Neural and cognitive mechanisms underlying adaptation
van den Berg, Berry

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
van den Berg, B. (2018). Neural and cognitive mechanisms underlying adaptation: Brain mechanisms that change the priority of future information based on their behavioral relevance. [Groningen]: University of Groningen.

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
A key role for stimulus-specific updating of sensory cortices in the learning of stimulus-reward associations

Berry van den Berg
Benjamin R. Geib
Rene San Martin
Marty G. Woldorff

Submitted - Social Cognitive Affective Neuroscience
Abstract

Successful adaptive behavior requires the learning of associations between stimulus-specific choices and rewarding outcomes. Most research on the mechanisms underlying such processes has focused on subcortical reward-processing regions, in conjunction with frontal circuits. Given the extensive stimulus-specific coding in the sensory cortices, we hypothesized they would play a key role in the learning of stimulus-specific reward associations. We recorded electrical brain activity (EEG) during a learning-based, decision-making, gambling task where, on each trial, participants chose between a face and a house and then received feedback (gain or loss). Within each 20-trial set, either faces or houses were more likely to predict a gain. Results showed that early feedback processing (200-1400ms) was independent of the choice made. In contrast, later (1400-1800ms) that processing became stimulus-specific, reflected by decreased alpha power (less cortical activity) over face-selective regions (reflecting increased cortical activity) for winning-versus-losing after a face choice, but not after a house choice. Finally, as the reward association was learned in a set, there was increasingly stronger attentional bias towards the more likely winning stimulus, reflected by increasing attentional-orienting-related brain activity and increasing likelihood of choosing that stimulus. These results delineate the processes underlying the updating of stimulus-reward associations during feedback-guided learning, which then guides future attentional allocation and decision making.

Keywords: EEG; oscillatory alpha; reward; learning; decision making

Introduction

To achieve successful adaptive behavior, humans and other animals need to learn to associate specific stimuli and choices with potential rewards, a process that requires the continuous monitoring and incorporation of feedback information. Learning such associations could facilitate predicting whether one will like a new food product based on previous experience with similar foods, or whether walking or taking public transportation is a quicker option to get home from work at certain times of day. Choosing adaptively in probabilistic settings requires an organism to learn, store, and continuously update stimulus-reward associations to be available to guide future behavior (e.g., Anderson, 2017).

Previous research investigating the neural processes underlying probabilistic learning has shown an important role for subcortical reward regions, such as the ventral tegmental area (VTA)/sustantia nigra (W. Schultz, 2016; W. Schultz, Dayan, & Montague, 1997) and nucleus accumbens (nAcc) (Alexander & Crutcher, 1990; Floresco, 2015), in conjunction with frontal regions, such as the anterior cingulate cortex (ACC) (Botvinick, 2007; Bush et al., 2002; Klein-Flugge, Kennerley, Friston, & Bestmann, 2016) and orbital frontal cortex (FitzGerald, Seymour, & Dolan, 2015; Leong, Radulescu, Daniel, DeWoskin, & niv, 2017; Sul, Kim, Huh, Lee, & Jung, 2010; Wallis, 2011), in the processing of value and feedback-related information. Activation in these brain regions has been found to vary according to the difference in value between the actual and the expected outcome (i.e., reward-prediction error), as shown both in humans with functional MRI and in nonhuman primates with single-unit recording (Yael niv, 2009; W. Schultz et al., 1997).

In addition, in humans, brain activity reflecting the evaluation of feedback information, has been tracked noninvasively at the scalp with high temporal resolution by using electroencephalogram (EEG) recordings and analyzing the associated event-related potentials (ERPs). This research has consistently found the feedback-related negativity (FRN) response (Gehring & Willoughby, 2002; Mittner, Braun, & Coles, 1997), a fronto-central negative-polarity deflection that peaks in the scalp-recorded ERP ~250ms following feedback and is thought to arise from the ACC, is larger for negative compared to positive outcomes. Following the FRN, outcome feedback also elicits a positive-polarity P3 deflection (starting at ~300ms) that tends to be larger in response to suboptimal outcomes and has been postulated to reflect the general enhancement of cognitive resources in response to a relative loss (Nieuwenhuis, Aston-Jones, & Cohen, 2005; O’Connell, Dockree, & Kelly, 2012; Polich, 2007; R San Martín, Appelbaum, Pearson, Huettel, & Woldorff, 2013; René San Martín, 2012). In addition, the P3 is observed to be larger when stimulus values have to be updated, and its amplitude has been found to predict future choice adjustments (Fischer & Ullsperger, 2013; R San Martin et al., 2013).
The subcortical and frontal regions mentioned above, however, seem to typically respond to loss-versus-gain feedback, mostly independently of the specific stimulus with which the feedback was associated. Accordingly, our understanding of the neural mechanisms through which feedback information guides the establishment and updating of stimulus-reward associations remains incomplete, particularly with regard to the possible role of the sensory cortices during the learning and updating of stimulus-choice contingencies. More specifically, we hypothesized that feedback-guided stimulus-specific updating processes would modulate the activity in areas in the sensory cortices specifically involved in the processing of the stimuli whose reward associations are currently being learned.

Relatedly, recent studies have reported that attention can be influenced by specific stimuli that have been previously associated with reward. Specifically, stimulus-reward associations have been shown to bias attention to be rapidly oriented towards those stimuli when they appear in a visual scene (Hickey et al., 2010a; Krebs, Boehler, Roberts, et al., 2011; San Martin, Appelbaum, Huettel, & Woldorff, 2016). The general idea is that by biasing attention, stimulus-reward associations can potentially improve environmental responsivity, decision-making, and other behavior (San Martin et al., 2016); on the other hand, such associations can also misguide attention when they are associated with a task-irrelevant distractor stimulus or feature (Hickey et al., 2010a; Krebs et al., 2013, 2010; Krebs, Boehler, Egner, et al., 2011). In the spatial attention domain, such biases in attentional orienting can be measured by examining the hallmark ERP component known as the N2pc, a lateralized negative-polarity deflection peaking at ~250ms over posterior cortex contralateral to the direction of an attentional shift (Hickey et al., 2010a; Kappenman & Luck, 2012). However, the stimulus-specific neural updating processes involved in feedback-guided learning that create these attentional biases remains poorly understood.

One possibility for such feedback-based updating would be that neurons in sensory cortices that are involved in stimulus-specific processing would be involved in such learning (Folstein, Palmeri, & Gauthier, 2013). A classic example of stimulus-specific coding is the set of cortical regions that respond selectively to images of faces, which have been delineated by fMRI and electrophysiological measures (Handwerker et al., 2006; McCarthy, Puce, Gore, & Allison, 1997; Perrett, Hietanen, Oram, Benson, & Rolls, 1992; Puce, Allison, & McCarthy, 1999). These face-specific regions include the occipital face area (OFA), the fusiform face area (FFA), and the superior temporal sulcus (STS) (Gobbini & Haxby, 2007; Pitcher, Walsh, & Duchaine, 2011). In addition, in humans the processing of faces elicit selective scalp-recorded ERPs, with the largest being the N170 (latency 170ms), a lateral-inferior occipital negative-polarity deflection that is greater for faces compared to other objects, generally thought to reflect increased activation from the lateral face-selective cortical processing regions (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Rossion & Jacques, 2012). Face-selective activity is typically extracted by comparing responses to face images to responses to images of other objects, most commonly images of houses and other buildings. fMRI studies have shown that house and building images elicit selective activity in the parahippocampal place area (PPA) (Epstein & Kanwisher, 1998), a medial-inferior temporal brain region, but such stimuli do not seem to produce a very distinctive marker in the ERPs.

Accordingly, here we focused on stimulus-specific cortical activity for faces (extracted by face-image versus building-image response contrasts) and its distinctive scalp ERP/EEG markers. The stimulus and cortical specificity of scalp-recorded electrophysiological measures, such as described above for the face-selective N170, was for the ERPs that are extracted by time-locked averaging the EEG. Also extractable from the EEG, however, are the time-locked power changes in oscillatory EEG activity, which can provide an additional useful window into cortical brain processing. In particular, changes in oscillatory EEG activity in the alpha band (8 - 14 Hz) have been found to index the directionality of cortical activation, in that decreased alpha power arising from a cortical brain region is typically associated with increased cortical activation, and vice versa (Jensen & Mazaheri, 2010; Peterson & Kleinschmidt, 2012; Scheeringa et al., 2012; van den Berg, Appelbaum, Clark, Lorist, & Woldorff, 2016b; van den Berg et al., 2014; van Dijk, Schoffelen, Oostenveld, Jensen, et al., 2008a). A classic example of this effect in spatial attention is the relative decrease in occipital alpha contralateral to vs. ipsilateral to a cued direction of attention (Foxe & Snyder, 2011; Grent-’t-Jong et al., 2011; Worden, Foxe, Wang, & Simpson, 2000b), inversely paralleling relative lateralized increase in fMRI activity observed under a similar contrast (Green et al., 2017; Grent-’t-Jong & Woldorff, 2007; Hopfinger et al., 2000; Kastner et al., 1999).

Here, to investigate the feedback-guided stimulus-specific updating processes, we leveraged the high temporal resolution of EEG, using both ERPs and oscillatory EEG power changes, to track the spatiotemporal cascade of activity over face-processing-selective cortex while participants performed a probabilistic decision-making gambling task. On each trial (Figure 1), participants were asked to choose between a face and a house, after which they were given feedback indicating that they would receive either a monetary gain or loss on that trial. Within each 20-trial set, either faces or houses were more likely to lead to a gain (probability bias in each set was randomly chosen between 0.50 and 0.75), with the participants instructed to try to learn the likelihood in that set and thereby improve their reward-gaining performance. To identify spatially discernable signals related to face processing, we used a separate localizer task from which we delineated scalp regions that reflect differential processing for faces versus houses. Additionally, we analyzed the power changes in oscillatory EEG activity in the alpha band as an inverse index of cortical activations. In particular, modulations in alpha
activity were used as a marker for face-selective cortical activation to index the trial-to-trial updating of stimulus-specific rewards associations during learning.

Figure 1. Probabilistic gambling task: each trial started with a dual-object cue-choice stimulus, with a face on one side and a house on the other. Following the cue, participants made a choice of one of these images by selecting the appropriate arrow. Feedback with respect to loss versus gain for that choice was then presented.

Methods

Participants
Thirty healthy volunteers (mean age [sd]: 23.5 [2.9], 18 female, 29 right-handed) participated in the study. All had normal or correct-to-normal vision and all gave written informed consent. The study was conducted in accordance with protocols approved by the Duke Institutional Review Board. Participants were paid 15 dollars/hour, plus an additional reward-related bonus (mean [sd]: 9.6 USD [7.8]). One participant was excluded due to poor EEG data quality (>40% of the trials rejected due to artifacts), leaving 29 subjects in the final analysis.

Tasks and Stimuli
Participants were asked to do a learning-based decision-making gambling task built upon one we had employed previously (San Martin et al., 2016; R San Martin et al., 2013). Stimuli were presented on a 24-inch LCD monitor using the Presentation software package (Neurobehavioral Systems, Inc., Albany, CA). Participants were seated in a comfortable chair with eyes about ~60 cm from the screen.

Learning-based gambling task
The probability-learning (Plearn) gambling task (Figure 1) consisted of 24 sets of 20 trials each. Within each set either the faces or houses were more likely to win, with an average probability bias of 0.625 (ranging randomly from 0.50-0.75). As a result, the final probability bias distribution across all sets was roughly Gaussian (mean [sd]: 0.625 [0.075]). Each trial started with the presentation of a image-pair cue (duration 1200ms), consisting of the image of a face (randomly selected from 20 male faces) and the image of a house (randomly selected from 20 houses), both with a resolution of 72 by 96 pixels. On each trial, the house and face were randomly assigned to appear on either the left or right side of a central fixation. After some delay (1200 ms), a screen for choosing between the two cue stimuli (the choice screen) was then presented, which consisted of two arrows randomly pointing to the left and right, just above and below the fixation, remaining on the screen until a behavioral response was given (up to a maximum of 1200ms). Participants were instructed to choose the stimulus that they judged would lead to a gain by selecting the arrow that pointed in that direction. Participants pressed the top or bottom bumper of a gamepad (Logitech Rumblepad) corresponding to the top or bottom arrow. If no response was given, participants received feedback displaying “no response” on the screen, followed by a loss in points. If a response was made, the chosen arrow was highlighted. Then, following a jittered interval sampled from a uniform distribution between 700 and 900ms the gain/loss feedback appeared onscreen (duration 500ms), which was either a “+8” or a “-8” printed in a blue or orange square (counterbalanced across participants) indicating the valence of the feedback. After the feedback there was another jittered interval between 1500 and 2000ms (i.e., 2000 - 2500ms following the onset of the feedback stimulus). We intentionally chose a long interval after feedback to be able to track the stimulus-reward updating neural signals following the processing of the feedback stimulus.

Participants were instructed to try to learn on each set of 20 trials whether faces or houses were more likely to result in a reward in that set and were informed that they could win points, which would translate to real money after the task, by choosing the stimulus (i.e., face or house) that was most likely to win on each set. Each set was followed by feedback as to how many points they had accrued thus far. After a practice period in which the subject demonstrated understanding of how to do the task, the session started with the learning-based decision-making task runs. After completing these, the participants performed the face-versus-house localizer task (see below).

Localizer task
To extract potential regions of interest (ROIs) for assessing the updating in the Plearn paradigm of the face-responsive sensory cortices, we used a face/house localizer task. This task consisted of presenting 480 faces and houses to the participants, which were sampled randomly from the same set of 20 faces and 20 houses that were used in the learning-based decision-making task. Twenty percent of these stimuli were presented
as a blurred version, which served as infrequent targets, while the rest were clear. Blurry faces and houses were created by means of convolving a symmetric two-dimensional Gaussian kernel (with a 10-pixel width and height) with the face and house images. The participants’ task was to detect the blurry images and indicate whether they are a face or a house by pressing a button on the response-pad (left button for a blurry face and right button for a blurry house). Only the 80% of trials in which a clear face or house was presented (i.e., the nontargets) were used for the EEG analysis. Each face or house was presented for 300ms, and the next trial started after a jittered interval of 600-800ms. Twenty-eight participants of the original 30 completed the localizer task. The data from three participants in the localizer task (out of 28) were discarded due to excessive noise (>40 percent of the trials rejected), leaving 25 participants whose data in this task were used to delineate the face-selective ROIs.

EEG recording and preprocessing
EEG was DC-recorded (500 Hz sampling rate, with a three-stage cascaded integrator comb low-pass filter [CIC-filter] with a corner frequency of 130 Hz) using a 64-channel, custom-layout, equidistant, extended-coverage electrode cap (Woldorff et al., 2002) and a Brainproducts Actichamp amplifier with active electrodes (Acticap). The data was recorded referenced to the average of all the electrodes. Data was filtered offline using a 0.05 highpass causal FIR filter. Independent component analysis (ICA) was used to correct for eyeblinks, which were extracted using the extended infomax algorithm as implemented in the EEGLab software package (using EEGLab13.4.4b [Delorme & Makeig, 2004]). Independent components (ICs) that reflected eyeblinks (1 or 2 ICs per participant) were removed from the data.

For the Plearn task, epochs were extracted from -1000ms before until 2000ms after the cue-pair onset and from -1000ms before until 3500ms after the onset of the feedback stimulus. For the localizer task, epochs were extracted from -500ms before until 1000ms after the onset of the face or house image. Data was baseline-d from -200ms to stimulus onset. Epochs containing any remaining artifacts (e.g., horizontal eye movements, muscle noise) were detected using an average threshold (±110μV from -500 to 1000ms or to 2500ms respectively for the localizer and Plearn task) a 30μV step function from -200 to 1000ms (to 500ms for the localizer) around the target (for horizontal eye movements) and excluded from further analysis. Finally, an additional low-pass filter (20Hz, butterworth 4th order filter) was applied to the epoched and averaged ERP Plearn and localizer data.

Time-frequency analysis
Frequency decomposition was performed on the EEG average-referenced data using multitaper methods as implemented in the analysis software package Fieldtrip (Oostenveld et al., 2011), in which discrete prolate slepian sequences were used to estimate the power in logarithmically spaced frequencies from 3 to 80Hz. The window widths for the tapers were 3 cycles for 3Hz, 4 cycles for 4-7Hz, 5 cycles for 8-14Hz, 7 cycles for 15-20Hz, 10 cycles for 21-50Hz, and 15 cycles for 51-80Hz. Smoothing by means of multitapers was specified as 5×log10 of each frequency. Log10 transformed power spectra for each subject were subsequently binned and averaged according to the various conditions (see below). No baseline correction was performed for the oscillatory power analyses.

Data binning and averaging
The epoched EEG data and time-frequency power spectra were binned according to feedback (gains and losses) and choice (face and house), which resulted in an average number of trials for each subject (mean[std]) for face gain (111[22]), face loss (96[20]), house gain (108[23]), and house loss (92[19]), after rejection of noisy epochs. As expected, due to learning there were significantly more gain trials than loss trials (F(1,28) = 51.1, p=0.0001), especially towards the latter parts of each set. On the other hand, there was no significant difference in the number of gain or loss trials for choosing a face versus a house (F(1,28) = 0.39, p = n.s).

Analysis of cue-evoked ERP data
In order to calculate the cue-evoked N2pc response for indexing the attentional bias as it was modulated by reward learning across a set, the following procedure was implemented. To calculate the attentional bias as a function of trial number, we extracted cue-evoked activity averaged in bins of two consecutive trials (i.e., trials 1 and 2 were binned together, trials 3 and 4 were binned together, etc.), separately for each side (left, right) and for each experimentally fixed set-winner (e.g., if a face would have a 0.60 probability of winning in a 20-trial set, that would be a face-winner set). The cue-related attentional bias was assayed by standard N2pc contralateral-versus-ipsilateral analysis (Luck & Kappenman, 2012), that is, by subtracting the activity in the contralateral channels (relative to the set-winner) minus the ipsilateral channels, and then collapsing over the left and right sides.

Statistical analysis
For statistical analysis, we first calculated the time-locked ERPs and time-frequency responses as a function of the various event types and conditions. Next, the mean ERP component or Alpha-power (8-14Hz) amplitudes were extracted from specific scalp regions of interest (ROIs) and time-windows of interest (TOIs) based on the literature
and the localizer task. The FRN and P3 were measured from 200-300ms in a fronto-central ROI (analogue to Cz and FCz) and from 400-600ms in a parietal ROI (around Pz and neighboring channels), respectively (Fischer & Ullsperger, 2013; R San Martín et al., 2013; René San Martin, 2012). Subsequently, the extracted values were analyzed with a repeated-measures ANOVA to test for statistical significance (p < 0.05) using the R statistical software package (R Core Team, 2015b).

For potential alpha effects following feedback, reflecting a possible stimulus-reward cortical-updating process, we chose a period after the feedback-evoked P3 and before the start of the next trial (600-2000ms following feedback; recall that the next trial started 2000-2500ms following the onset of the feedback). The specific scalp ROIs for the potential alpha cortical-updating effect were based on the data extracted from the separate localizer task (averaged across participants). Statistical tests of alpha power over time was done using a cluster-based permutation testing approach (Maris & Oostenveld, 2007). T-tests were performed on the average power in the alpha band (8 to 14Hz) and on each time point within the 600 to 2000ms time window of interest separately for each ROI (i.e. the interval following the P3 activity and preceding the next trial (R San Martín et al., 2013)); if the resulting statistic exceeded a p-value of 0.05, then that time point was included into a cluster that was formed by including significant adjacent points. Cluster statistics were obtained by summing all t-values within a cluster. Statistical significance of a cluster was obtained by comparing the cluster statistic to a permutation distribution (created from 1000 iterations by randomly switching the labels of conditions at subject level), and was considered significant at p < 0.05 and reported at a trend level if p < 0.1.

The cue-related attentional bias (N2pc) was measured from 200-300ms after onset of the cue image pair, from corresponding left and right occipital ROIs (around O1/O2, PO5/PO6, PO3/PO4) (Hickey et al., 2010a; Kappenman & Luck, 2012; San Martin et al., 2016). Subsequently we calculated difference in voltage contralateral versus ipsilateral to the side on which the set-winner was presented.

To calculate the behavioral learning rate and neural attentional bias, we employed, on a trial-by-trial number basis, a mixed-modelling approach using the lmtest statistical package (Bates, Maechler, Bolker, & Walker, 2015). A varying slope of condition per subject (the random effect) was included in the model if the Akaike Information Criterion (a measure of the quality of the model) improved. To obtain statistical significance the degrees of freedom were approximated using the Satterwaithe approximation of degrees of freedom as given by the R package lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017). Statistical tests were considered significant at p<0.05.

1) \[ \text{proportion choosing set-winner} = \beta_0 + \beta_1 \cdot \text{trial}_n + \beta_2 \cdot \text{trial}_n^2 + \beta_3 \cdot \text{chosen} + \epsilon_n \]

2) \[ \text{attentional bias}_n = \beta_0 + \beta_1 \cdot \text{trial}_n + \epsilon_n \]

In formula (1), the dependent variable “proportion” reflects the proportion of choosing the winner on each trial\(\#\) (1-20), and the predictor variables consisted of an intercept, the first and second degree polynomials for trial\(\#\) (1-20 and 1-20 squared) and chosen stimulus (face versus house). The polynomial term was included in the model to account for non-linearities in the learning rate, notably that subjects’ performance gradually increases early on in the block before tending to asymptote when the most likely set-winner had been learned. In formula (2), the attentional bias (in µV) reflected by the N2pc was estimated by trial number (binned according to the average across two consecutive trials to achieve better signal to noise). For both the behavioral learning rate and neural attentional bias, the final regression model allowed for a varying intercept and a varying slope of trial number for each participant (as indicated by \(j\)).

Results

Localizer task: three regions of interest

The first difference between face and house processing in evoked activation patterns in the localizer task was a midline occipital positive deflection that was larger for faces compared to houses and peaked around 110ms (100-120ms: t(24) = 6.02 - Figure 2). This positive deflection has been observed in previous studies and tends to be bigger for upright and inverted faces compared to other object types (Eimer, 1998; Itti & Taylor, 2004). Following this midline occipital positivity, faces versus houses elicited the large, hallmark, face-selective N170 response (Bentin et al., 1996), which started at ~140ms as a differential, slightly right lateralized, lateral-inferior occipital negativity that was greater for faces compared to houses, peaked at around 160ms, and lasted until around 200ms (150 - 170ms: t(24) = -12.6). Subsequently we observed a second, more superior, bilateral, occipital, negative-polarity activation that was greater for faces compared to houses (200-240ms: t(24) = -3.6). Lastly, after 300ms there was a temporally extended activation in both the inferior and superior occipital ROIs that showed a negative-polarity difference between faces and houses (300-500ms: t(24) = -5.32). These patterns of activations from the localizer task resulted in three face-versus-house-selective ROIs that could be used to assess activity patterns that might reflect the updating and storing of face-specific stimulus-reward associations in the probability learning task: one occipital midline, one inferior occipital bilateral, and one
more superior occipital bilateral. These ROIs could potentially reflect activation from the OFA (midline occipital), FFA (inferior-lateral occipital), and the STS (superior occipital) (Pitcher et al., 2011), respectively. Importantly, the localizer results confirmed that our stimuli evoked classic patterns of face-selective activity as indicated by the occipital midline and inferior occipital ROI (midline positive deflection, and the N170).

Figure 2. Face-versus-house activity from the localizer task: differential activity for faces versus houses included the hallmark N170 negative-polarity response over ventrolateral visual cortex (for face responses minus house responses), which was preceded by a midline occipital modulation, and followed by a more superior occipital activation.

Behavioral results and the development of attentional biases toward the cue-choice stimulus with higher reward likelihood

The behavioral results from the Plearn task (Figure 3A) showed that across each set of 20 trials participants were able to learn whether the face or the house was the more likely winner for that set (main effect trial#: F(2,56) = 80, p<0.0001) which was not observably different for faces and houses (main effect chose: F(1,1094) = 0.1, p = n.s., chose × trial#: F(1,1094) = 0.5, p = n.s.). At the beginning of each 20-trial set, the participants chose the most likely winner for that set at chance (0.50), as would be expected. But by the end of the set, this proportion increased to ~0.75, indicating that the participants had utilized the feedback across the set to learn to choose the most likely probability-based winner. In other words, participants increased the probability of receiving gains by choosing faces more often when the set-winner was a face, and by choosing houses more often when the set-winner was a house.

Table 1. Model estimates for analysis of the behavioral learning rate: proportion choosing set-winner = β0 + β1 trial# + β2 trial#2 + β3 chose + β4 trial# × chose + ε. The levels in each predictor was: trial# 1:20 and chose; face and house.

<table>
<thead>
<tr>
<th>fixed effect</th>
<th>estimate (SE)</th>
<th>t(df), p</th>
</tr>
</thead>
<tbody>
<tr>
<td>β0</td>
<td>0.48 (0.02)</td>
<td>23.3 (107) &lt; 0.001</td>
</tr>
<tr>
<td>β1</td>
<td>0.14 (0.004)</td>
<td>9.5 (817) &lt; 0.001</td>
</tr>
<tr>
<td>β2</td>
<td>-0.001 (0.0002)</td>
<td>-6.6 (1094) &lt; 0.001</td>
</tr>
<tr>
<td>β3</td>
<td>0.02 (0.02)</td>
<td>0.9 (1094) = n.s.</td>
</tr>
<tr>
<td>β4</td>
<td>-0.005 (0.005)</td>
<td>-1 (1094) = n.s.</td>
</tr>
</tbody>
</table>

Neurally, at the moment of the presentation of the cue stimuli, this learning process was reflected by attention-sensitive lateralized N2pc activity evoked by the cue (Figure 3B). More specifically, in the first half of each 20-trial set the N2pc showed little or no reflection of attentional orienting toward the probabilistic set winner stimulus type in the cue-pair presentation, but in the second half of the set there was clearly an attentional bias, as reflected by a modulation of the lateralized orienting-related N2pc contralateral to the set winner. That is, consistent with previous studies showing with these neural measures the attentional orienting toward reward-associated stimuli (Hickey et al., 2010a; San Martin et al., 2016), during the learning of the reward-association here a negative deflection (200-300ms) that increased in size over the course of each set was elicited contralateral versus ipsilateral to the object in the cue-pair stimulus that predicted a higher probability of reward (F(1,106) = 7.1, p = 0.009).
Feedback processing

ERP results: evaluation of losses versus gains

ERP analysis of feedback processing focused on the feedback-evoked potentials, and more precisely the feedback-related negativity (FRN) and the P3. Based on previous literature, we anticipated processing of losses versus gains would be reflected by a larger FRN, with the later P3 wave also being larger following losses. The feedback-elicited ERPs confirmed these hypotheses (Figure 4A and B). Following the feedback stimulus, the first difference between losses and gains manifested itself as the hallmark, fronto-central, negative-polarity FRN deflection from around 200 to 300ms and peaking around 250ms (F(1,28) = 30.1, p < 0.0001). Additionally, there was no observable difference in the amplitude of the FRN as a function of whether the participant had chosen a face as compared to a house on a trial, confirming that the FRN was sensitive to a loss versus a gain and not to the particular stimulus type that had been chosen (feedback × choice: F(1,28) = 0.3, p = n.s.). Similarly, the P3 was larger for losses compared to gains, as has previously been reported (R San Martín et al., 2013) (F(1,28) =22.5, p < 0.0001). In contrast to the FRN, the loss-vs-gain effect for the feedback-elicited P3 did show some difference as a function of the stimulus type that had been chosen, being somewhat larger when the choice had been a house than when it had been a face (feedback × choice: F(1,28) = 7.7, p = 0.01). Moreover, specific contrasts with respect to this interaction indicated that the P3 modulation effect following losses was larger when participants had chosen...
a house vs. a face \((t(28) = 3.0, p = 0.006; 4.50 \mu V\) when choosing a house as compared to \(4.02 \mu V\) when choosing a face), with no difference for these choices following a gain \((t(28) = -0.1, p = n.s.; \sim 3.22 \mu V\) for both).

Oscillatory EEG results: updating of stimulus-reward associations

We had hypothesized that, following the generic (i.e., non-stimulus-specific) feedback processing (i.e., reflected by the FRn and, to a lesser extent, the P3), a more stimulus-specific cortical signal, as measured by means of oscillatory alpha power, would reflect processes related to the updating of specific stimulus-reward associations in sensory regions in posterior cortex. Figure 5 shows the frequency spectra for gains minus losses following feedback for when participants had chosen a face versus when they had chosen a house, along with the corresponding interaction effect (double difference wave). The frequency spectra in Figure 5, the alpha-power traces in Figure 6A, and the alpha-power topographical distributions (Figure 6B) indicate that there was lower amplitude alpha (starting \(\sim 900\)ms after feedback onset) following feedback of a gain as compared to a loss, both for when a face had been chosen and when a house had been chosen. These cluster-based analyses showed that during the initial part of this time frame (\(\sim 900\) to \(1400\)ms following feedback), participants had lower alpha power following gains as compared to losses within the inferior occipital ROI and occipital midline ROI but not the superior occipital ROI (main effect of feedback: inferior-lateral occipital: 900-1400ms, \(p = 0.006\); occipital midline: 1000-1400ms, \(p = 0.042\); superior occipital: no significant clusters found). In this earlier time period, however, significant interactions with chosen stimulus type were not observed.

In contrast, the long-latency alpha effect over the face-selective ROIs following the feedback was strikingly larger and lasted longer when participants had chosen a face versus a house. More specifically, topographical plots of alpha power (Figure 6B: interaction between feedback and choice) showed clear stimulus-related specificity for feedback processing over the face-selective ROIs delineated by the localizer task, suggesting the instantiation of stimulus-specific neural activity during feedback processing in these face-specific cortical regions. Specifically, for activity during this later post-feedback time period (but well before the next cue stimulus pair), we observed lower alpha power (i.e., higher cortical activation) for gains as compared to losses when the participant had chosen a face, an effect mostly absent when the participant had chosen a house. This observation was confirmed by cluster-based statistical analysis in the face-specific ROIs as defined by the localizer task (Table 2).

Figure 5. Stimulus-specific feedback related processes – frequency spectra: (A) Cortical updating of stimulus-reward associations: time-frequency plots of the differential activity to the feedback stimulus following gains versus losses (note the subtraction direction here) over occipital channels in the face ROIs (previously delineated by the localizer task), shown separately following having chosen a face versus having chosen a house. The plots reveal a difference in alpha power in these ROIs between 800 and 1800ms following the feedback for gains versus losses, but this effect was significantly larger when participants had chosen a face as compared to a house. (B) This late, cortically specific, alpha-power effect over the face areas can be seen more directly in the double subtraction (interaction) of activity, namely gains-minus-losses for faces minus gains-minus-losses for houses.

Table 2. Cluster based statistics related to the updating of stimulus-specific reward associations in the alpha (8 - 14Hz) frequency range.

<table>
<thead>
<tr>
<th>Region of interest</th>
<th>Time window, p</th>
<th>Inferior-lateral occip</th>
<th>Midline occip</th>
<th>Superior occip</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choosing x feedback</td>
<td></td>
<td>1400-1750ms, (p = 0.042)</td>
<td>1300-1850ms, (p = 0.004)</td>
<td>1400-1750ms, (p = 0.008)</td>
</tr>
<tr>
<td>Chose house gain versus loss</td>
<td>950-1150ms, (p = 0.008)</td>
<td>No sig. clusters</td>
<td>No sig. clusters</td>
<td></td>
</tr>
<tr>
<td>Chose face gain versus loss</td>
<td>900-1900ms, (p &lt; 0.002)</td>
<td>1000-2000ms, (p = 0.004)</td>
<td>1150-1850ms, (p = 0.012)</td>
<td></td>
</tr>
</tbody>
</table>
Considering that decreased alpha power in a cortical region typically reflects increased cortical activation (Green et al., 2017; Laufs et al., 2003; Mantini, Perrucci, Del Gratta, Romani, & Corbetta, 2007; Scheeringa et al., 2012), this pattern is consistent with the view that there was more cortical activation in the face processing areas after having chosen a face and received feedback of a gain than after having chosen a house and received feedback of a gain. In sum, the results revealed that, in response to feedback of a gain vs. a loss, there was an early general (i.e., stimulus-nonspecific) posterior cortical activation followed by a stimulus-specific one. It is important to note that this stimulus-specific alpha modulation following the feedback occurred even though the face/house stimuli had disappeared from view several seconds earlier.

**Figure 6.** Stimulus specific feedback related processes - alpha power: (A) Overall alpha power traces show a stimulus-related decrease in power followed by an increase. Cluster based analysis revealed statistical significance for various contrasts. Shaded area represents mean ± SEM. (B) ROI-specific alpha-power traces following gain minus loss feedback. The early part of this long-latency alpha modulation was stimulus-nonspecific (no difference for whether a face versus a house had been chosen on that trial. Later, the alpha activity showed stimulus-specific difference between gain and loss trials as a function of whether a face versus a house had been chosen on that trial.

**Discussion**

In the present study we sought to shed light on the neural mechanisms by which stimulus-reward associations might be updated in stimulus-specific cortex. We used a probability-learning gambling task with feedback. On each trial, participants chose between a face and a house stimulus and had to learn across each set of 20 trials which stimulus a higher probability of gaining a reward in that set. Because face and house stimuli elicit spatially distinct neural activations in scalp-recorded EEG, we were able to investigate the neural processes related to stimulus-specific feedback processing in the sensory cortices. Behaviorally, the results indicated that participants were able to learn which stimulus yielded a higher probability of reward in each 20-trial set. Neurally, this learning was marked by a cascade of changes in the brain electrical responses. First, feedback evaluation was most quickly reflected at ~250ms by the FRN, a negative fronto-central deflection, which was then followed by a modulation of the centro-parietal P3 (~400ms), likely reflecting a general increase in cognitive resources. Both of these cortical activations were larger for losses compared to gains. Subsequently (~900-1400ms after feedback onset), we observed either a decrease (for gains) or increase (for losses) of stimulus-nonspecific posterior oscillatory alpha activity that was mostly stimulus-nonspecific, but this was then followed by a strongly enhanced stimulus-specific activity for gains compared to losses (~1400-1800ms) over the sensory face areas when the participant had chosen a face, an effect that was mostly absent when they had chosen a house. Then on the next trial, the learning was further marked by a rapid attentional orienting towards the more-likely-to-be-rewarded stimulus when these stimuli were later presented as a cue pair, an effect that increased across the 20-trial set. These results expand our understanding of the cortical mechanisms by which stimulus-specific regions are activated during feedback learning in service of establishing and/or updating stimulus-reward associations.

Classically, the cascade of neural processes underlying feedback processing has been delineated as starting with feedback evaluation, as reflected by the FRN component for losses relative to gains, and continuing general increase in cognitive resources, as reflected by an enhanced P3 wave (Polich, 2007; René San Martín, 2012), if the feedback indicated a suboptimal outcome and thus that the choice behavior potentially needs to be adjusted (Fischer & Ullsperger, 2013; O’Connell et al., 2012). These effects were also seen in the present study. We also showed here that the FRN activity was stimulus non-specific, in that it did not differ as a function of whether the feedback was preceded by a face or house selection. This result is consistent with the idea that the ACC, a likely contributor to the FRN (Gehring & Willoughby, 2002), reflects the detection/evaluation of outcome information independent of the specific stimulus and/or event context.
One possibility would be that these stimulus-specific activations reflect a saliency to those stimulus types, the brain seems to then bias attention towards those stimuli. During learning is followed by stronger orienting of attention towards those stimuli when next encountered, and in shaping future choice behavior with respect to those stimuli. Accordingly, the present results provide new insight into the neural mechanisms underlying the updating of stimulus-reward associations in the brain, and thus in turn for understanding how the brain implements complex adaptive behavior.

Alpha power effects began as a general posterior cortical activation, greater for gains compared to losses, but irrespective of whether participants had chosen a face or a house. This effect might be regarded as an increased activation of visual processing regions following gain feedback. Following this general visual cortex activation, between 1400-1800 ms following feedback after having made a face choice, we observed activity over the sensory face regions that was substantially larger following gain feedback compared to loss feedback, with the corresponding effect not being observed following gain-vs-loss feedback after participants had chosen a house. Importantly, this stimulus-specific activation occurred even though the face (vs. house) stimuli had disappeared from the screen seconds before. The activation pattern suggests that these effects, occurring over the occipital-midline, occipital-superior, and occipital-inferior face-selective localizer-defined ROIs, potentially reflect activity in the occipital face area (OFA), superior temporal sulcus (STS), and fusiform face area (FFA), which have all been found to selectively process face stimuli (Gobbini & Haxby, 2007; Pitcher et al., 2011).

In the context of these results, we interpret the reduced alpha power over the face ROIs after a gain following a face choice to reflect cortical activation in the face-processing-selective cortical brain regions, as part of the establishing, updating, and storage of stimulus-reward associations. Arguably this principle extends to a general mechanism of cortical updating for stimulus-reward association for any stimulus type with respect to feedback. It seems likely that as feedback-guided updating establishes and strengthens stimulus-reward associations, the brain also establishes and strengthens a subsequent reactive attentional bias toward the more likely-to-be-rewarded stimulus. In our results this manifested as an enhanced N2pc towards the stimulus more likely to win on its next presentation in the cue image pair, paralleling the increased chance that this stimulus would be chosen subsequently at the choice screen. These findings are in line with previous studies investigating effects of reward on attentional orienting, which have similarly shown that once a stimulus-reward association has been learned, it tends to induce a larger attentional shift when presented and also predicts the likelihood of subsequent choice behavior (Hickey et al., 2010a; Hickey & van Zoest, 2012; San Martin et al., 2016). Our results substantiate these findings by implicating a key role of the sensory cortices in the cascade of processes involved in the formation and neural updating of stimulus-reward associations.

In response to stimulus-specific areas being activated following rewarded trials for those stimulus types, the brain seems to then bias attention towards those stimuli. One possibility would be that these stimulus-specific activations reflect a saliency adjustment of the rewarding objects. That is, in addition to coding for stimuli and features, it has been hypothesized that the visual system contains a saliency map for objects and their features (Itti & Koch, 2000; Itti, Koch, Way, & Angeles, 2001). Such a saliency map could serve as a prioritizing scheme that codes for the importance of various stimuli. By adjusting the saliency of each object category (here, faces or houses) the brain would prioritize the processing of certain objects over other ones. The neuroanatomical properties and location of such saliency maps are under debate but there is evidence of involvement of sensory visual areas (Çukur, Nishimoto, Huth, & Gallant, 2013; Li, 2002; Madden, 2007; Mazer & Gallant, 2003; Torralba, Oliva, Castelhano, & Henderson, 2006). If indeed the observed cortical activation in response to feedback information is a reflection of saliency map adjustment, then the results from the present study would have implications for understanding the mechanisms by which previously encountered outcome information might be involved in attentional prioritization and cognitive control more generally.

The present study examined mechanisms underlying the learning of reward associations along a relatively short stretch of trials (i.e., each set of 20 trials lasting several minutes), which would then change in the next stretch. Long-term learning in a stable environment would typically lead to value information being updated much more slowly. Notably, learning of object values with different stabilities has been shown to impact activity in the subcortical regions differentially (Kim & Hikosaka, 2013), suggesting different neural mechanisms are involved based on how stable the learning environment is. Thus, understanding how the processes delineated in this paper differ relative to within a more stable environment is an important goal for future studies.

In conclusion, we have shown that in response to positive feedback to choices made on specific stimulus types, the brain activates the sensory cortical regions coding for those specific stimulus types. This activation appears to reflect processes related to the establishment and neural updating in stimulus-specific cortical brain regions of the reward associations for those specific stimuli, perhaps by way of enhancing their representation in a cortically based saliency map. Moreover, the establishment and updating of these reward associations in stimulus-specific cortical processing regions during learning is followed by stronger orienting of attention towards those stimuli when next encountered, and in shaping future choice behavior with respect to those stimuli. Accordingly, the present results provide new insight into the neural mechanisms underlying the updating of stimulus-reward associations in the brain, and thus in turn for understanding how the brain implements complex adaptive behavior.
Author Contributions

Funding
This work was supported by NIH grant RO1-N5051048 to M.G.W.

Acknowledgements
We thank the editors and two anonymous reviewers for helpful comments on an earlier version of the paper.