-M.S. and T.M. performed research; A.-M.S., T.M., N.Y., and F.W. analyzed data; A.-M.S., N.Y., and F.W. wrote the paper.

This work is supported by Biotechnology and Biological Sciences Research Council Grant BB/1019847/1, awarded to N.Y. and F.W.

The authors declare no competing financial interests.

This article is freely available online through the JNeurosci Author Open Choice option.

Correspondence should be addressed to Anne-Marike Schiffer at the above address. E-mail: anne-marike.schiffer@psy.ox.ac.uk.

DOI:10.1523/JNEUROSCI.1640-14.2014

Copyright © 2014 Schiffer et al.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

mizing the possible effects of neural adaptation (Grill-Spector et al., 2006; FitzGerald et al., 2012).

We sought evidence of reward-dependent, stimulus-specific cortical activity; that is, activity in our regions of interest (ROIs), the fusiform face area (FFA) and the parahippocampal place area (PPA), at the time of reward delivery. This was our first criterion for a neural signature that could serve as an eligibility trace. Our second criterion was task dependency. Postreward activation should be stronger for stimuli that were used to gain reward. We therefore compared postreward activation in the perceptual decision task with activation in a second, instructed response task, hypothesizing that stimulus activity for reward outcomes would be restricted to trials in which outcomes were experienced as a consequence of a perceptual decision.

In summary, we predicted that activity in the ROIs would show a positive correlation with reward size for associated decisions (stimulus specificity), and would be more influenced by reward following a perceptual decision than following an instructed response (task dependency).

Materials and Methods

Eighteen right-handed, healthy participants (10 women; ages 20–32 years mean age 24 years) took part in the study. The participants reported no psychiatric or neurological past or present condition. All procedures were approved by the local ethics committee of the University of Oxford and all participants gave written informed consent.

Stimulus material

Grayscale photographs of faces and front views of houses (Fig. 1) served as stimulus material. In a first step, all images were adjusted in luminance and spatial frequency to the mean of the stimulus pool using the SHINE (Spectrum, Histogram and Intensity Normalization and Equalization) Matlab tool (Willenbockel et al., 2010). This measure was taken to prevent categorization based on surface similarities (Schyns and Oliva, 1994; Rajimehr et al., 2011). Images were Fourier transformed and a variable percentage of all phases in each Fourier transformed image was scrambled. Images were then back-transformed into native space. Three degrees of phase scrambling were applied per category to yield stimuli producing easy, medium, or hard levels of difficulty in perceptual discrimination. Face images were phase scrambled to 70, 75, and 85%. House images were phase scrambled to 50, 65, or 75%. These degrees of scrambling were chosen based on pilot testing to produce comparable performance for house and face stimuli across the three levels of degradation. In addition to the three difficulty levels, half of the images were pure noise images with 100% of all phases scrambled.

Task

In each trial, participants were first presented with a stimulus image for 2 s. Stimulus presentation was followed by a task cue, displayed on the screen for 1.5 s; the task cue was either a question mark or an exclamation mark. Question marks instructed participants to press the left or right button to indicate whether they had seen a face or a house (perceptual decision task). Participants were unaware that half of the images were noise images and instructed to always decide and respond. Images of exclamation marks contained a darkened box on the left or right side underneath the exclamation mark (Fig. 1A). In these trials, participants had to press the button on the side corresponding to the darkened box (instructed response task). Importantly, since participants did not know the trial type at the time of stimulus presentation, they had to make a perceptual decision in all trials. Participants had 1.5 s to respond, after which the task image stayed on the screen for the remaining response–stimulus interval (RSI). A cross in a box to the left or right of the task image was displayed during this interval, indicating their previous response. The length of this RSI was randomly drawn from a Poisson distribution with $\lambda = 4$ s, minimum 2 s, maximum 6 s, and jittered in steps of 500 ms. This interval was followed by feedback, which could be rewarding, neutral, or penalizing.

Rewards consisted of images showing either one or two moneybags, indicating gains of 10 or 20 points, respectively. Penalties were shown as one or two bombs, indicating loss of either 10 or 20 points. Participants were told that rewards and penalties were contingent on the correctness of their previous response, but were in fact randomly assigned in noise trials. Participants received feedback on their accumulated score every 50 trials. Their final score was converted into a monetary bonus of $\leq \text{£}5$ after the scan, which was added to their usual remuneration of $\text{£}20$. The experimental sequence in the scanner consisted of 275 trials, 138 of which were noise trials. Thirty-five noise trials and 33 signal trials were instruction trials; 24–26 noise question trials were followed by a neutral outcome, the remaining noise trials were followed in equal numbers by large rewards, small rewards, small penalties, and large penalties. Outcomes in the signal trials were contingent on performance, but outcome size was randomly determined. Just before the scanning session, participants performed 16 practice trials of the experimental paradigm.

The experimental task was followed by a functional localizer task to determine the ROIs for all planned contrasts. Participants performed a one-back task while they were presented with two sequences of six blocks of 18 images that appeared on the screen for 150 ms each, followed by an interstimulus interval of 400 ms. Participants had a short break after the first six blocks. They had to switch from making responses with one hand to the other after the break. Each block contained images of only one category; these categories were as follows: unscrambled face images; unscrambled house images; easy, medium, and hard face images; easy, medium, and hard house images; pure noise images; and unscrambled object images.

Behavioral analysis

Behavior in the task was recorded to establish that participants showed the expected performance modulation by the degree of phase scrambling of the signal stimuli, as well as to assess changes in performance over time and win–stay/lose–shift behavior as markers of learning. We also tested whether participants made both face and house judgments in trials with pure noise stimuli.

fMRI procedure

The functional imaging session took place in a 3T Siemens Magnetom Trio scanner (Siemens). During the scan, participants lay supine on the scanner bed with their left and right index fingers resting on two buttons of a centrally placed response box. Participants wore sound-attenuating headphones that allowed communication with the experimenter. They viewed the stimuli on the screen via a mirror built into the head coil. Stimuli were displayed at 5° of visual angle to prevent head and eye movements. The functional session engaged a single-shot gradient echo-planar imaging (EPI) sequence sensitive to blood oxygen level-dependent (BOLD) contrast (32 slices, 192 mm field of view, 4 mm slice thickness, $3 \times 3 \times 4$ in-plane resolution, orientation parallel to the bicommissural plane, 30 ms TE, 90° flip angle, 2000 ms TR, interleaved, descending recording). After the functional session was completed, high-resolution 3D T1-weighted whole-brain modified driven equilibrium Fourier transformation (MDEFT) sequences were recorded for every participant (128 slices, 256 mm field of view, 256×256 pixel matrix, 1 mm slice thickness, 0.25 mm spacing).

fMRI data analysis

fMRI data analysis was conducted with the LIPSIA (Leipzig Image Processing and Statistical Inference Algorithms) processing tool (Lohmann et al., 2001). For spatial registration, EPI data and 3D MDEFT data were first oriented along the ac – pc axis. The matching parameters (six degrees of freedom, three rotational, three translational) of the functional data onto the individual 3D MDEFT reference set were used to calculate the transformation matrices for linear registration. These matrices were subsequently normalized to Talairach brain size ($x = 135$ mm, $y = 175$ mm, $z = 120$ mm; Talairach and Tournoux, 1988) by linear scaling. The normalized transformation matrices were then applied to the functional slices, to transform them using trilinear interpolation and align them with the 3D reference set in the stereotactic coordinate system. The generated output had a spatial resolution of $3 \times 3 \times 3$ mm. Cubic-spline interpolation was used to correct for the temporal offset between the

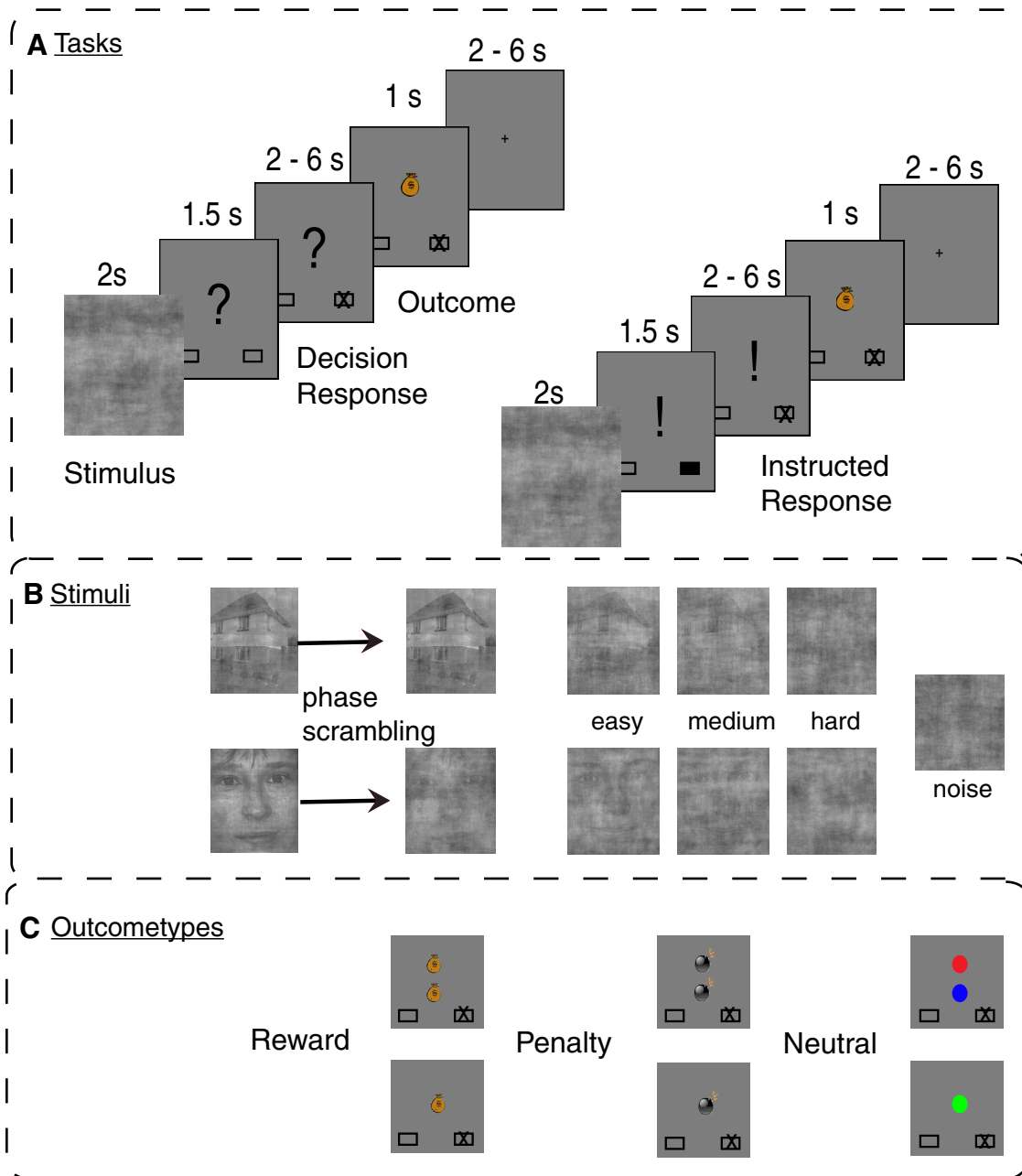


Figure 1. *A*, The main task was a perceptual decision task in which participants were first presented with a stimulus that they had to classify as either face or house. When the question mark appeared, participants indicated their decision with a left or right button press. They then received positive (as shown), negative, or neutral feedback. The second task made up 25% of all trials. Here, the initial stimulus was followed by an exclamation mark on top of two boxes, one of which was darkened; participants had to press the button corresponding to the side of the darkened box. Feedback was again delivered in the same format as for the perceptual decision task and contingent on response accuracy. *B*, Luminance-adjusted grayscale images of faces and houses were Fourier transformed and phase scrambled for use as stimuli. Within the experiment, three levels of degradation and noise trials were included, yielding graded performance levels (the displayed levels of phase scrambling for signal stimuli are for illustrative purposes only and deviate from the actual levels in the main experiment; see Materials and Methods). Noise stimuli were 100% phase scrambled. Participants were unaware of the existence of noise stimuli. *C*, Participants experienced five levels of outcome valence, two levels of reward, two levels of penalty, and a neutral outcome. Large and small rewards or penalties resulted in the gain or loss of 20 or 10 points, respectively. Outcome valence (reward or penalty) was performance contingent on signal trials and in the instructed response task, but randomly assigned for perceptual decisions in noise trials. A quarter of all trials were followed by neutral outcomes instead independent of performance.

slices acquired in one scan. To remove low-frequency signal changes and baseline drifts, a high-pass filter of 1/75 Hz was applied for event-related analysis and a high-pass filter of 1/125 Hz was applied to the analysis of the localizer blocks. Filter lengths were chosen based on the optimal filter length for a specific design file as suggested by LIPSIA. Statistical evaluation was based on a least-square estimation using the general linear model (GLM) for serially autocorrelated observations (Worsley and Friston, 1995). Temporal Gaussian smoothing (4 s FWHM) was applied to deal with temporal autocorrelation and determine the degrees of free-

dom (Worsley and Friston, 1995). A spatial Gaussian filter of 5.65 mm FWHM was applied. Unless otherwise stated, the design matrix was generated by hemodynamic modeling using a γ -function in all contrasts. The onset vectors were modeled with a duration of 1 s in a time-locked event-related fashion unless otherwise stated. No derivatives were included in the models.

For group analyses, *t* tests and repeated-measures ANOVAs were performed on β values from the contrast of regressors in the respective GLMs. Acquired *t* values were transformed to *z* scores. For whole-brain

analyses, an initial z threshold of 2.56 ($p < 0.01$, one tailed) was then applied to the activation map. All voxels showing a positive activation above this threshold entered the second step of the correction. Here, a Monte Carlo simulation was used to define thresholds for cluster size and cluster value at a significance level of $p < 0.05$ (one tailed). The combination of cluster size and cluster value decreases the risk of neglecting true activations in small structures. Thus, all reported activations were significant at $p < 0.05$, corrected for multiple comparisons at the cluster level.

ROI definition. Functional ROIs were determined in a two-step approach. As a first step, all blocks from the functional localizer that contained house images, all blocks that contained face images, and the object image blocks were entered separately as regressors into a GLM (GLM1). Events were modeled with a box-car function and event length set to block length. House blocks were contrasted with face blocks (HouseLocalizerBlock > FaceLocalizerBlock) and vice versa (FaceLocalizerBlock > HouseLocalizerBlock) on the single-subject level and averaged into t map contrast images. In a separate analysis, the face and house signal trials from the main experiment were entered separately into another GLM (GLM2). A parametric increase for event amplitudes was determined by signal strength, with amplitude increasing from 1 to 3 for hard to easy trials. The regressor accounting for house trials was then contrasted with the face-trial regressor (HouseSignalStrength > FaceSignalStrength) and vice versa (FaceSignalStrength > HouseSignalStrength). The resulting contrast images were masked with the contrast images generated based on the functional localizer task. The masked images then entered second-level random-effects analysis. One-sample t tests were used for the group analyses across the contrast images of all subjects to test whether observed differences between conditions were significantly different from zero. The bilateral peak voxels of activity in the parahippocampal gyri were used as centers for the PPA ROI. ROIs were established as $2 \times 2 \times 2$ voxel cubes centered on the bilateral peak coordinates. Peak voxels for the FFA ROI were generated in a parallel approach, locating peak voxels in the fusiform gyrus; the bilateral ROI was set as a cube of $2 \times 2 \times 2$ voxels around these centers.

Decision-specific activation at noise stimulus presentation. To establish whether noise stimuli were treated as if they contained signal, a first contrast tested whether the ROIs would show significant activation in line with the perceptual decision on noise trials. Two regressors were entered into one GLM (GLM3), one accounting for the presentation of noise stimuli that were followed by a house decision and the corresponding regressor for noise stimuli that were followed by a face decision. Events were time-locked to noise stimulus presentation, modeled with an event length of 1 s and event amplitude of 1. We estimated the main effect of each regressor separately and contrasted face noise trials with house noise trials (FaceDecision > HouseDecision). The mean β scores extracted from the FFA and PPA ROIs entered a repeated-measures ANOVA to estimate main effects and interactions of decision (face or house) and ROI (FFA or PPA).

Reward network response. The second contrast aimed to show that reward after noise trials would result in the network response associated with learning from rewards. We parametrically modeled BOLD increase from neutral to large reward trials after both face and house decisions to noise stimuli. Events were time-locked to reward presentation and modeled with an event length of 1 s and an event amplitude ranging from 1 (neutral) to 3 (large reward).

Stimulus-specific activation at reward outcome. The most critical features for postreward activation to be qualified as a signature of reward-driven learning were stimulus specificity and task dependency. Analyses for both effects were conducted in the same GLM (GLM5), which contained the following eight regressors. The four main regressors of interest implemented a parametric modulation of BOLD by reward size: (1) after house responses in noise decision trials, (2) after face responses in noise decision trials, (3) after instructed “house” responses, and (4) after instructed “face” responses. Large rewards were modeled with an amplitude of 2 and small rewards were modeled with an amplitude of 1. Further, this GLM contained separate regressors of no interest for penalty outcomes modeled by size after face decisions and house decisions and separate regressors for neutral outcomes after each type of decision.

Events were modeled with a 1 s duration, time-locked to reward presentation. The main parametric contrasts to estimate stimulus specificity were FaceDecisionReward and HouseDecisionReward. The estimated β values for postreward activity scaling with reward size in the FFA and PPA for face versus house decisions respectively were entered into a 2×2 repeated measures ANOVA to test for stimulus-specific postreward activation. Because these analyses were calculated on noise trials, with pseudorandomized response–reward contingencies, there was no significant correlation between reward size and stimulus-specific regressors related to the earlier perceptual decision ($r = 0.034$).

Task-dependent activation at reward outcome. The second parametric contrast calculated from this GLM (GLM5) tested for the assumption that reactivation should be task dependent, i.e., depend on a perceptual decision, as opposed to an instructed response. To test this hypothesis, we performed a repeated-measures ANOVA on the β values from the four regressors of interest: FaceDecisionReward, FaceInstructionReward, HouseDecisionReward, and HouseInstructionReward. Further, we directly contrasted FaceDecisionReward > FaceInstructionReward in the FFA ROI. Events were modeled time-locked to reward presentation with an event duration of 1 s. Reward size for instructed responses was modeled parametrically with the same amplitude vector as in the stimulus-specificity analysis.

Reward-driven activation of stimulus-specific areas: psychophysiological interaction analysis. To investigate whether proposed effects of stimulus-specific BOLD increase with reward size were linked to activity in regions associated with reward processing, we conducted a psychophysiological interaction analysis (PPI) to complement the interaction analysis outlined above. The PPI procedure aimed to establish whether activity in areas coding for reward magnitude would correlate with activity in areas showing stimulus-specific activation. This analysis was implemented in three steps. First, we established the peak voxel of activity in the left and right nucleus accumbens in the contrast measuring the reward-network response (above). We then conducted two PPI analyses for each hemisphere. One of these correlated the seed voxel with whole-brain activation in a contrast of house reward after noise trials versus face reward after noise trials, while the second correlated the same seed voxel with the reverse contrast. PPI analyses used an unconvolved regressor and modeled event lengths of 3 TRs (O’Reilly et al., 2012). These four PPI analyses (two hemispheres \times two contrasts) were then combined in two conjunction analyses. The first conjunction combined the left nucleus accumbens-seed PPI in house reward after noise trials versus face reward after noise trials with the corresponding right nucleus accumbens-seed PPI (house-vs-face PPI, hereafter). The second conjunction analogously combined the face reward after noise trials versus house reward after noise trials PPIs from both hemispheres (face-vs-house PPI, hereafter). LIPSIA controls for the inflated α -error in conjunction analyses (Lohmann et al., 2001). The second step was to determine areas within the vicinity of the PPA ROI that were significantly activated in the house-vs-face PPI analysis and determining areas within the vicinity of the FFA ROI that were significantly activated in the face-vs-house PPI analysis. Selection of these ROIs was performed on not cluster-corrected data, as the search radius was limited based on previous ROIs and anatomy. Last, the areas identified in step two were used as ROIs (masks of a $2 \times 2 \times 2$ cube in each hemisphere) in the localizer contrast of face blocks versus house blocks and house blocks versus face blocks (GLM1). This ensured that these ROIs not only correlated with nucleus accumbens activity after rewards corresponding to the perceptual decision in favor of one specific type of stimulus, but were also globally more sensitive to one stimulus category than the other. For example, areas that correlated with the nucleus accumbens more after face decisions on noise trials than house decisions on noise trials were required also to be more active during blocks of face images than blocks of house images in the functional localizer, thus supporting the argument that reward-related activity is stimulus specific.

Trial-type sensitivity. The last contrast tested whether noise stimuli would deliver a more sensitive context for post-reward activation effects than signal stimuli. We therefore defined a final GLM (GLM6), which included the following four main regressors to assess stimulus specificity and task dependency. These modeled reward magnitude parametrically

for house responses in noise decision trials, face responses in noise decision trials, instructed “house” responses in noise trials, and instructed “face” responses in noise trials. To allow comparison, the GLM further included the corresponding regressors for signal trials. Mean β s from all parametric contrasts were entered into a $2 \times 2 \times 2 \times 2$ repeated-measures ANOVA for further analysis.

Results

Behavioral analysis

Analysis of participants’ performance indicated that they engaged with the task, and confirmed that the paradigm effectively created three different levels of difficulty, with performance in the hardest level of difficulty being close to the fixed chance performance in noise trials. Participants made on average 75.8% correct responses (SD, 12.9%) on signal trials. One dataset was excluded from the analysis because performance was <2 SDs from the mean. The remaining 17 participants achieved on average 88.9, 79.9, and 63.3% correct on easy, medium, and hard face signal trials, respectively, and 92.1, 79.4, and 58.4% correct on the corresponding house signal trials. In instructed response trials, participants reached on average 89.27% correct responses (SD, 9.67%). According to postexperiment reports, errors on these instructed trials typically occurred when participants failed to suppress a prepared response to the stimulus. To test that signal trials created a plausible context for noise trials, but were at the same time not clearly distinguishable from noise trials, we assessed how many levels of degradation participants thought they had encountered. Of 17 participants, 9 indicated that the experiment implemented three levels of difficulties, while 4 participants believed that there had been “3–4” levels of difficulty. Only 2 participants correctly estimated that there had been four levels of difficulty, while the remaining participants indicated five, and 50 levels of degradation, respectively. It thus appears as if for most subjects, noise trials were not clearly distinguishable from signal trials. Incidentally, only one participant reported noticing that a few trials did not contain any signal, but nevertheless did not realize that fully half of all trials were noise trials.

In a next step, we assessed whether participants made use of feedback to adapt their behavior. We therefore assessed participants’ performance changes over the course of the experiment. As expected, performance improved, as revealed in a 2×3 repeated-measures ANOVA with the factors TIME (level: first half of experiment/second half of experiment) and DEGRADATION (level: easy/medium/hard). This analysis revealed a marginally significant main effect of time ($F_{(1,15)} = 4.41$, $p = 0.051$), a significant main effect of degradation ($F_{(1,16)} = 42.85$, $p = 0.000$), and a significant interaction between the two main factors ($F_{(1,16)} = 4.23$, $p = 0.032$). Descriptively, participants’ performance improved particularly on hard trials (Fig. 2A). It thus seemed that participants learned from feedback integrating this information to modify their behavior.

Because feedback was assigned randomly in noise trials (for which learning was impossible), performance modification by feedback for these trials was assessed on a trial-by-trial basis instead. To this end, we assessed for successive noise trials how likely participants were to repeat a rewarded response or switch

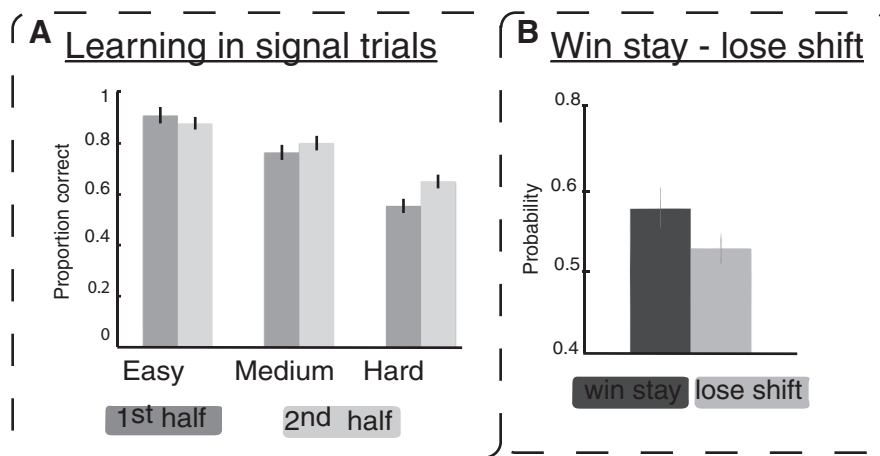


Figure 2. *A*, Feedback integration for signal stimuli was apparent in the performance improvement from the first half (dark gray bars) of the experiment to the second half (light gray bars) on perceptual decision trials. *B*, Feedback integration for noise trials was evident in win–stay/lose–shift behavior. Participants were significantly more likely to stay with a response than to shift response when the response was rewarded and more likely to shift than to stay when a response was penalized.

away from a penalized one. Such a win–stay/lose–shift behavior would indicate pseudolearning from positive feedback. A one-sample t test revealed significant difference in stay probabilities for rewarded compared with penalized noise trials ($t_{(16)} = 2.99$, $p = 0.004$), with participants being more likely to repeat a rewarded response. While this trial-by-trial analysis was primarily aimed at uncovering correlates of feedback integration in noise trials, which do not allow assessment of true learning, we also performed the same analysis on signal trials. We expected the effect in signal trials to be weaker, because use of feedback to guide behavior seems more plausible if the rewarded or penalized trial is visually similar to the current trial. Seven of 17 subjects made no errors on easy, or easy and medium trials, preventing us from comparing effects of noise trials separately for each level of degradation. We therefore averaged win–stay probability, win–shift probability, lose–stay probability, and lose–shift probability across all levels of degradation, including noise trials. To test whether win–stay/lose–shift behavior was influenced by signal strength, we calculated a repeated-measures ANOVA with the two-level factor DEGRADATION on trial $N - 1$ (signal/noise), VALENCE on trial $N - 1$ (reward/penalty), and BEHAVIOR on trial N (stay/shift). This analysis revealed no reliable main effects, a marginally significant interaction VALENCE on $N - 1 \times$ BEHAVIOR ($F_{(1,16)} = 4.251$, $p = 0.056$) and no three-way interaction. Thus, although win–stay/lose–shift behavior was numerically smaller after signal trials than noise trials, the difference was not significant.

As a final assessment of the credibility of the manipulation, and also to establish the comparability of BOLD effects across face versus house decisions, we compared the distribution of perceptual judgments on noise trials. Participants showed balanced judgments, with no strong preferences on the group level. Face decisions were on average made on 49% of all trials (SD, 7.8 standardized percentage; range, 31–61%).

fMRI analysis

The FFA ROI for all analyses except the PPI was derived by masking the FaceSignalStrength $>$ HouseSignalStrength contrast with the FacelocalizerBlock $>$ HouselocalizerBlock contrast, and was centered on the peak coordinates $x = -38$, $y = -51$, $z = -15$, and $x = 34$, $y = -60$, $z = -15$. The PPA ROI for all analyses except the PPI was derived by masking the HouseSignal-

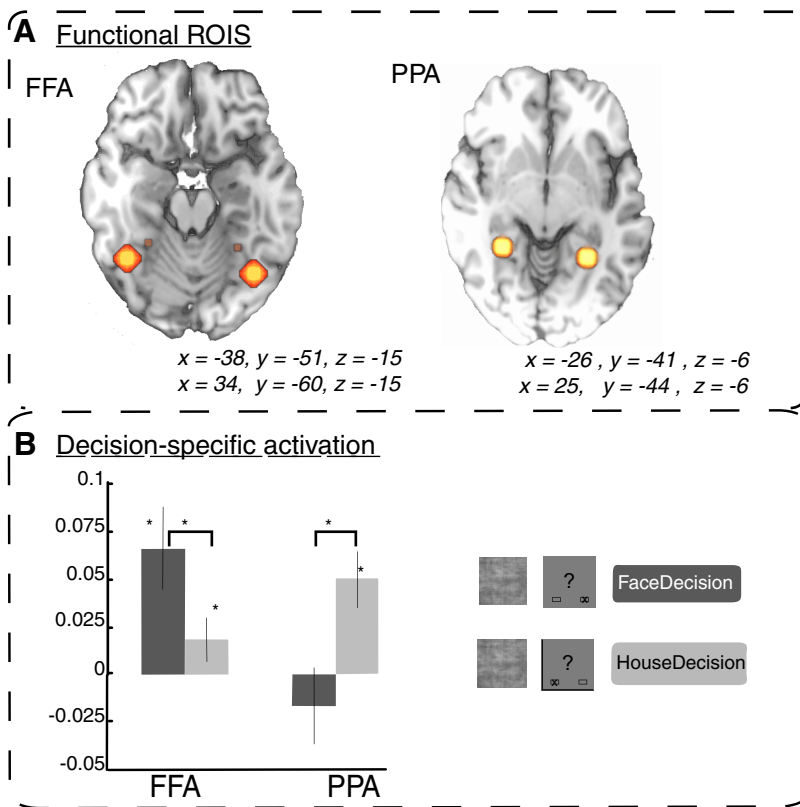


Figure 3. *A*, Functional ROIs were created by masking the contrast $\text{FacelocalizerBlock} > \text{HouselocalizerBlock}$ and $\text{HouselocalizerBlock} > \text{FacelocalizerBlock}$, respectively, with the corresponding parametric contrast for BOLD increase with signal strength from the main experiment. *B*, Contrasting BOLD response for stimuli preceding face decisions (dark gray) on signal trials with those preceding house decisions (light gray) revealed significantly stronger activity for face decisions in the FFA (left) and for house decisions in the PPA (right) compared with the respective alternative decision.

Strength $>$ FaceSignalStrength contrast with the $\text{HouselocalizerBlock} > \text{FacelocalizerBlock}$ contrast, and was centered on the peak coordinates in $x = -26, y = -41, z = -6$, and $x = 25, y = -44, z = -6$ (Fig. 3).

To determine whether participants treated noise stimuli as if they contained signal, we estimated the BOLD activity in the ROIs at the time when noise stimuli were presented in relation to the subsequent perceptual judgment. Activity in these stimulus-specific ROIs provided clear evidence that noise stimuli were treated as if they contained some (albeit weak) signal. Participants' individual β values for the two conditions—viewing noise stimuli that were then judged to be faces (FaceDecision) and viewing noise stimuli that were judged to be houses (HouseDecision)—were estimated in the two ROIs (Fig. 3A).

These individual β values were then entered into a repeated-measures ANOVA with the factors DECISION (face/house) and ROI (FFA/PPA). This yielded no significant effect of decision, but a significant main effect of ROI ($F_{(1,16)} = 5.93, p = 0.027$) and a significant interaction between DECISION and ROI ($F_{(1,16)} = 38.33, p < 0.001$). The significant interaction is further illustrated by the direct contrasts of conditions within the ROIs. These contrasts showed significantly more activity in the FFA for pending face versus house judgments ($t_{(16)} = 2.25, p = 0.019$) and significantly more activity in the PPA preceding house versus face judgments ($t_{(16)} = 3.39, p = 0.002$; Fig. 3B).

We further established that positive outcomes to noise trials would activate the network of brain regions associated with learning from reward (O'Doherty et al., 2003; O'Doherty, 2004). After correction for multiple comparisons at the whole-brain

level, a positive correlation of BOLD signal with reward magnitude was established in the hypothesized network of areas, classically associated with reward processing, including the right nucleus accumbens and right subgenual anterior cingulate gyrus/ventromedial PFC (vmPFC). The network further included bilateral hippocampal activation (see Fig. 5A).

Stimulus-specific activation at reward outcome

The primary aim of the present study was to identify neural correlates of reward-driven learning, reflected in postreward activation in ROIs that represent stimulus categories. This stimulus specificity was defined as the first criterion to make postreward activation a plausible correlate of credit assignment. The stimulus-specificity effect was assessed in two separate contrasts that modeled the parametric effect of reward size. The two parameters modeled BOLD activity increase in the ROIs separately for reward after face and house decisions, respectively. A repeated-measures ANOVA on the mean β values from the parametric contrasts with the factors ROI (PPA/FFA) and RESPONSE (house/face) yielded a significant main effect of ROI ($F_{(1,16)} = 5.104, p = 0.038$), no significant main effect of RESPONSE ($F_{(1,16)} = 1.59, p = 0.22$), and a statistically significant interaction ($F_{(1,16)} = 8.98, p = 0.009$), in line

with the hypothesis of stimulus-specific postreward ROI activity (Fig. 4). In a follow-up analysis, we investigated the degree to which both areas contributed to this overall effect. To this end, we assessed β weights within both ROIs separately for the following two parametric contrasts: (1) parametric increase in BOLD response with reward size after face decisions (FaceDecisionReward) and (2) parametric increase in BOLD response with reward size after house decisions (HouseDecisionReward). The FaceDecisionReward parameter yielded a significant result in the FFA ROI ($t_{(16)} = 2.45, p = 0.013$), but no significant result in the PPA ROI. No significant parametric BOLD increase could be established in either ROI for HouseDecisionReward. Thus, visually identical reward images (pictures of money bags) activated stimulus representations differentially in a decision-contingent manner, an effect mostly carried by an increase in FFA activity following reward stimuli to face decisions compared with reward stimuli following house decisions (Fig. 4).

Task-dependent activation at reward outcome

Our second criterion for postreward activation to be a marker of credit assignment was task dependency. Task dependency requires postreward activation to be specific to perceptual decision tasks. We modeled four separate parametric contrasts: reward size for face decisions, house decisions, instructed "face" responses, and instructed "house" responses. Entering the β values from the RESPONSE (face/house), ROI (FFA/PPA), and TASK (decision/instructed) conditions of the parametric analysis into a $2 \times 2 \times 2$ repeated-measures ANOVA revealed no significant main effects, but a marginally significant interaction for RE-

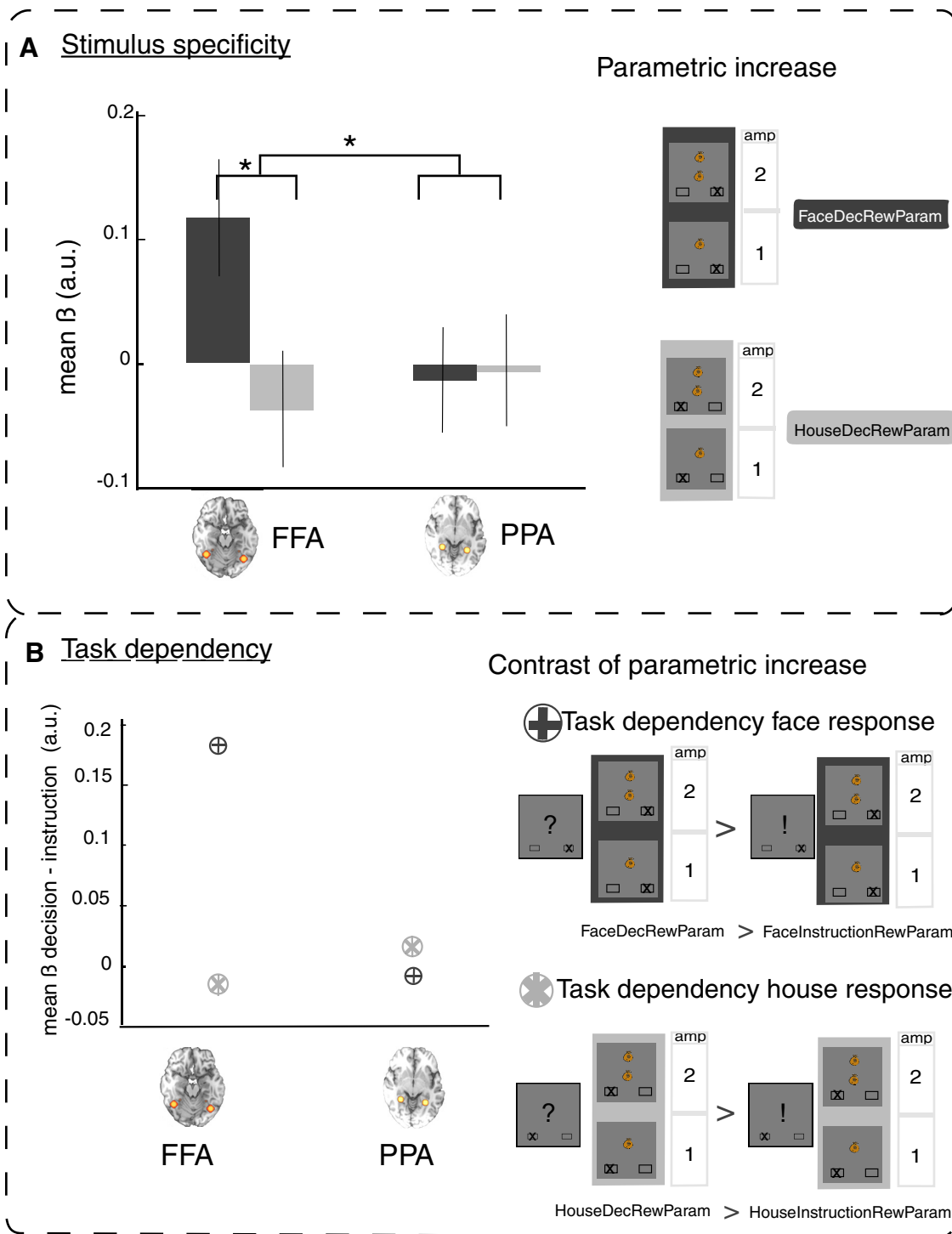


Figure 4. *A*, Stimulus-specific activity in noise trials was measured as an increase in BOLD activity with reward size, which was significantly stronger for the decision associated with an ROI (face for FFA, house for PPA) than for the opposite decision. Dark gray bars show mean β s for parametric increase with reward magnitude for face decisions; light gray bars show mean β s for parametric increase with reward magnitude for house decisions. Left, Activity in the FFA ROI. Right, Activity in the PPA ROI. *B*, Task dependency was reflected in a larger parameter estimate (mean β) for BOLD increase with reward size in the associated ROIs (FFA left, PPA right) in the perceptual decision, compared with the instructed response tasks. Markers refer to the difference of mean β s in the perceptual decision and instructed response tasks; the difference for face responses are shown in dark gray, house responses in light gray.

SPONSE \times TASK ($F_{(1,16)} = 3.125, p = 0.096$) and the predicted significant interaction of ROI \times TASK \times RESPONSE ($F_{(1,16)} = 10.84, p = 0.005$). This significant three-way interaction is indicative of a stronger positive relationship between reward size and response-specific ROI activity in the perceptual decision task than in the instructed response task (Fig. 4), satisfying the task-

dependency criterion. Because activity in the FFA ROI was modulated to a higher degree by stimulus specificity, we assessed within the FFA whether this stimulus-specific activity was also task dependent. Therefore, face decisions were contrasted with trials in which participants made an instructed response with the same key. This contrast (FaceDecisionReward > FaceInstruc-

tionReward) yielded a significant result in the FFA ROI ($t_{(16)} = 2.02$, $p = 0.03$), showing that the significant effect of stimulus specificity in the FFA was task dependent (Fig. 4*b*).

Reward-driven activation of stimulus-specific areas: PPI analyses

As a parallel to the preceding ROI analysis, which assessed whether stimulus-selective regions in sensory cortex would show activity at the time of reward presentation, our PPI analyses aimed to identify whether regions showing reward-related modulation would simultaneously exhibit evidence of stimulus specificity. Consistent with this notion, the face-vs-house PPI analysis yielded significant activation ($z = 2.3$, $p = 0.01$) in an area of the posterior fusiform gyrus centered on $x = 34$, $y = -78$, $z = -15$ (Talairach coordinates). As a complementary left-hemisphere ROI, we chose the closest peak of activation to the mirrored coordinates in the left hemisphere ($x = -34$), which identified a peak voxel at $x = -38$, $y = -78$, $z = -6$ that was marginally significant ($z = 1.63$; $p = 0.053$). A mask consisting of two cubes of $2 \times 2 \times 2$ voxels around these centers was then applied to the face block versus house block functional localizer and yielded a significant result ($z = 2.09$, $p = 0.018$). Thus, these bilateral ROIs in the fusiform gyrus (1) correlated more positively with the nucleus accumbens after reward delivery for face than house decisions on noise trials, and (2) were more involved in face than house processing in a separate task. These ROIs were somewhat posterior to the FFA ROIs discussed above. PPI activation peaks were evident close to these FFA ROIs in the right hemisphere ($x = 31$, $y = -60$, $z = -15$) and left hemisphere ($x = -35$, $y = -48$, $z = -12$), but these activations were only marginally reliable in the PPI analysis (right hemisphere: $z = 1.39$, $p = 0.082$; left hemisphere: $z = 1.506$, $p = 0.066$), while of course exhibiting evidence of stimulus specificity ($z = 1.66$, $p = 0.047$).

The corresponding house-vs-face PPI analysis yielded significant activation ($z = 2.14$, $p = 0.016$) in an area in the hippocampal gyrus centered on $x = 16$, $y = -48$, $z = 3$ (Talairach coordinates). As a complementary left-hemisphere ROI, we chose the closest peak of activation to the mirrored coordinates in the left hemisphere ($x = -16$), yielding a peak voxel at $x = -17$, $y = -45$, $z = -9$ that was significantly activated ($z = 1.72$; $p = 0.041$). A mask consisting of two cubes of $2 \times 2 \times 2$ voxels around these centers was then applied to the house block versus face block functional localizer and yielded a significant result ($z = 2.507$, $p = 0.006$). Thus, we identified two ROIs in the parahippocampal region which correlated more positively with the nucleus accumbens after reward delivery for house than face decisions on noise trials, and these ROIs were more involved in house than face processing in a separate (localizer) task (Fig. 5).

Trial-type sensitivity

The present paradigm aimed to establish stimulus specificity and task dependency by focusing on trials with noise stimuli, using stimuli with true (house or face) signal primarily to create a credible context for those critical noise trials within our perceptual judgment task. It is nevertheless instructive to analyze reward-induced activity following the signal trials for comparison with other recent studies of reward-related activation in sensory cortex (Pleger et al., 2008; Weil et al., 2010; FitzGerald et al., 2012). Signal trials differ from noise trials for several notable reasons. They differ, for example, because these perceptual decisions would involve more reliance on bottom-up features and because reward probability and neural adaptation effects are confounded with signal strength.

To compare the two types of trials, we modeled reward parameters for noise as well as signal trials in each task, implementing eight separate regressors for the factorial combination of TRIAL-TYPE (noise/signal), RESPONSE (house/face), and TASK (decision/instructed). As a first pass, we established whether the documented stimulus-specificity and task-dependency effects were again reliable in noise trials in a $2 \times 2 \times 2$ (ROI \times RESPONSE \times TASK) repeated-measures ANOVA. This yielded a marginally significant interaction of ROI and TASK ($F_{(1,16)} = 3.09$, $p = 0.098$) and a significant three-way interaction between ROI, RESPONSE, and TASK ($F_{(1,16)} = 10.19$, $p = 0.006$), with ROI activity being greater for the associated response in decision tasks. Thus, we find support for the result from the original stimulus-specificity and task-dependency analyses in this alternative GLM.

Given this confirmation, we next compared postreward BOLD responses for the two different trial types in a $2 \times 2 \times 2 \times 2$ repeated-measures ANOVA, including the factor ROI (FFA/PPA) crossed with RESPONSE, TASK, and TRIAL-TYPE. This analysis replicated the stimulus-specificity effect in a marginally significant interaction of ROI \times RESPONSE ($F_{(1,16)} = 3.47$, $p = 0.081$). Further, it showed a significant four-way interaction ($F_{(1,16)} = 18.11$, $p = 0.001$). The significant four-way interaction indicates trial-type sensitivity of the established effects, expressed as a difference between noise and signal trials with regard to task dependency. Because the differential activation of the ROIs was repeatedly shown to be decision dependent, and not pronounced for instruction trials, we focused our further comparison of noise and signal trials on decision trials only. We therefore investigated the effect of trial-type sensitivity for decision trials in a $2 \times 2 \times 2$ repeated-measures ANOVA with the factors ROI (FFA/PPA), RESPONSE (face/house), and TRIAL-TYPE (noise/signal) and found a significant interaction of ROI and RESPONSE ($F_{(1,16)} = 7.32$, $p = 0.016$), indicating stimulus specificity in decision tasks, a significant interaction of ROI and TRIAL-TYPE ($F_{(1,16)} = 10.13$, $p = 0.006$), and a significant three-way interaction between ROI, RESPONSE, and TRIAL-TYPE ($F_{(1,16)} = 7.45$, $p = 0.015$), with more stimulus-specific, reward-modulated ROI activity after perceptual decisions on noise trials than on signal trials. Thus, we find confirmation of the stimulus-specificity and task-dependency effect, but a difference between noise and signal trials.

Discussion

Humans learn how to acquire rewarding outcomes and the neural basis of reward processing has been studied extensively. However, we know very little about how reward-yielding tasks are represented in the brain. Machine learning describes how reward-driven learning depends on eligibility traces that signal which events are predictors of reward (Weil et al., 2010; cf. credit assignment, FitzGerald et al., 2012). It has been proposed that these learning processes are evident at the neural level as activation of the sensory cortices that represent components of the rewarded task. We investigated the neural correlates of postreward task representations in visual association cortices in a perceptual decision task. We defined two criteria for potential neural correlates of eligibility traces: postreward activation in a sensory association cortex should be stimulus specific (i.e., reflect the stimulus category of the rewarded response) and task dependent (i.e., should only occur if the stimulus was relevant to the task). We indeed found the representation of a stimulus to be activated postreward, especially if it was relevant for the correct response. This effect was established in a significant interaction of

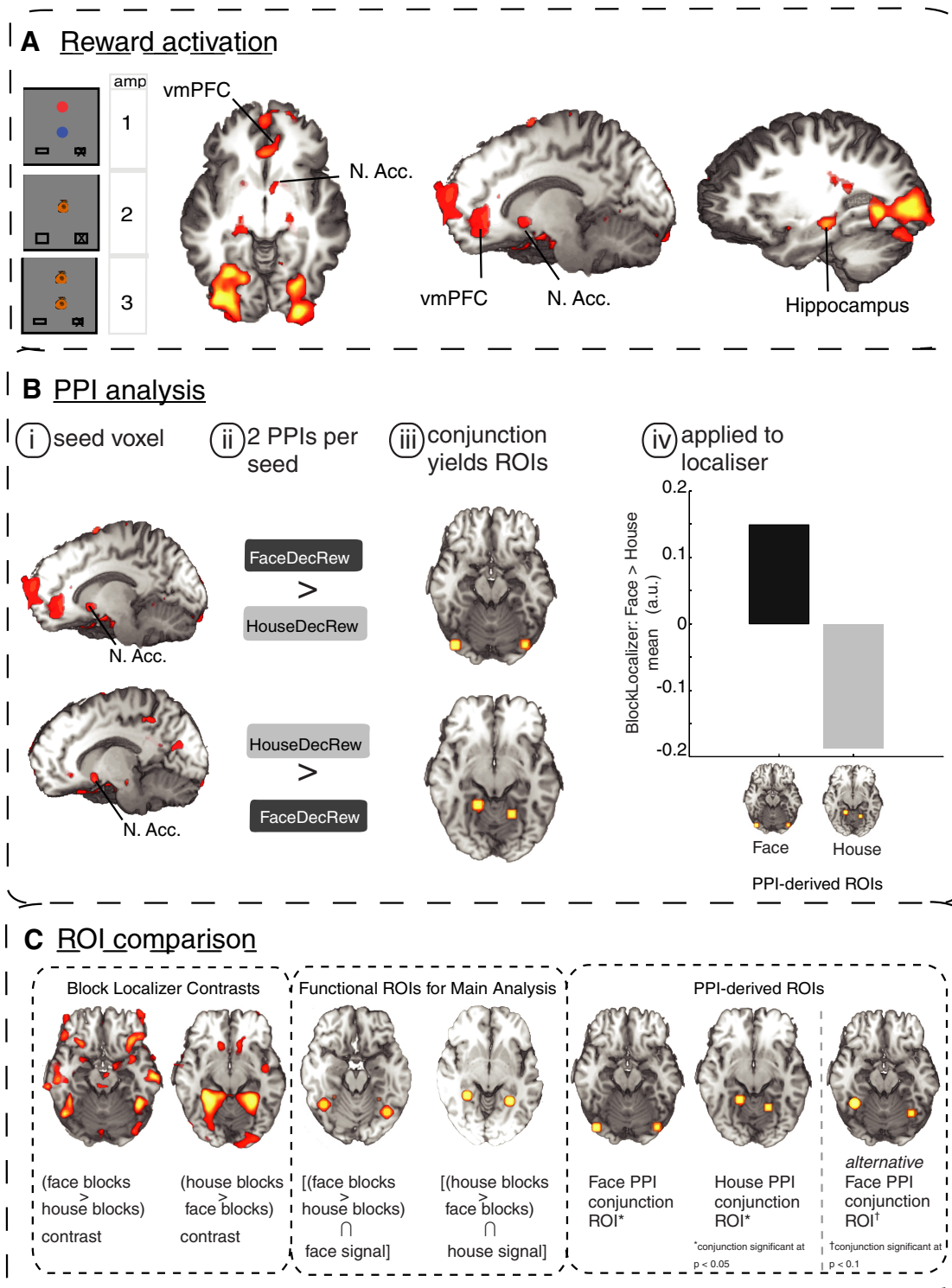


Figure 5. *A*, Rewarding outcomes after decisions on noise trials activated a network classically associated with reward delivery, including the ventromedial PFC (vmPFC) and nucleus accumbens (N. Acc.), as well as the anterior hippocampus. *B*, Each N. Acc. peak voxel (*i*) was used as a seed voxel in two PPI analyses (*ii*). The first PPI analysis implemented a contrast of the increase in BOLD response after face decisions to noise trials versus BOLD increase after house decisions to noise trials. The other PPI implemented the reverse contrast. A conjunction of the PPIs derived from different hemispheres for FaceDecisionReward > HouseDecisionReward contrasts yielded ROIs in the posterior fusiform gyrus. The conjunction of the bilateral PPIs implementing the HouseDecisionReward > FaceDecisionReward contrasts yielded ROIs in the parahippocampal region (*iii*). Applying these ROIs to the functional localizer block showed that areas that correlated with the N. Acc. in the FaceDecisionReward > HouseDecisionReward contrast were also selectively activated by blocks of face images (*iv*). Conversely, areas that correlated with the N. Acc. in the reverse contrast were more selectively activated by blocks of house images than face images. *C*, While the ROIs derived from the PPI approach (right) diverge from the ROIs that were selected on the criteria: responsiveness to a specific category in the functional localizer and correlation with signal strength (middle), both sets of areas lie clearly within the large parts of cortex that show either preferential response to face blocks or preferential response to house blocks, respectively. The most strongly activated face-selective area in the FaceDecisionReward > HouseDecisionReward PPI was located posterior to classical FFA coordinates. An area closer to these coordinates (denoted as alternative Face PPI conjunction ROI) showed marginally significant activity in the FaceDecisionReward > HouseDecisionReward PPI and also displayed face selectivity in the block localizer.

response category (face or house decision) and BOLD activity in the FFA or PPA. Moreover, the effect was specific to trials in which the stimulus was task relevant, fulfilling the criterion of task dependency. The effects of stimulus-specific postreward activation received further support from a PPI analysis. This analysis identified an area in the parahippocampal region that correlated with nucleus accumbens activity after rewarded house responses and that was sensitive to house processing in the separate localizer task, whereas an area in the fusiform gyrus that correlated with nucleus accumbens activity after rewarded face responses was sensitive to face processing in the separate localizer task.

Credit assignment

Reinforcement learning theory explains how reward-predicting events are assigned a higher value and become targets of behavior (Sutton, 1988; Sutton and Barto, 1990; Daw and Doya, 2006; Dayan and Niv, 2008). However, the neural underpinnings of this credit assignment mechanism are unclear. In particular, it is known that reward prediction and prediction errors elicit neural activity in the basal ganglia and vmPFC and result in increased firing of dopamine neurons in the midbrain, but it is yet to be established how this reward response fosters the representations of rewarded tasks. One proposal is that reward signals increase synaptic plasticity in sensory areas (Jay, 2003; Lisman et al., 2011; Pennartz et al., 2011). Supporting evidence comes from studies showing modulation of neural activity by anticipated reward (Serences, 2008; Brosch et al., 2011). These findings indicate that pairing with reward changes the neuronal representation of an event's value in sensory cortices; this may explain why associations between stimuli and reward can prime behavior (Hickey et al., 2010; Hickey and van Zoest, 2012; Wimmer and Shohamy, 2012). However, while these studies show that credit assignment takes place, and may be linked to dopamine release, they do not reveal how neural activity representing relevant stimuli is linked to neural correlates of reward during learning. The present study sheds light on this question by demonstrating stimulus specificity and task dependency of postreward activation.

Stimulus specificity

If postreward activation of the task underlies learning, an exact representation of the previous stimulus should be traceable after reward delivery. Although tested in a number of studies (Pleger et al., 2008, 2009; Weil et al., 2010; FitzGerald et al., 2012), results of paradigms studying reward-related activation in visual association areas have been ambiguous. Using fMRI in monkeys, Arsenault and colleagues (2013) successfully showed postreward activation in sensory cortices, but only in trials that did not entail the visual stimulus itself. Conversely, several human fMRI studies have found no evidence for stimulus-specific activity in sensory cortex following reward delivery (Weil et al., 2010; FitzGerald et al., 2012).

In the present study, we measured effects following decisions on pure noise stimuli. This may have rendered the design especially sensitive to stimulus-specific activation for a number of reasons. First, adaptation to individual stimuli has been suggested to explain the observation of nonspecific activation effects (FitzGerald et al., 2012). In contrast, we analyzed postreward activation following perceptual decisions on noise. Activity in the respective ROIs was thus dependent on the judged category representations, not on low-level features of the individual stimuli. Category representations might be less prone to sensory adaptation than low-level features. Second, activation in noise trials and

top-down-driven postreward activation rely both on feedback projections, which differ from feedforward projections conveying sensory input. They might therefore activate the same level in the cortical hierarchy of stimulus representation (Markov et al., 2013). This could increase the overlap of the locus of BOLD response measured for decisions under noise and top-down-driven activation, increasing positive correlation between decision-specific activity and postreward stimulus-specific activity. Third, in studies with true "signal" stimuli, strong anticipation of reward may modulate activity in the sensory cortices before reward delivery (Serences, 2008; Brosch et al., 2011). Thus, reward delivery may have had little effect on activity given that it was delivered in a performance-dependent manner in tasks where participants performed above chance (Pleger et al., 2008; Weil et al., 2010; FitzGerald et al., 2012). Here, however, rewards in noise trials could not be anticipated, as feedback was assigned randomly, limiting a positive correlation between signal strength, reward anticipation, and reward. Collectively, these features of noise stimuli in our design may have made our study very sensitive to stimulus-specific postreward effects.

The stimulus-specific effects we identified in terms of interactions between decisions and ROIs were reliable only within the FFA. A possible explanation for this finding is that we chose the ROIs based on a localizer that correlated BOLD increase with the increase in signal strength in the stimulus. We thus biased our analysis toward ROIs that responded more strongly if more category-specific bottom-up input was present. However, whether, for example, subregions of the PPA are differentially involved in top-down-driven versus bottom-up-driven scene processing has yet to be empirically tested (cf. Park et al., 2010).

Task dependency

Most objects require a specific manipulation to yield desired outcomes. However, at any given moment, many objects are present in our environment, and any given object may afford different actions, depending on the task. Hence, reward needs to activate the specific representations of only those objects that were involved in the current task. Global postreward activation, including irrelevant stimulus representations, would yield a new credit assignment problem (Roelfsema and van Ooyen, 2005). In fact, stimulus specificity and task dependency relate to the difference between Pavlovian and instrumental conditioning. Pavlovian conditioning on the one hand requires a stimulus-specific association between the conditional stimulus and the unconditional stimulus to be formed; this association needs to be stimulus specific, but is not necessarily related to a specific action. On the other hand, instrumental conditioning requires the association of a stimulus, an appropriate response, and the eventual reinforcer. Thus, instrumental conditioning should crucially incorporate task dependency.

This reasoning implies that if postreward activation observed in the current study was a marker of a credit assignment, we would expect it to reflect whether the preceding stimulus was relevant to the rewarded task. Previously, postreward activation has been tested using designs in which stimuli that preceded reward were always task relevant, rendering it difficult to interpret the established effects as correlates of either credit assignment or less specific reward-driven activation. In the present study, only one of two tasks required a response that was stimulus specific. We showed that stimulus-specific postreward activation was dependent on the relevance of the stimulus for the reward-yielding task. Thus, our study shows that reward selectively increases activity in sensory areas representing objects that have been used to

perform a task, and not globally in sensory areas representing any object present in the current context.

Conclusion

The present study has established stimulus-specific and task-dependent activation following reward delivery in a perceptual decision task. The established features of stimulus specificity and task dependency suggest that postreward activation may be a cortical signature of eligibility traces for credit assignment. This finding is a substantial step toward closing the gap between well defined computational concepts in reward-based learning and their neural implementation. The next important step will be to clarify how information is maintained during the intervals between the stimulus, the response, and the ensuing feedback. Representations could be maintained as persistent activation in sensory cortices, which would be modulated by the eventual feedback. Alternatively, increased activation of representations might only persist in PFC during the delay and be reinstated in sensory cortices by reward. Further, an important task for future research will be to test the suggested functional relationship between postreward activation and the dopaminergic modulation of synaptic plasticity directly. Establishing such a link would open the field to further exciting questions regarding the mechanisms of maintaining relevant representations until reward delivery, the nature of dopaminergic circuits that mediate this form of learning, and the degree to which this learning depends on the availability of dopamine, or the interplay between dopamine and other neurotransmitters that influence behavior.

References

- Arsenault JT, Nelissen K, Jarraya B, Vanduffel W (2013) Dopaminergic reward signals selectively decrease fMRI activity in primate visual cortex. *Neuron* 77:1174–1186. [CrossRef Medline](#)
- Brosch M, Selezneva E, Scheich H (2011) Representation of reward feedback in primate auditory cortex. *Front Syst Neurosci* 5:5. [CrossRef Medline](#)
- Daw ND, Doya K (2006) The computational neurobiology of learning and reward. *Curr Opin Neurobiol* 16:199–204. [CrossRef Medline](#)
- Dayan P, Niv Y (2008) Reinforcement learning: the good, the bad and the ugly. *Curr Opin Neurobiol* 18:185–196. [CrossRef Medline](#)
- FitzGerald TH, Friston KJ, Dolan RJ (2012) Action-specific value signals in reward-related regions of the human brain. *J Neurosci* 32:16417–16423a. [CrossRef Medline](#)
- Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci* 10:14–23. [CrossRef Medline](#)
- Hickey C, van Zoest W (2012) Reward creates oculomotor salience. *Curr Biol* 22:R219–R220. [CrossRef Medline](#)
- Hickey C, Chelazzi L, Theeuwes J (2010) Reward changes salience in human vision via the anterior cingulate. *J Neurosci* 30:11096–11103. [CrossRef Medline](#)
- Jay TM (2003) Dopamine: a potential substrate for synaptic plasticity and memory mechanisms. *Prog Neurobiol* 69:375–390. [CrossRef Medline](#)
- Lisman J, Grace AA, Duzel E (2011) A neoHebbian framework for episodic memory; role of dopamine-dependent late LTP. *Trends Neurosci* 34:536–547. [CrossRef Medline](#)
- Lohmann G, Müller K, Bosch V, Mentzel H, Hessler S, Chen L, Zysset S, von Cramon DY (2001) LIPSIA—a new software system for the evaluation of functional magnetic resonance images of the human brain. *Comput Med Imaging Graph* 25:449–457. [CrossRef Medline](#)
- Markov NT, Ercsey-Ravasz M, Van Essen DC, Knoblauch K, Toroczkai Z, Kennedy H (2013) Cortical high-density counterstream architectures. *Science* 342:1238406. [CrossRef Medline](#)
- O’Doherty JP (2004) Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol* 14:769–776. [CrossRef Medline](#)
- O’Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337. [CrossRef Medline](#)
- O’Reilly JX, Woolrich MW, Behrens TE, Smith SM, Johansen-Berg H (2012) Tools of the trade: psychophysiological interactions and functional connectivity. *Soc Cogn Affect Neurosci* 7:604–609. [CrossRef Medline](#)
- Park S, Chun MM, Johnson MK (2010) Refreshing and integrating visual scenes in scene-selective cortex. *J Cogn Neurosci* 22:2813–2822. [CrossRef Medline](#)
- Pennartz CM, Ito R, Verschure PF, Battaglia FP, Robbins TW (2011) The hippocampalstriatal axis in learning, prediction and goal-directed behavior. *Trends Neurosci* 34:548–559. [CrossRef Medline](#)
- Pleger B, Blankenburg F, Ruff CC, Driver J, Dolan RJ (2008) Reward facilitates tactile judgments and modulates hemodynamic responses in human primary somatosensory cortex. *J Neurosci* 28:8161–8168. [CrossRef Medline](#)
- Pleger B, Ruff CC, Blankenburg F, Klöppel S, Driver J, Dolan RJ (2009) Influence of dopaminergically mediated reward on somatosensory decision-making. *PLoS Biol* 7:e1000164. [CrossRef Medline](#)
- Rajimehr R, Devaney KJ, Bilenko NY, Young JC, Tootell RB (2011) The parahippocampal place area responds preferentially to high spatial frequencies in humans and monkeys. *PLoS Biol* 9:e1000608. [Medline](#)
- Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: *Classical conditioning II: Current research and theory* (Black AH, Prokasy WF, eds), pp 64–69. New York: Appleton Century Crofts.
- Roelofsma PR, van Ooyen A (2005) Attention-gated reinforcement learning of internal representations for classification. *Neural Comput* 17:2176–2214. [CrossRef Medline](#)
- Schultz W (2007) Behavioral dopamine signals. *Trends Neurosci* 30:203–210. [CrossRef Medline](#)
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599. [CrossRef Medline](#)
- Schyns PG, Oliva A (1994) From blobs to boundary edges: Evidence for time- and spatial-scale-dependent scene recognition. *Psychol Sci* 5:195–200. [CrossRef](#)
- Serences JT (2008) Value-based modulations in human visual cortex. *Neuron* 60:1169–1181. [CrossRef Medline](#)
- Sutton RS (1988) Learning to predict by the methods of temporal differences. *Machine Learning* 3:9–44. [CrossRef](#)
- Sutton RS, Barto AC (1990) Time-derivative models of Pavlovian reinforcement. In: *Learning and Computational Neuroscience: Foundations of Adaptive Networks* (Gabriel M, Moore J, eds), pp 497–537. Cambridge, MA: MIT.
- Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: an approach to cerebral imaging. New York: Thieme Medical.
- Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412:43–48. [CrossRef Medline](#)
- Weil RS, Furl N, Ruff CC, Symmonds M, Flandin G, Dolan RJ, Driver J, Rees G (2010) Rewarding feedback after correct visual discriminations has both general and specific influences on visual cortex. *J Neurophysiol* 104:1746–1757. [CrossRef Medline](#)
- Willenbockel V, Sadr J, Fiset D, Horne GO, Gosselin F, Tanaka JW (2010) Controlling low-level image properties: the SHINE toolbox. *Behav Res Methods* 42:671–684. [CrossRef Medline](#)
- Wimmer GE, Shohamy D (2012) Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science* 338:270–273. [CrossRef Medline](#)
- Worsley KJ, Friston KJ (1995) Analysis of fMRI time series revisited—again. *Neuroimage* 2:173–181. [CrossRef Medline](#)