Dynamic modulation of neural feedback processing and attention during spatial probabilistic learning

Pütz, Celina; van den Berg, Berry; Lorist, Monicque M

[Link to publication in University of Groningen/UMCG research database](#)

Article

Dynamic modulation of neural feedback processing and attention during spatial probabilistic learning
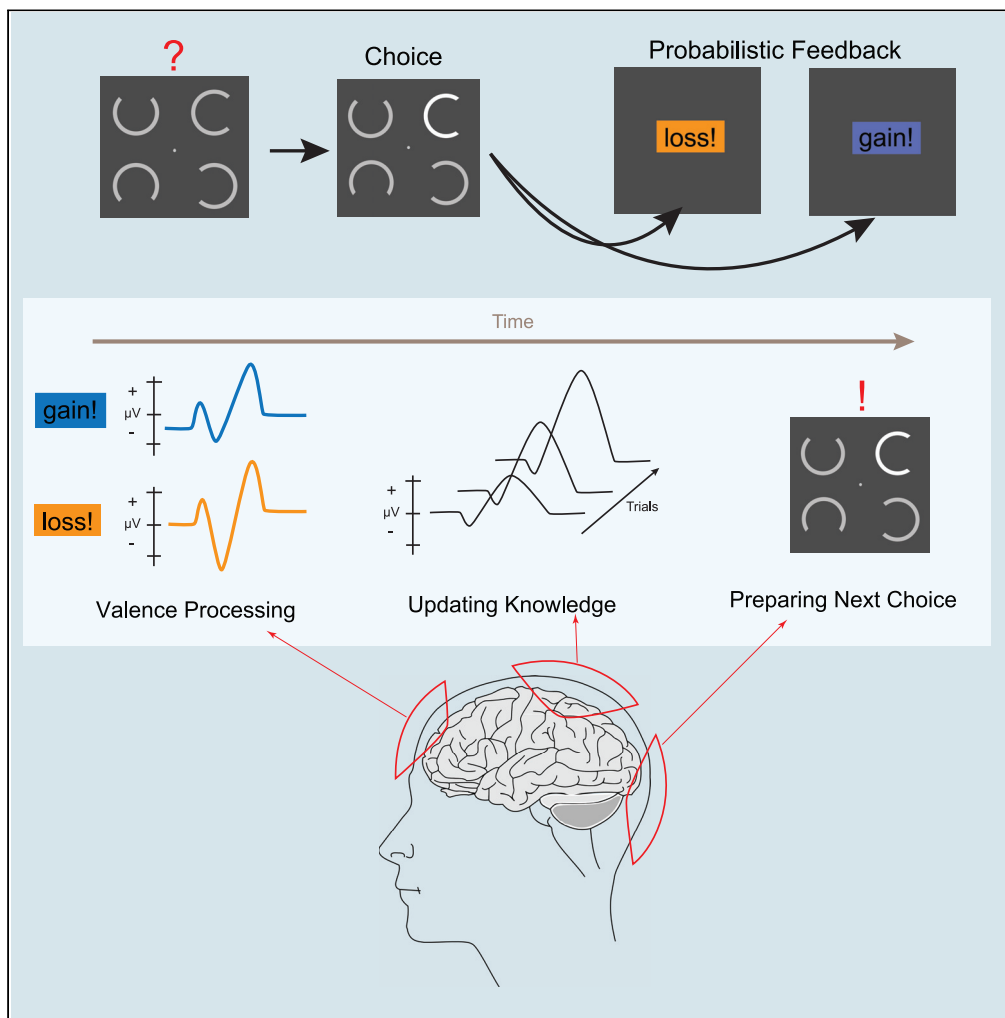
Celina Pütz, Berry van den Berg, Monicque M. Lorist

m.m.lorist@rug.nl

Highlights

We can learn which *spatial* location relates to the highest probability of reward

Neural processing of feedback valence was not influenced by learning

LPC amplitude was dynamically modulated by learning, reflecting context updating

Feedback-locked alpha power was predictive of ensuing orientation of attention

Article

Dynamic modulation of neural feedback processing and attention during spatial probabilistic learning

Celina Pütz,[1,2,3] Berry van den Berg,[1] and Monicque M. Lorist[1,4,*]

SUMMARY

Learned stimulus-reward associations can modulate behavior and the underlying neural processing of information. We investigated the cascade of these neuro-cognitive mechanisms involved in the learning of *spatial* stimulus-reward associations. Using electroencephalogram recordings while participants performed a probabilistic spatial reward learning task, we observed that the feedback-related negativity component was more negative in response to loss feedback compared to gain feedback but showed no modulation by learning. The late positive component became larger in response to losses as the learning set progressed but smaller in response to gains. In addition, feedback-locked alpha frequency oscillations measured over occipital sites were predictive of N2pc amplitudes—a marker of spatial attention orienting—observed on the next trial. This relationship was found to become stronger with learning set progression. Taken together, we elucidated neurocognitive dynamics underlying feedback processing during spatial reward learning, and the subsequent effects of these learned spatial stimulus-reward associations on spatial attention.

INTRODUCTION

Our ability to optimize future behavior based on information we received as the result of our previous behavior is an essential part of navigating through everyday life. Especially, stimuli in our environment that we associate with rewarding outcomes shape our behavior, for instance by capturing our attention, a process that has been termed value-based attentional capture (Anderson, 2016). Research on value-based attentional capture consistently showed that reward-associated nonspatial stimulus features can modulate attention voluntarily or involuntarily, depending on their task relevance (Bourgeois et al., 2017). This is manifested behaviorally through faster response times (RTs) in selecting target items that carry reward-associated features (Kiss et al., 2009) and slower RTs when distractor items do so (Anderson et al., 2016). However, behavioral findings on the effects of reward-associated *spatial* features on attention are conflicting, with some showing that reward-associated locations in our visual field can attract our attention similarly to reward-associated nonspatial features (Mine et al., 2021; Sisk et al., 2020; Anderson and Kim, 2018; Chelazzi et al., 2014), whereas others do not find an effect (Won and Leber, 2018). This discrepancy has been attributed to the role of awareness of participants of the contingencies between spatial locations and reward, with successful spatial bias only occurring when we are explicitly aware of this association (Anderson et al., 2021; Mine et al., 2021).

Studies investigating the effect of stimulus-reward associations on behavior generally train their participants first, before testing the effect of these associations in an unrelated subsequent experimental task (Sali et al., 2014). Here, we explicitly sought to investigate the cascade of neurocognitive mechanisms underlying the learning of spatial stimulus-reward associations, with a specific focus on feedback processing and related modulations of spatial attention in subsequent encounters with these reward-related stimuli.

Learning stimulus-reward associations is guided by feedback information that follows a behavior. The presentation of feedback elicits a cascade of neural processes that ranges from feedback valence processing in subcortical areas of the brain (Delgado et al., 2005) to the monitoring and updating of stimulus-reward contingencies and subsequent decision making with regard to whether or not future behavior needs to be adjusted in the frontal cortices (Rushworth et al., 2011). In addition, the sensory cortices show changes in neural activity dependent on the sensory features of the reward-associated stimuli (van den Berg et al., 2019; Henschke et al., 2020; Schiffer et al., 2014; Folstein et al., 2013). The result of these neurocognitive

[1]Department of Experimental Psychology, University of Groningen, Grote Kruisstraat 2/1, Groningen 9712TS, the Netherlands

[2]Department of Neurobiology, University of Groningen, P.O. Box 11103, Groningen 9700CC, the Netherlands

[3]Department of Neurology, University Medical Center Groningen, Postbus 30001, Groningen 9700RB, the Netherlands

[4]Lead contact

*Correspondence:
m.m.lorist@rug.nl

https://doi.org/10.1016/j.isci.2022.104302

processes might ensure that if the rewarded stimulus is encountered again in the future, attention will be directed to this information, resulting in faster and/or more accurate performance (e.g., Hickey et al., 2010).

Electroencephalography (EEG) and event-related potential (ERP) recordings allow us to study the temporal sequence of the cascade of processes involved in feedback processing during spatial stimulus-reward learning. For example, peaking around 200 ms after feedback presentation, the fronto-central negative-polarity feedback-related negativity (FRN) component was found to be more negative in response to loss feedback compared to gain feedback (Miltner et al., 1997; Hajcak et al., 2006; Holroyd and Coles, 2002; Nieuwenhuis et al., 2004; for a review see: San Martín, 2012). Interestingly, however, little is known about whether FRN amplitude tracks the learning progress. Traditionally, the FRN has been regarded as a neural correlate of the reward prediction error (RPE), reflecting the difference between expected versus actual reward outcomes (Sambrook and Goslin, 2015; Cohen et al., 2007; Nieuwenhuis et al., 2004). The RPE has been argued to be an essential part of learning, as the comparison between expected and actual outcomes is what makes individuals adjust their behaviors to match the actual reward outcome (Schultz, 2016).

The FRN is followed by the late positive component (LPC), a positive-going wave starting around 400 ms after stimulus presentation, with maxima occurring at parieto-central electrode sites (Trimber and Luhmann, 2017; Muller-Gass et al., 2019). The LPC has originally been studied in the context of affective processing (Paller et al., 1995). In the context of stimulus-reward learning, LPC amplitudes were observed to be larger in response to loss feedback than to gain feedback (Trimber and Luhmann, 2017). Furthermore, LPC amplitudes were found to be sensitive to reward expectancy, with largest amplitudes elicited by unexpected losses (Trimber and Luhmann, 2017; Muller-Gass et al., 2019; Donaldson et al., 2016). In addition, LPC amplitudes in response to losses are found to be linked to subsequent behavioral adjustments, suggesting that the LPC plays a role in behavioral optimization based on feedback (San Martín et al., 2013; von Borries et al., 2013). Thus, the LPC is reflective of the affective-value processing of feedback, which incorporates active context updating based on the feedback that was received to optimize behavior (Polich, 2007; Glazer et al., 2018).

As a result of feedback processing and learning of stimulus-reward associations, the sensitivity of the sensory cortex for reward-related information might increase. Evidence supporting this account stems from neuroimaging studies that indeed report enhanced neural discriminability of relevant stimulus features in the sensory cortex after learning (Henschke et al., 2020; Schiffer et al., 2014; Folstein et al., 2013). Moreover, the post-feedback re-activation of the sensory cortex was found to be localized to those areas of the visual cortex that were involved in initial processing of the reward-related stimuli. This re-activation was therefore argued to serve as a signal to enhance the sensory representations of rewarded information, which in turn might function as a preparatory mechanism facilitating future processing of reward-related information (Fitzgerald et al., 2013; Weil et al., 2010).

Intriguingly, similar findings have been obtained with alpha power in the EEG signal (8–14 Hz), which is inversely linked to the fMRI BOLD signal (Goldman et al., 2002; Scheeringa et al., 2011), indicating that activity in the alpha band might be related to cortical arousal. Alpha power was found to be significantly suppressed at occipital electrode sites contralateral to the to-be-attended stimulus in latency ranges following the LPC (van den Berg et al., 2019). Importantly, these modulations in alpha power occurred in the absence of any sensory stimulation after the presentation of feedback. One possible explanation for this post-feedback alpha activity in the sensory cortex is that the suppression of alpha power might reflect a preparatory activity bias in the visual cortex, which leads to enhanced neural processing of reward related stimuli (Worden et al., 2000; Zhao et al., 2019). Similar neural mechanisms are observed in the context of preparatory spatial attention. Here, top-down influence of higher order brain areas is thought to induce a preparatory bias in the visual cortex before the onset of a stimulus array (Hopfinger et al., 2000). However, whether feedback-related brain activity in the sensory cortex underlies a similar preparatory mechanism and whether this is in turn modulated by learning is yet to be determined.

The learning of stimulus-reward associations ultimately are expected to induce changes in attentional deployment. Spatial attentional orienting can be detected neurally via the N2pc ERP (Luck and Hillyard, 1994), a neural reflection of spatial orienting of attention that starts within 250 ms after the presentation of a bilateral stimulus array. The N2pc is a lateralized component that is larger on electrode sites contralateral to the attended location and that shows larger amplitudes if the attended target is associated with reward (Kiss et al., 2009). N2pc amplitudes were found to be predicted by preparatory alpha power

Table 1. Overview of mixed effects models used for our statistical analyses.

| Analysis	Model
Behavioral Choice	$p(\text{choosing set} - \text{winner}) = \beta_{0i[n]} + \beta_{0l[n]} + trial\# * \beta_{1i[n]} + trial\#^2 * \beta_{2i[n]} + \varepsilon_{il[n]}$
RTs	$RT = \beta_{0i[n]} + trial\# * \beta_{1i[n]} + prob_{Choice} * \beta_{2[n]} + trial\# * prob_{Choice} * \beta_{3[n]} + \varepsilon_{i[n]}$
FRN analysis	$\mu V_{FRN\ latency\ rage} = \beta_{0i[n]} + trial\# * \beta_{1[n]} + feedback * \beta_{2i[n]} + trial\# * feedback * \beta_{3[n]} + \varepsilon_{i[n]}$
LPC analysis	$\mu V_{LPC\ latency\ range} = \beta_{0i[n]} + trial\# * \beta_{1i[n]} + feedback * \beta_{2i[n]} + trial\# * feedback * \beta_{3i[n]} + \varepsilon_{i[n]}$
N2pc analysis (t-test)	$\mu V_{N2pc\ latency\ range} = \beta_{0i[n]} + \varepsilon_{i[n]}$
Lateralized alpha power analysis (t-test)	$alpha\ power = \beta_{0i[n]} + \varepsilon_{i[n]}$
N2pc predicted by alpha power analysis	$\mu V_{N2pc(next\ choice)\ latency\ range} = \beta_{0i[n]} + alpha\ power * \beta_{1i[n]} + trial\# * \beta_{2i[n]} + alpha\ power * trial\# * \beta_{3i[n]} + \varepsilon_{i[n]}$

suppression observed during performance of attentional selection tasks (Zhao et al., 2019). Thus, the N2pc makes a suited candidate to study neural attention orienting in space and reward-association learning.

In the present study, we sought to investigate whether reward could be coupled to a spatial location, using a probabilistic reward-learning task. In addition, we examined the cascade of neurocognitive mechanisms involved in learning spatial stimulus-reward associations based on feedback information, and how learning these associations in turn modulates subsequent spatial attention and behavioral performance. First of all, we expected that participants would learn the spatial stimulus-reward association as read out by an increased probability of choosing the highest gain-associated location stimulus as the learning set progresses. In addition, we expected that RTs would become faster in response to choosing the set winning location stimulus, indicating faster information processing as the reward-association is learned.

Neurally, we expected to observe several modulations of neural markers of feedback processing, starting with amplitude differences between loss and gain feedback trials reflected in the FRN component. Here, we expected to find more negative FRN amplitudes in response to loss feedback, replicating the vast literature on the FRN and thereby providing an indicator that neural processing of feedback information occurred as expected using a *spatial* stimulus-reward learning task. We further expected that if the FRN represents a neural RPE, the difference between loss and gain feedback would become larger with trial progression as reward expectancy increases with learning. That way, we aimed to explore the FRN's role in active trial-by-trial learning. In addition, we expected to observe significant modulations in LPC amplitude, which is used as a marker of context-updating following the FRN. We expected to observe more positive LPC amplitudes for losses relative to gains, and that LPC amplitudes reduce in size once the location-reward association has been learned because less context-updating is needed when stimulus-reward associations have been established. Following the LPC, but not before the next choice has been made, we hypothesized to observe significantly lower alpha power on contralateral electrode sites relative to the upcoming location choice, indicating stimulus-specific preparatory cortical activity. If indeed lateralized alpha activity serves as a preparatory mechanism modulating the sensitivity of the sensory cortex to reward-associated stimulus features, we further expected that feedback-locked alpha power suppression contralateral to the to-be-attended location would be significantly related to the subsequent orientation of attention, indexed by N2pc amplitude, toward that location. Lastly, we hypothesized that this relationship will become stronger with learning, indicating the strengthening of visual cortex representations of the reward-associated location stimulus and providing a potential framework on how post-feedback cortical activity eventually leads to neural activity changes in the visual cortex after learning.

RESULTS

All mixed-effects models are summarized in Table 1 in the STAR methods section.
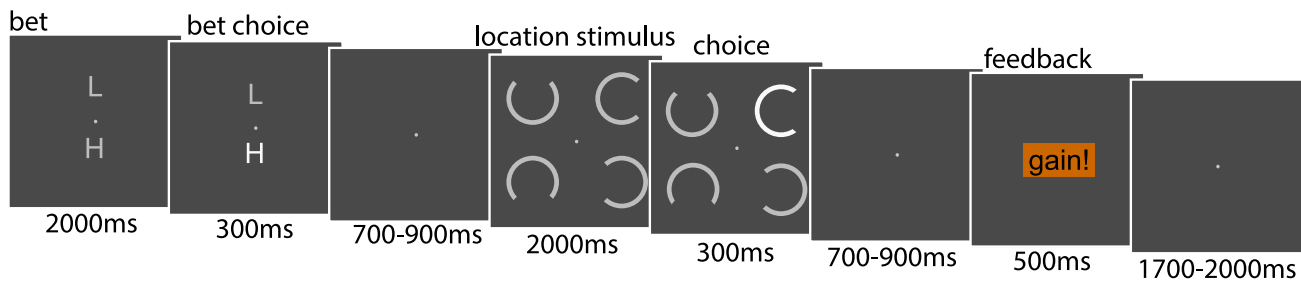
Figure 1. Trial-sequence of the probabilistic reward-based spatial learning task

First, participants had to choose between betting high or low by choosing the corresponding letter. Then, their choice of bet was highlighted. Next, participants chose a location by selecting the opening of the circle that was placed in the location on that trial. The choice of the opening of the circle was also highlighted after. Lastly, feedback in the form of the words "gain!" and "loss!" appeared to guide their choice. Participants underwent the trial-sequence 40 times in each set before the set winner changed.

Behavior

Learning and response speed

The learning curves were not dependent on the set winning location, supporting that participants had no general bias towards one of the four stimulus locations (interaction 'learning set half by location' $F(3,188)$ = 0.57, p = 0.6; Figure 1B). In line with our expectations, the proportion of times that participants chose the set winner compared to choosing a non-set winner increased with increasing trial number (Wald's $X^2(2, 10806)$ = 204.07, p < 0.001, Figure 2A). Furthermore, the average amount of money that participants earned was 13.20 € (SD = 6.76 €), indicating that they indeed aimed at maximizing their monetary gains and learned the most rewarding location stimulus. In addition, subject's RTs decreased with increasing trial number (main effect 'trial number' $F(2, 29.1)$ = 7.3, p < 0.01) and were significantly different between trials in which participants chose the set winning location stimulus versus when they chose a non-set winning location stimulus (main effect 'choosing set winner versus not choosing set winner' $F(1, 10180.8)$ = 20.95, p < 0.0001; M_{diff} = −17.51 ms, SE = 3.8) (Figure 2C). However, the change of RTs across the learning set between trials in which the set winner was chosen versus when the set winner was not chosen was not statistically significant (interaction effect 'RT by choosing set winner versus not choosing set winner': $F(2, 7829.6)$ = 0.082, p = 0.92).

Betting behavior

Our analyses of participant's betting behavior yielded a significant effect for 'trial number' on bet (Wald's $X^2(1, 10806)$ = 27.67, p < 0.001, Figure 1A). Participants placed higher bets increasingly more often with increasing trial numbers. Because participant's betting curves showed large similarities to their learning curve, we did not include the bet in our EEG analyses. A descriptive figure of betting behavior can be found in the supplementary material S1.

EEG

Feedback processing

We observed a frontocentral ERP in the FRN latency range which amplitude was, as expected, significantly more negative in response to loss feedback compared to gain feedback (main effect 'feedback' $F(1, 30)$ = 34.1, p < 0.001; $M_{gain\ minus\ loss}$ = - 1.38, SE = 0.24). However, we did not observe a statistically significant effect of trial number on mean ERP amplitudes in the FRN latency range (main effect 'trial number' $F(1, 10190)$ = 1.25, p = 0.26) nor did we observe a statistically significant change of mean amplitudes with trial number depending on loss or gain feedback (interaction effect 'feedback by trial number,' $F(1, 10174)$ = 3.04, p = 0.08). Lastly, model comparison between a mixed model including an interaction effect of 'feedback' by 'trial number' and a model without an interaction effect remained inconclusive based on the difference in AIC alone ($AIC_{Interaction}$ = 66121; $AIC_{Main\ effects\ only}$ = 66122). Therefore, we also took the BIC of both models into account during model selection, which revealed that the best fitted model was the more parsimonious main effects only model ($BIC_{main\ effects\ only}$ = 66172, $BIC_{Interaction}$ = 66179).

In line with our expectations, mean amplitudes in the LPC latency range showed significant changes with trial number (main effect 'trial number' $F(1, 10180.3)$ = 9.34, p = 0.002). In addition, mean amplitudes in
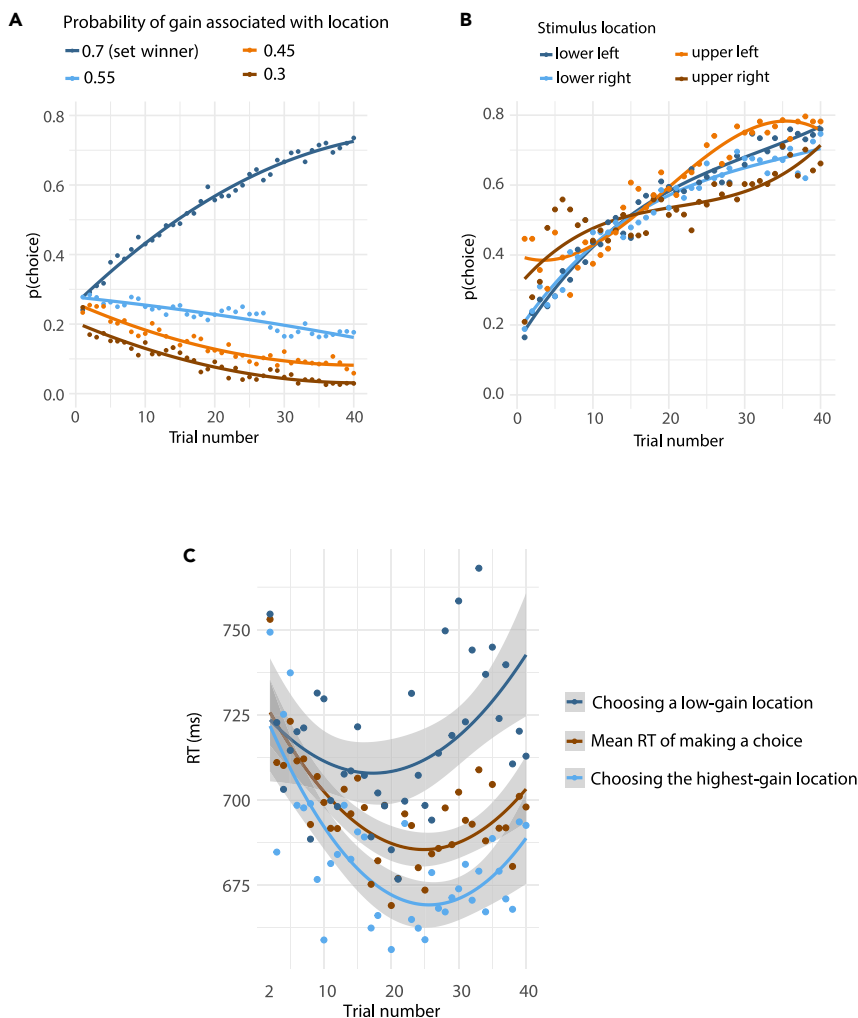
Figure 2. Behavioral results

(A) The probability of choosing a location with the highest gain probability increased with trial number, whereas the probability of choosing all other locations decreased.

(B) All four locations showed similar probability curves when they had the highest gain probability assigned.

(C) Participants were faster in choosing the highest gain location stimulus compared to the other location stimuli across all trials. The difference in RTs between gain probabilities remained stable with trial numbers.

the LPC latency range significantly differed between gains and losses dependent on trial number (interaction effect 'trial number by feedback', $F(1, 10172.9) = 19.82$, $p < 0.0001$). As illustrated in Figure 3B, average amplitudes became more positive in response to losses with increasing trial number, whereas gain feedback elicited progressively smaller amplitudes ($M_{slope(gains)} = -0.02$, $SE = 0.006$; $M_{slope(losses)} = 0.02$, $SE = 0.007$).

Alpha power

The cluster-based permutation test showed a significant cluster in post-feedback lateralized alpha on occipital electrode sites contralateral to the upcoming location stimulus of choice between 600 and 2000 ms (cluster statistic $p < 0.001$, Figures 4B and 4C). A subsequent t-test on average alpha power in the 600 to 1200 ms post-feedback interval (using the TOI extracted for our N2pc-alpha power model analyses) revealed a confirmatory result ($t(28.35) = -2.08$, $p = 0.046$), indicating that alpha power is significantly reduced at contralateral channel sites relative to the upcoming location of choice within our TOI. Thus, we proceeded to use the average alpha power extracted from our TOI in the model analysis.

FRN

LPC

Gain minus Loss
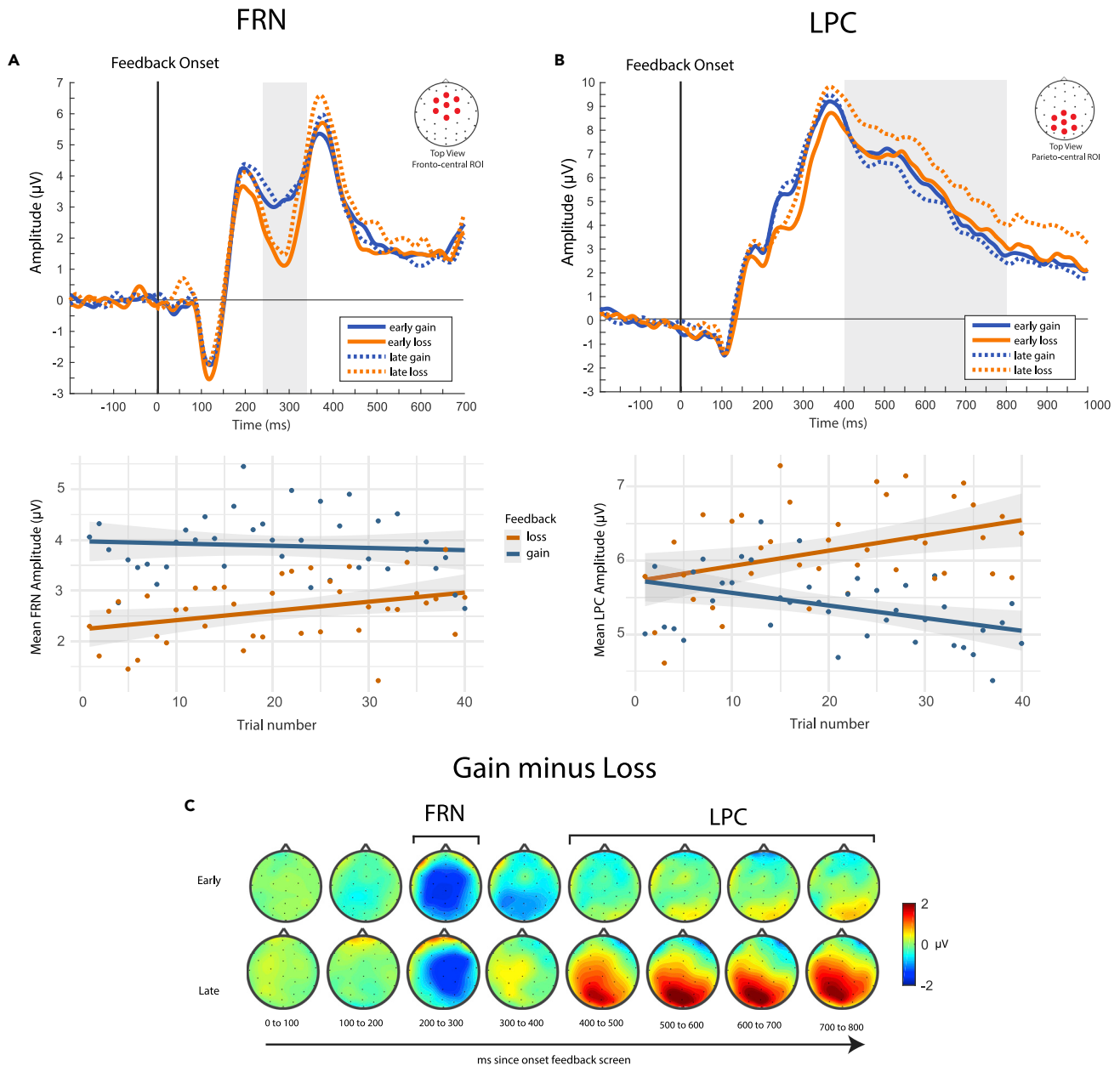
Figure 3. Feedback-locked ERPs early (first 8 trials) and late (last 8 trials) in the learning set

(A) Fronto-central amplitudes between 240 and 340 ms were more negative in response to loss feedback and showed no difference between early and late trials in the learning set.

(B) Parietal amplitudes between 400 and 800 ms did not distinguish between losses and gains at the beginning of the learning set but showed larger responses to losses later on.

(C) Topography distribution of the gain minus loss effect indicates a fronto-central FRN (200 ms to 300 ms) and a parieto-central LPC (400 to 800 ms) early and late in the set. Values show the difference between gain and loss feedback (gain minus loss).

Relationship between alpha power and the N2pc across the learning set

The presentation of the location stimuli elicited an N2pc that was significantly larger on electrode sites that were contralateral to the stimulus location of choice, reflecting the orienting of spatial attention towards the location of choice ($t(30.18) = -7.06$, p < 0.0001, Figure 4A). Next, we examined whether average post-feedback alpha power and post-stimulus N2pc amplitude relative to the upcoming location choice would show a significant relationship that was modulated by trial number. This was indeed the case
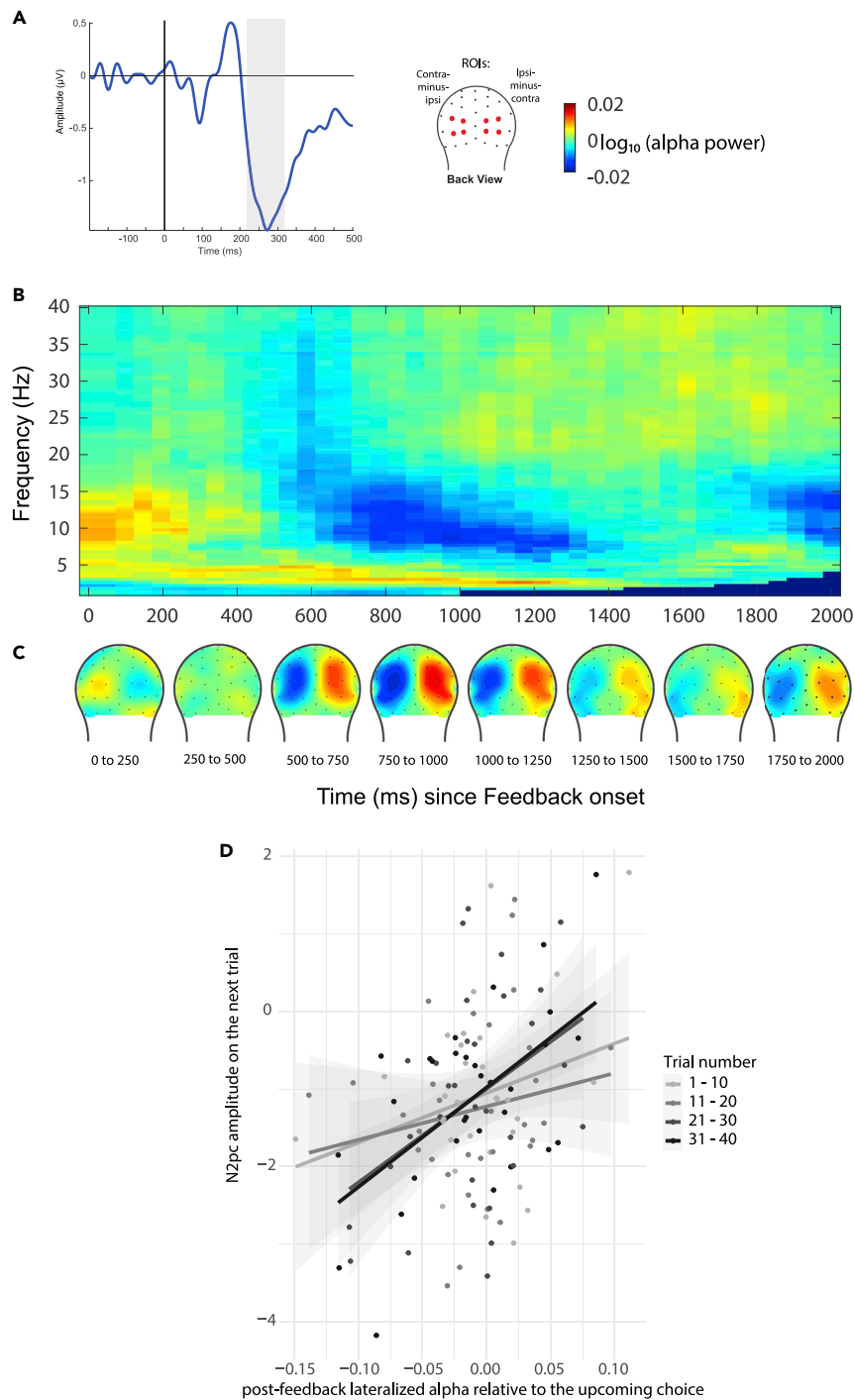
Figure 4. The N2pc component and feedback-locked lateralized alpha power, calculated as a function of contralateral-versus-ipsilateral electrode sites to choice

(A) Right: Difference wave of the N2pc component relative to the location stimulus of choice. The difference was calculated by subtracting ipsilateral amplitude values from contralateral amplitude values relative to the choice location. Left: Graphical representation of ROIs and alpha power scale.

(B) Time-frequency spectrum of our time-frequency analyses shows that the decrease in contralateral power is specific for the alpha band (8-14 Hz).

Figure 4. *Continued*

(C) Topographical distribution of feedback-locked lateralized alpha power shows that contralateral-to-upcoming choice alpha activity is significantly lower compared to ipsilateral activity at occipital electrode sites.

(D) Post-feedback lateralized-to-choice alpha power was significantly predictive of N2pc amplitudes on the next trial, in that the smaller the lateralized alpha power, the more negative N2pc amplitudes. This relationship was the strongest in later trials of the learning set.

(interaction effect 'trial number by lateralized alpha' $F(1, 9687.7) = 4.72$, p = 0.03). Specifically, the more pronounced suppression of post-feedback alpha power relative to the upcoming choice, the larger the N2pc amplitudes relative to the choice on the upcoming trial (Figure 4D). In addition, as shown in Figure 4D, lateralized alpha power predicted upcoming N2pc amplitudes stronger in later trials compared to the beginning of the learning set.

DISCUSSION

Rewarded stimulus features in our environment can voluntarily or involuntarily capture our attention (Anderson, 2016; Bourgeois et al., 2017). Yet, whether the same holds true for rewarded spatial locations is less straightforward (Sisk et al., 2020; Anderson and Kim, 2018; Chelazzi et al., 2014; Won and Leber, 2018). Here, we aimed to map out the cascade of neurocognitive mechanisms that underlie the learning of spatial stimulus-reward associations in order to shed more light on the role of spatial reward-associations in value-based attentional capture. Our focus lies on neural reward-feedback processing, which is an essential part of learning (Schultz, 2015). Our findings delineate how explicitly learned spatial reward associations influence subsequent orientation of attention towards reward related information.

We found significant differences in amplitudes between gain and loss feedback conditions for the FRN and LPC ERP components, reflecting feedback processing and context-updating, respectively. Only the LPC showed significant modulations in amplitudes as the learning set progressed. Subsequently, we observed a reduction in lateralized alpha power, reflecting activity in the visual cortex contralateral to the upcoming location of choice, which was significantly related to the N2pc—a neural correlate for the orienting of spatial attention—elicited by subsequent stimuli. This relationship between alpha power suppression and N2pc amplitude became stronger with learning. These findings provide insights into the cascade of neurocognitive mechanisms involved in the learning of *spatial* stimulus-reward associations affecting subsequent behavior.

By using a design with multiple unique learning sets of 40 trials, we were able to map the behavioral learning curves of spatial stimulus-reward associations. Similar behavioral results are also commonly reported in two-choice reward learning tasks which deploy nonspatial target-features as their to-be-rewarded stimuli (van den Berg et al., 2019; Bourgeois et al., 2017). In addition, the learning of stimulus-reward associations is a prerequisite in most studies investigating their attentional capture (Anderson et al., 2011; Anderson and Kim, 2018; Chelazzi et al., 2014). Given that the average net gain was positive in our study, our data support that participants actively used the feedback information to learn the spatial stimulus-reward association.

The first neural difference between loss and gain feedback in our study occurred in the form of a modulation of the FRN, an ERP component that is commonly associated with the RPE (Bellebaum and Daum, 2008; Nieuwenhuis et al., 2004; Cohen et al., 2007). Conceptually, the RPE reflects the difference in expected versus delivered feedback outcomes (Schultz, 2016). Given this interpretation, we expected to observe FRN amplitudes to be more negative for losses than for gains, and that this difference would become larger when the stimulus-reward association is learned and reward expectancies become stronger. Although we did observe FRN amplitudes to be different between loss and gain feedback, this difference did not change with trial number—that is with learning. Therefore, our findings suggest that during learning, the FRN represents feedback processing on a valence-level only, without taking expected feedback outcome into account (see also van Borries et al., 2013).

Following the FRN, we observed that LPC amplitudes were influenced by trial number; in response to loss feedback the LPC amplitude increased as the learning set progressed, whereas gain feedback elicited gradually smaller LPC amplitudes. The LPC has been suggested to reflect affective feedback processing

and subsequent context updating (Polich, 2007; Glazer et al., 2018). Late losses could signal that the current choice might need re-evaluation to stay on track of the most rewarding stimulus choice. This interpretation ties in with findings that report increased LPC amplitudes in response to unexpected losses (Trimber and Luhmann, 2017; Muller-Gass et al., 2019; Donaldson et al., 2016) and links between the LPC and subsequent behavioral adjustments (San Martín et al., 2013; Chase et al., 2011). Gain feedback received later in a learning set, on the other hand, does not provide any additional or novel information about the context or the choice that is being made that necessitates updating; hence, LPC amplitudes become smaller. Notably, research on the LPC in the context of learning from rewarding feedback is sparse, and the literature is just beginning to grow (Glazer et al., 2018). Originally, the LPC was associated with processes related to memory and affect (Hajcak et al., 2009). Our findings do not exclude the interpretation that late loss feedback simply creates greater affective responses because the reward association is known, but people still experience a loss, thereby eliciting greater responses in the LPC. Nonetheless, here we report distinct differences in LPC amplitudes in response to gain and loss feedback with learning, suggesting that LPC is dynamically modulated by learning stimulus-reward associations.

In latencies overlapping with the LPC and extending to more than 1000 ms after feedback presentation, we observed that post-feedback alpha power was lower on occipital electrode sites contralateral to the to-be-attended spatial location compared to ipsilateral sites. In general, power in the alpha band has been found to be inversely linked to the BOLD signal (Goldman et al., 2002; Scheeringa et al., 2011), wherein a reduction in alpha power is associated with increased cortical activity. Within this framework, our findings could indicate increased preparatory activity in visual brain areas contralateral to the upcoming choice, similar to preparatory spatial attention elicited by cue stimuli providing information about the location of subsequent stimuli (Worden et al., 2000). Similarly, an increase in occipital alpha power contralateral to the presentation of irrelevant information is linked to cortical inhibition (Schneider et al., 2019), most likely to maximize the processing of upcoming spatial reward information (Heuer et al., 2017). The alpha power suppression on contralateral sites that we observed in our data logically meant that alpha power on ipsilateral sites was increased. It is therefore possible that inhibition of cortical activity related to irrelevant choice locations helped in the enhanced response toward the eventually chosen location. Here, more research is needed to shed light on the interplay of hemispheric asymmetries in preparatory alpha power.

Interestingly, contralateral alpha power following feedback information was predictive of N2pc amplitudes elicited by the subsequent location stimuli, suggesting that it could be related to attentional orienting towards the next location of choice. Specifically, preparatory activity in the sensory cortices is thought to facilitate attentional processing towards task relevant information, as it is correlated with subsequent N2pc amplitudes relative to the attended location (Zhao et al., 2019) but also with decreased RTs (van den Berg et al., 2016). Multivariate pattern analyses also recently showed that preparatory contralateral alpha power boosts subsequent stimulus processing (Barne et al., 2020). Here, we extend those findings by reporting that preparatory alpha power also occurs in stimulus-reward learning tasks, as a direct function of the presentation of feedback related to the choice that was made. In the present study, however, participants received probabilistic feedback, which means that there is no direct mapping between performance and feedback on a single trial; however, to learn the statistical regularities between actual performance and probabilistic feedback, participants had to integrate information about stimulus-reward associations over trials. Most importantly, in the present study, we observed that the relationship between alpha power and N2pc becomes stronger with learning. Should the contralateral alpha power that we observe indeed reflect sensory cortical activity, our findings could potentially reflect the neural process in which sensoricortical activity is modulated by learning, thereby leading to the neural activity changes observed in the sensory cortices post-learning (Folstein et al., 2013; Schiffer et al., 2014).

Given the temporal order of neural feedback processing, disclosed by FRN and LPC amplitudes, the outcome of feedback processing could lead to an "educated" decision on which choice should be made next. This, in turn, induces preparatory activity in the visual cortex related to the upcoming choice. This interpretation would be in line with results from invasive electrophysiological animal studies that have shown that neurons in the sensory cortices become tuned towards the reward-predicting stimulus as a result of reward learning, by showing more discriminability and excitability in response to it (Jurjut et al., 2017). Additional findings in mice have shown that visuospatial attention is accompanied by enhanced neuronal responses and synaptic activity in the visual cortex (Speed et al., 2020).

Few studies to date have investigated the role of the sensory cortices in reward feedback processing, specifically, the neurotemporal order of when they come into play. Neuroimaging studies demonstrated that the sensory cortices show activity post-feedback but not *when* in the processing cascade this activation occurs (Schiffer et al., 2014), whereas invasive animal studies mainly focus on changes in neuronal firing patterns in the sensory cortices only but not on cognitive processing that occurs beforehand and in higher-order brain areas (Jurjut et al., 2017; Speed et al., 2020). Lastly, one recent paper found that activity in the stimulus-specific sensory cortex occurred post-feedback of the *previous* choice (van den Berg et al., 2019). Our data suggests that activity in the visual cortex after feedback presentation is significantly related to subsequent attentional processing and therefore could reflect preparatory mechanisms as a function of feedback processing. Moreover, we showed that the relationship between preparatory activity and attentional processing becomes stronger with reward learning, suggesting that the strength of the stimulus-reward association influences the preparatory response of the relevant sensory regions towards the to-be-attended stimulus location.

Although our findings show that spatial stimulus-reward associations can be learned and also affect subsequent attentional processing, one open question remains of whether the learned stimulus-reward association generalizes to attentional orientation in an unrelated task. Research on feature-based stimulus-reward association learning suggests that reward-associations induce more permanent changes in the feature-related sensory cortex, which subsequently leads to attentional capture (Tankelevitch et al., 2020; MacLean and Giesbrecht, 2015). A similar proposition has been made for spatial stimulus-reward associations (Chelazzi et al., 2014) but not yet been tested neurally. Recent papers suggest that spatial value-based attentional capture is only successful if participants are aware of the spatial reward contingencies during learning (Anderson et al., 2021; Mine et al., 2021). Our study shows that if those contingencies are known, significant changes in neural activity are observed in sensory areas, suggesting that a similar mechanism as observed in feature-based attentional capture is at play under those circumstances. What is left to discover is whether implicit reward learning is also accompanied by post-feedback neural activity changes in the visual cortex and subsequent attentional orienting or whether those responses are absent.

To sum up, we set out to investigate the neurocognitive cascade involved in feedback processing and subsequent modulations of spatial orientation of attention that occur during probabilistic spatial stimulus-reward learning. First, we observed that feedback valence processing as reflected through the FRN is not modulated by learning, and we conclude that the FRN only signals binary feedback outcome processing. The LPC, on the other hand, was dynamically modulated based on feedback and learning, supporting the view that it reflects neural updating of stimulus-reward-associations. Next, we found that preparatory brain activity in the visual cortex was lateralized toward the choice that was about to be made and predictive of later attentional orienting read out via the N2pc—a relationship that became stronger with learning.

LIMITATIONS OF THE STUDY

There are some limitations to this study that should be addressed. Most importantly, we analyzed most neural components as a function of trial number to draw conclusions about their modulations in learning. Learning is a complex process that involves the incorporation of previous and current knowledge on each trial, taking feedback and choice history into account. Our analysis of trial numbers simplifies this process, and future research can build on our findings to get a detailed picture of how the FRN, LPC, and alpha power are modulated by learning.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAIL
  - Participants

- METHOD DETAILS
  - Apparatus
  - Procedure
  - Probabilistic spatial learning task
  - EEG recording and pre-processing
  - ERP analysis and data extraction
  - Time-frequency analysis and data extraction
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Trial exclusion and model selection
  - Behavioral analyses
  - EEG analyses
  - Time-frequency analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104302.

AUTHOR CONTRIBUTIONS

Conceptualization, C.P., B.v.d.B., and M.L.; Investigation, C.P., and B.v.d.B.; Writing – Original Draft, C.P., B.v.d.B., and M.L.; Writing – Review & Editing, C.P. and M.L.

DECLARATION OF INTEREST

The authors declare no competing interests.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. IEEE Trans. Automat. Contr. *19*, 716–723. https://doi.org/10.1109/tac.1974.1100705.

Anderson, B.A. (2016). The attention habit: how reward learning shapes attentional selection. Ann. N. Y. Acad. Sci. *1369*, 24–39. https://doi.org/10.1111/nyas.12957.

Anderson, B.A., and Kim, H. (2018). Mechanisms of value-learning in the guidance of spatial attention. Cognition *178*, 26–36. https://doi.org/10.1016/j.cognition.2018.05.005.

Anderson, B.A., Folk, C.L., Garrison, R., and Rogers, L. (2016). Mechanisms of habitual approach: failure to suppress irrelevant responses evoked by previously reward-associated stimuli. J. Exp. Psychol. Gen. *145*, 796–805. https://doi.org/10.1037/xge0000169.

Anderson, B.A., Kim, H., Kim, A.J., Liao, M.R., Mrkonja, L., Clement, A., and Grégoire, L. (2021). The past, present, and future of selection history. Neurosci. Biobehav. Rev. *130*, 326–350. https://doi.org/10.1016/j.neubiorev.2021.09.004.

Anderson, B.A., Laurent, P.A., and Yantis, S. (2011). Value-driven attentional capture. Proc. Natl. Acad. Sci. U S A *108*, 10367–10371. https://doi.org/10.1073/pnas.1104047108.

Barne, L.C., de Lange, F.P., and Cravo, A.M. (2020). Prestimulus alpha power is related to the strength of stimulus representation. Cortex *132*, 250–257. https://doi.org/10.1016/j.cortex.2020.08.017.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). lme4: Linear Mixed-Effects Models Using Eigen and S4. R Package Version 1.1–7. Preprint at arXiv. 1406-5823. 2014.

Bell, A.J., and Sejnowski, T.J. (1995). An information-maximization approach to blind separation and blind deconvolution. Neural Comput. *7*, 1129–1159. https://doi.org/10.1162/neco.1995.7.6.1129.

Bellebaum, C., and Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. Eur. J. Neurosci. *27*, 1823–1835. https://doi.org/10.1111/j.1460-9568.2008.06138.x.

Bourgeois, A., Neveu, R., Bayle, D.J., and Vuilleumier, P. (2017). How does reward compete with goal-directed and stimulus-driven shifts of attention? Cogn. Emot. *31*, 109–118. https://doi.org/10.1080/02699931.2015.1085366.

Chase, H.W., Swainson, R., Durham, L., Benham, L., and Cools, R. (2011). Feedback-related Negativity codes prediction error but not behavioral adjustment during probabilistic reversal learning. J. Cogn. Neurosci. *23*, 936–946. https://doi.org/10.1162/jocn.2010.21456.

Chelazzi, L., E to inova, J., Calletti, R., Lo Gerfo, E., Sani, I., Della Libera, C., and Santandrea, E. (2014). Altering spatial priority maps via reward-based learning. J. Neurosci. *34*, 8594–8604. https://doi.org/10.1523/jneurosci.0277-14.2014.

Cohen, M.X., Elger, C.E., and Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. Neuroimage *35*, 968–978. https://doi.org/10.1016/j.neuroimage.2006.11.056.

Delgado, M.R., Miller, M.M., Inati, S., and Phelps, E.A. (2005). An fMRI study of reward- related probability learning. Neuroimage *24*, 862–873. https://doi.org/10.1016/j.neuroimage.2004.10.002.

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single- trial

EEG dynamics including independent component analysis. J. Neurosci. Methods *134*, 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

Donaldson, K.R., Ait Oumeziane, B., Hélie, S., and Foti, D. (2016). The temporal dynamics of reversal learning: P3 amplitude predicts valence-specific behavioral adjustment. Physiol. Behav. *161*, 24–32. https://doi.org/10.1016/j.physbeh.2016.03.034.

Fitzgerald, T.H., Friston, K.J., and Dolan, R.J. (2013). Characterising reward outcome signals in sensory cortex. Neuroimage *83*, 329–334. https://doi.org/10.1016/j.neuroimage.2013.06.061.

Folstein, J.R., Palmeri, T.J., and Gauthier, I. (2013). Category learning increases discriminability of relevant object dimensions in visual cortex. Cereb. Cortex *23*, 814–823. https://doi.org/10.1093/cercor/bhs067.

Glazer, J.E., Kelley, N.J., Pornpattananangkul, N., Mittal, V.A., and Nusslock, R. (2018). Beyond the FRN: broadening the time-course of EEG and ERP components implicated in reward processing. Int. J. Psychophysiol *132*, 184–202. https://doi.org/10.1016/j.ijpsycho.2018.02.002.

Goldman, R.I., Stern, J.M., Engel, J., Jr., and Cohen, M.S. (2002). Simultaneous EEG and fMRI of the alpha rhythm. Neuroreport *13*, 2487–2492. https://doi.org/10.1097/00001756-200212200-00022.

Hajcak, G., Dunning, J.P., and Foti, D. (2009). Motivated and controlled attention to emotion: time-course of the late positive potential. Clin. Neurophysiol. *120*, 505–510. https://doi.org/10.1016/j.clinph.2008.11.028.

Hajcak, G., Moser, J.S., Holroyd, C.B., and Simons, R.F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. Biol. Psychol. *71*, 148–154. https://doi.org/10.1016/j.biopsycho.2005.04.001.

Henschke, J.U., Dylda, E., Katsanevaki, D., Dupuy, N., Amvrosiadis, T., Currie, S.P., Pakan, J.M.P., and Rochefort, N.L. (2020). Reward association enhances stimulus-specific representations in primary visual cortex. Curr. Biol. *30*, 1866–1880.e5. https://doi.org/10.1016/j.cub.2020.03.018.

Heuer, A., Wolf, C., Schütz, A.C., and Schubö, A. (2017). The necessity to choose causes reward-related anticipatory biasing: parieto-occipital alpha-band oscillations reveal suppression of low-value targets. Sci. Rep. *7*, 14318–14411. https://doi.org/10.1038/s41598-017-14742-w.

Hickey, C., Chelazzi, L., and Theeuwes, J. (2010). Reward changes salience in human vision via the anterior cingulate. J. Neurosci. *30*, 11096–11103. https://doi.org/10.1523/jneurosci.1026-10.2010.

Holroyd, C.B., and Coles, M.G.H.H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol. Rev. *109*, 679–709. https://doi.org/10.1037/0033-295x.109.4.679.

Hopfinger, J.B., Buonocore, M.H., and Mangun, G.R. (2000). The neural mechanisms of top-down attentional control. Nat. Neurosci. *3*, 284–291. https://doi.org/10.1038/72999.

Jurjut, O., Georgieva, P., Busse, L., and Katzner, S. (2017). Learning enhances sensory processing in mouse V1 before improving behavior. J. Neurosci. *37*, 6460–6474. https://doi.org/10.1523/jneurosci.3485-16.2017.

Kiss, M., Driver, J., and Eimer, M. (2009). Reward priority of visual target singletons modulates event-related potential signatures of attentional selection. Psychol. Sci. *20*, 245–251. https://doi.org/10.1111/j.1467-9280.2009.02281.x.

Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). lmerTest package: tests in linear mixed effects models. J. Stat. Softw. *82*, 1–26. https://doi.org/10.18637/jss.v082.i13.

Lenth, R.V. (2021). Emmeans: Estimated Marginal Means, Aka Least-Squares Means. R Package Version 1.5.5-1. https://CRAN.R-project.org/package=emmeans.

Luck, S.J., and Hillyard, S.A. (1994). Spatial filtering during visual search: evidence from human electrophysiology. J. Exp. Psychol. Hum. Percept. Perform. *20* (5), 1000–1014. https://doi.org/10.1037/0096-1523.20.5.1000.

MacLean, M.H., and Giesbrecht, B. (2015). Neural evidence reveals the rapid effects of reward history on selective attention. Brain Res. *1606*, 86–94. https://doi.org/10.1016/j.brainres.2015.02.016.

Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. J. Neurosci. Methods *164*, 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024.

Miltner, W.H.R., Braun, C.H., and Coles, M.G.H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. J. Cogn. Neurosci. *9*, 788–798. https://doi.org/10.1162/jocn.1997.9.6.788.

Mine, C., Yokoyama, T., and Takeda, Y. (2021). Awareness is necessary for attentional biases by location–reward association. Atten. Percept. Psychophys. *83*, 2002–2016.

Muller-Gass, A., Duncan, M., Tavakoli, P., and Campbell, K. (2019). Individual differences in feedback processing affect perceptual learning. Personal. Individ. Differ. *143*, 145–154. https://doi.org/10.1016/j.paid.2019.01.017.

Nieuwenhuis, S., Holroyd, C.B., Mol, N., and Coles, M.G.H. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. Neurosci. Biobehav. Rev. *28*, 441–448. https://doi.org/10.1016/j.neubiorev.2004.05.003.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. *2011*, 156869.

Paller, K.A., Kutas, M., and McIsaac, H.K. (1995). Monitoring conscious recollection via the electrical activity of the brain. Psychol. Sci. *6*, 107–111. https://doi.org/10.1111/j.1467-9280.1995.tb00315.x.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. *118*, 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019.

Pornpattananangkul, N., and Nusslock, R. (2015). Motivated to win: relationship between anticipatory and outcome reward-related neural activity. Brain Cogn. *100*, 21–40. https://doi.org/10.1016/j.bandc.2015.09.002.

R Core Team (2020). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing). https://www.R-project.org/.

Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E., and Behrens, T.E. (2011). Frontal cortex and reward-guided learning and decision-making. Neuron *70*, 1054–1069. https://doi.org/10.1016/j.neuron.2011.05.014.

Sali, A.W., Anderson, B.A., and Yantis, S. (2014). The role of reward prediction in the control of attention. J. Exp. Psychol. Hum. Percept. Perform. *40* (4), 1654–1664. https://doi.org/10.1037/a0037267.

Sambrook, T.D., and Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. Psychol. Bull. *141*, 213–235. https://doi.org/10.1037/bul0000006.

San Martín, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. Front. Hum. Neurosci. *6*, 304. https://doi.org/10.3389/fnhum.2012.00304.

San Martín, R., Appelbaum, L.G., Pearson, J.M., Huettel, S.A., and Woldorff, M.G. (2013). Rapid brain responses independently predict gain maximization and loss minimization during economic decision making. J. Neurosci. *33*, 7011–7019. https://doi.org/10.1523/jneurosci.4242-12.2013.

Scheeringa, R., Mazaheri, A., Bojak, I., Norris, D.G., and Kleinschmidt, A. (2011). Modulation of visually evoked cortical FMRI responses by phase of ongoing occipital alpha oscillations. J. Neurosci. *31*, 3813–3820. https://doi.org/10.1523/jneurosci.4697-10.2011.

Schiffer, A., Muller, T., Yeung, N., and Waszak, F. (2014). Reward activates stimulus- specific and task-dependent representations in visual association cortices. J. Neurosci. *34*, 15610–15620. https://doi.org/10.1523/jneurosci.1640-14.2014.

Schneider, D., Göddertz, A., Haase, H., Hickey, C., and Wascher, E. (2019). Hemispheric asymmetries in EEG alpha oscillations indicate active inhibition during attentional orienting within working memory. Behav. Brain Res. *359*, 38–46. https://doi.org/10.1016/j.bbr.2018.10.020.

Schultz, W. (2015). Neuronal reward and decision signals: from theories to data. Physiol. Rev. *95*, 853–951. https://doi.org/10.1152/physrev.00023.2014.

Schultz, W. (2016). Dopamine reward prediction error coding. Dialogues Clin. Neurosci. *18*, 23–32. https://doi.org/10.31887/dcns.2016.18.1/wschultz.

Schwarz, G. (1978). Estimating the dimension of a model. Ann. Stat. *6*, 461–464. https://doi.org/10.1214/aos/1176344136.

Sisk, C.A., Remington, R.W., and Jiang, Y.V. (2020). A spatial bias toward highly rewarded locations is associated with awareness. J. Exp. Psychol. Learn. Mem. Cogn. *46*, 669–683. https://doi.org/10.1037/xlm0000749.

Speed, A., Del Rosario, J., Mikail, N., and Haider, B. (2020). Spatial attention enhances network, cellular, and subthreshold responses in mouse visual cortex. Nat. Commun. *11*, 505. https://doi.org/10.1038/s41467-020-14355-4.

Tankelevitch, L., Spaak, E., Rushworth, M.F., and Stokes, M.G. (2020). Previously reward-associated stimuli capture spatial attention in the absence of changes in the corresponding sensory representations as measured with MEG. J. Neurosci. *40*, 5033–5050. https://doi.org/10.1523/jneurosci.1172-19.2020.

Trimber, E.M., and Luhmann, C.C. (2017). Implicit predictions of future rewards and their electrophysiological correlates. Behav. Brain Res.

333, 184–191. https://doi.org/10.1016/j.bbr.2017.07.005.

van den Berg, B., Appelbaum, L.G., Clark, K., Lorist, M.M., and Woldorff, M.G. (2016). Visual search performance is predicted by both prestimulus and poststimulus electrical brain activity. Sci. Rep. *6*, 37718. https://doi.org/10.1038/srep37718.

van den Berg, B., Geib, B.R., San Martín, R., and Woldorff, M.G. (2019). A key role for stimulus-specific updating of the sensory cortices in the learning of stimulus–reward associations. Social Cogn. Affect. Neurosci. *14*, 173–187. https://doi.org/10.1093/scan/nsy116.

von Borries, A.K.L., Verkes, R.J., Bulten, B.H., Cools, R., and De Bruijn, E.R.A. (2013). Feedback-related negativity codes outcome valence, but not outcome expectancy, during reversal learning. Cogn. Affect. Behav. Neurosci. *13*, 737–746. https://doi.org/10.3758/s13415-013-0150-1.

Weil, R.S., Furl, N., Ruff, C.C., Symmonds, M., Flandin, G., Dolan, R.J., Driver, J., and Rees, G. (2010). Rewarding feedback after correct visual discriminations has both general and specific influences on visual cortex. J. Neurophysiol. *104*, 1746–1757. https://doi.org/10.1152/jn.00870.2009.

Won, B.Y., and Leber, A.B. (2018). Failure to exploit learned spatial value information during visual search. Vis. Cogn. *26*, 482–499. https://doi.org/10.1080/13506285.2018.1500502.

Worden, M.S., Foxe, J.J., Wang, N., and Simpson, G.V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific α-band electroencephalography increases over occipital cortex. J. Neurosci. *30*, RC63.

Zhao, C., Guo, J., Li, D., Tao, Y., Ding, Y., Liu, H., and Song, Y. (2019). Anticipatory alpha oscillation predicts attentional selection and hemodynamic response. Hum. Brain Mapp. *40*, 3606–3619. https://doi.org/10.1002/hbm.24619.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Experimental Script and Analysis Code	This paper	OSF: https://doi.org/10.17605/OSF.IO/NBX9P
Software and algorithms		
Presentation	Neurobehavioral Systems, Inc., Ablany, CA	www.neurobs.com
Fieldtrip (MATLAB package)	Oostenveld et al. (2011)	http://fieldtriptoolbox.org
EEGlab (MATLAB package)	Delorme and Makeig (2004)	https://sccn.ucsd.edu/eeglab/index.php
Infomax ICA algorithm	Bell and Sejnowsky (1995)	Within the Fieldtrip toolbox
R	R Core Team (2020)	https://cran.r-project.org/bin/windows/base/
Lme4	Bates et al. (2015)	https://cran.r-project.org/web/packages/lme4/index.html
Emmeans	Lenth (2021)	https://cran.r-project.org/web/packages/emmeans/index.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for data and code generated for this study should be directed to and will be fulfilled by the lead contact, Monicque Lorist (m.m.lorist@rug.nl).

Materials availability

This study does not generate new unique reagents.

Data and code availability

- The data cannot be publicly shared because the data is considered to be personal data under the GDPR, and cannot be further anonymized. The research participants have not given their consent for sharing data beyond the original research purpose, which does include data access for independent verification of the results, but prohibits public sharing. These limitations are imposed by European law and verified by the data office of the faculty of Behavioral and Social Sciences at the University of Groningen. Data access can be requested via the data office of the faculty of Behavioral and Social Sciences of the University of Groningen: research-data-bss@rug.nl. Please quote study code PSY-1819-S-0157 in the request.

- Original code has been deposited at Open Science Framework and is publicly available as of the date of publication (Open Science Framework: https://doi.org/10.17605/OSF.IO/NBX9P).

- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAIL

Participants

Thirty-six adults participated in the experiment, of which two had to be excluded because of misunderstandings in the task instruction, and another three because of too many EEG artefacts (>30% of artefact trials). In addition, two subjects misunderstood the task instructions during the first two learning sets, and their data were nevertheless included in the analyses after removing the invalid learning sets from the dataset (<11% of trials and <23% of trials removed, respectively). Thus, the final sample comprised 31 adults (M_{age} = 24.2, SD = 3.6), of which 23 were female and 29 were right-handed. All participants reported normal or corrected-to-normal vision. Four subjects participated for course credit, whereas 27 participated received monetary compensation (7 €/hour). In addition, participants received a bonus that depended on their task performance. Written consent was obtained from all participants before the start

of the experimental session. All study protocols were approved by the Ethical Committee of Psychology of the University of Groningen.

METHOD DETAILS

Apparatus

Participants performed the experiment in a dimly lit, sound-attenuated experimental room. The task was created and presented using the software package 'Presentation' (Version 20.3, Neurobehavioral Systems, Inc., Albany, CA, www.neurobs.com) on a 27-inch monitor with a refresh-rate of 60 Hz. Participants made their responses using a gamepad (Logitech Rumblepad).

Procedure

After application of the EEG cap, participants were seated in a chair at approximately 60 cm distance from the monitor and instructed to keep their eyes on the fixation point at the center of the screen. In addition, participants were asked to limit movement during the experiment. They subsequently performed a probabilistic reward-based spatial learning task, in which they had to find a location stimulus with the highest associated gain probability using trial-by-trial gain/loss feedback as a guide, in order to maximize their monetary gains. On each trial they also had to bet either high or low on the to-be-chosen location. If participants placed a high bet and received gain feedback on that trial, the monetary reward was 90 eurocents, whereas high bets and loss feedback yielded a monetary loss of 90 eurocents. Low bets in combination with gain feedback yielded a monetary gain of 30 eurocents, and low bets followed by loss feedback led to a monetary loss of 30 eurocents. In case of no responses to the bet, participants still received valid feedback to the stimulus response. If participants did not select a location stimulus, they immediately were presented with loss feedback. An interim summary after each learning set of 40 trials provided subjects with the sum of money they had earned so far. Participants were paid out at the end of the experiment with a maximum of 15 €, the exact amount depending on their monetary gain during the experiment.

Probabilistic spatial learning task

The probabilistic reward-based spatial learning task consisted of nine learning sets, of 40 trials each, spread across three blocks, totaling to 360 trials for the entire experiment. Within each learning set, four locations (defined as four quadrants on the monitor screen) had a probability of monetary gain assigned to them. These four gain probabilities were predefined as 0.3, 0.45, 0.55, and 0.7, and were randomly assigned to the four locations at the start of each learning set.

Each trial started with a bet screen, which was shown until a response was made or 2000 ms elapsed. The bet screen consisted of a central fixation point with the letters "H" (meaning high bet) and "L" (meaning low bet) placed 50 pixels above or below the fixation. The letters randomly switched position on each trial, and participants had to press the upper left bumper button of the gamepad to choose the upper letter or the lower left bumper button to choose the lower letter. Following a response, the chosen bet option was highlighted for 300 ms. In case the participant made no response, a screen with the letters "no response!" was shown for 300 ms instead.

The bet screens were followed by a fixation point which remained onscreen between 700 and 900 ms. Next, the location stimulus screen appeared, in which four circle stimuli containing a gap were presented. The circles were 90 by 90 pixels wide and placed in each quadrant of the screen, at a distance of 50 pixels from the fixation point. The four circles differed with regard to the location of the gap (up, down, left, right), which was linked to a button response using the right button set on the gamepad. The buttons were arranged in the way that there was an upper, lower, left, and right button corresponding to the upper, lower, left, and right opening of the circle. Participants were instructed to choose a location by pressing the button that corresponded to the opening of the circle that was placed in the quadrant they wanted to choose on that trial. The position arrangement of the opening of the circles changed randomly on each trial. As a result, chosen stimulus location and button presses were unrelated.

The location stimulus screen was presented for 2000 ms, followed by a choice screen highlighting the location stimulus that was chosen (300 ms). After another central fixation screen (the duration again randomly ranging between 700 and 900 ms), feedback was presented for 500 ms. The words "gain!" or "loss!" occurred in black font on either an orange or blue background. The combination of feedback valence

and background color was counterbalanced across participants, that is half of the participants received gain feedback on a blue background and loss feedback on an orange background and the other half vice versa. The trial ended with a third fixation screen, with a duration randomly chosen between 1700 and 2000 ms. Participants were allowed to take short, self-timed breaks in between experimental blocks.

EEG recording and pre-processing

EEG was recorded using a 64-channel, ANT/Duke-layout, equidistant electrode cap and a DC amplifier at a sampling rate of 512 Hz and referenced to a central electrode channel (channel 5Z, corresponding to channel Cz in the 10–20 layout system). Four subjects were measured using a sampling rate of 500 Hz, and data of these subjects were up-sampled to 512 Hz during pre-processing. Pre-processing was done using the MATLAB EEGlab (Delorme and Makeig, 2004) and FieldTrip (Oostenveld et al., 2011, http://fieldtriptoolbox.org) toolboxes. The recorded data was re-referenced to the average reference and filtered offline using a non-causal bandpass filter (Hamming windowed finite response filter, low cut-off value of 0.01 Hz (-6 dB at 0.005), high cut-off value of 30 Hz (-6 dB at 33.75 Hz). For the time-frequency analyses, we only applied the highpass filter (same settings as for the ERP analysis data) in order to preserve frequency information at higher frequencies. To correct for eyeblinks and horizontal eye movements, we performed an independent component analysis using the algorithm implemented in EEGlab (infomax ICA algorithm, Bell and Sejnowsky (1995)) and reconstructed the EEG data without components that reflected eyeblinks or horizontal eye movements (a maximum of three components per subject – $M = 2.2$ components). In addition, we excluded trials that exceeded an amplitude threshold of $\pm 120\ \mu V$ per stimulus-locked and feedback-locked ERP interval, and excluded subjects that had more than 30% of the trials removed. We then extracted epochs from −500 ms prior to the onset of the stimulus presentation screen to 2000 ms post-stimulus presentation screen, and baselined the data to the mean voltage observed between − 200 and 0 ms pre-stimulus. Lastly, we extracted epochs spanning − 500 msbefore the onset of the feedback screen to 2000 ms after the onset of the feedback screen, again baselined the data to the mean voltage between − 200 and 0 ms before the onset of the feedback screen.

ERP analysis and data extraction

In line with previous work (van den Berg et al., 2019), our regions of interest (ROIs) for the N2pc analyses were left and right occipital electrode sites (corresponding to PO7, O1, O2, PO8 of the 10–20 system). We determined the time window of interest (TOI) by creating stimulus-locked grand averages and calculating the contralateral-minus-ipsilateral difference in amplitudes relative to the location of choice over all conditions and selecting a time window ranging from 50 ms before and after the grand average peak, resulting in a TOI of 220–320 ms. We then averaged amplitude values in the TOI over the left and right ROIs separately and calculated the contralateral-minus-ipsilateral amplitudes by subtracting amplitudes of ipsilateral-to-choice channel sites from contralateral ones, and then collapsing over left- and right-choice conditions (Luck and Hillyard, 1994).

We chose a fronto-central region of interest (ROI) for the FRN (corresponding to Fz, FCz, FC2, and FC1 in the 10–20 system, based on previous work; van den Berg et al., 2019). The latency interval was determined by creating a feedback-locked grand averages over trials and selecting a time-window corresponding to 50 ms before and after the grand average peak, resulting in an interval between 240 and 340 ms post-feedback onset. Our LPC ROIs were parieto-central channels, approximating Pz, CPz, P1, and P2 in the 10–20 system, and we selected amplitude values between 400 and 800 ms post-feedback onset. LPC ROIs and TOIs were derived from the literature (Glazer et al., 2018; Pornpattananangkul and Nusslock, 2015).

Time-frequency analysis and data extraction

Time-frequency analyses of alpha band power were done using the FieldTrip MATLAB toolbox (Oostenveld et al., 2011, http://fieldtriptoolbox.org). Frequency decomposition was performed on average-referenced EEG data time-locked to the presentation of feedback using a multitaper method based on discrete prolate spheroidal sequences. We performed the analysis on linearly spaced frequencies from 1 to 40 Hz with steps of 0.5 Hz, using taper window widths that increased by one cycle every 3 Hz, starting at 3 cycles for 1 to 3.5 Hz and ending with 15 cycles for 37–40 Hz. Window smoothing was defined as frequency multiplied by 0.4, leading to less frequency smoothing in lower frequencies compared to higher frequencies. A $\log_{10}$ conversion was performed on the resulting power spectra to correct for non-normality of power within each time-by-frequency bin. For the analyses examining the relationship between alpha power and the N2pc, we extracted average alpha band power (8–14 Hz) between 600 and 1200 ms post-feedback

presentation, from left and right occipital electrode channels (corresponding to PO7, O1, O2, PO8 in the 10–20 system) separately in order to analyze lateralized-to-choice alpha power. The ROIs for the analysis were chosen to be the same as for the N2pc analyses, and the TOI was estimated based on visual inspection of the data.

QUANTIFICATION AND STATISTICAL ANALYSIS
Trial exclusion and model selection

For our statistical analyses, we excluded no-response trials and trials in which RTs were faster than 200 ms (i.e., fast guesses). A no-response trial was defined as trials in which the bet or stimulus response was missing. Thus, for the behavioral analyses, an average of 356 trials (SD = 2.99) per subject remained. After removing trials marked with artefacts, an average of 338 trials (SD = 22.19) remained for the FRN and LPC analyses, and an average of 321 trials (SD = 26.84) remained for the N2pc and alpha power analyses.

In all analyses of behavioral and neural data, we applied a mixed-modelling approach using the lme4 statistical package (Bates et al., 2015), and for post-hoc analyses, the package emmeans (Lenth, 2021), within the software package *R* (R Core Team, 2020). All models reported here contained the random factor 'subject' as a varying intercept. We determined the inclusion of random slopes varying by subject via model comparison, by selecting the model with the lowest Akaike information criterion (AIC, Akaike, 1974). This process allows for selecting the best-fitted model by taking over- and underfitting into account. In case the AIC difference yielded inconclusive results, we also took the Bayesian information criterion (BIC, Schwarz, 1978) of each model into account during model section. All models that we used in our analyses are summarized in Table 1 and described in detail in the sections below. Statistical results are reported under 'Results' in the main body of the text. Degrees of freedom were approximated using Satterthwaite's method using the package lmerT-est (Kuznetsova et al., 2017). Statistical significance was defined at p < 0.05.

Behavioral analyses

We first tested whether participants showed a behavioral bias towards any of the locations by running a two-way ANOVA with the predictors 'learning set half' (with the two levels 'first half of trials' and 'second half of trials' of the learning set) and 'location' (with four levels corresponding to the four locations). The predictor was the probability of choosing a location stimulus when it was the assigned set winner.

In order to analyze participant's choice behavior with regard to the set winner, we fitted a logistic mixed effects model with the binary outcome variable 'choice: set winner' versus 'choice: not set winner' as the dependent variable. As fixed effect predictors, we used the variable 'trial number', ranging from 1 to 40, and the second-degree polynomial, to account for non-linearities in the data. The results of the behavioral model are tested with Wald's Chi-Square test. To analyse participant's betting behavior across the learning set, we fitted a generalized linear mixed model with the binary outcome variable 'high bet' versus 'low bet' and the predictor trial number, with varying intercept over subject and a varying slope of trial number over subject.

Initial observation of the RT data showed that responses made on the first trial of a set were substantially slower than all other trials ($M_{trial\#1}$ = 904 ms, $M_{trial\#2-40}$ = 696 ms, $t(10804)$ = 15.6, p < 0.0001), and we therefore excluded the first trial from our model analysis. Next, we fitted a mixed-effects model with the outcome variable 'Reaction Time', and the fixed effect predictors 'trial number', and 'choosing set winner' versus 'not choosing set winner', and an interaction between the two.

EEG analyses

In order to confirm the presence of a lateralized N2pc component, we applied a one-sided t-test on extracted N2pc amplitudes to test whether these differed significantly from zero. The FRN and LPC were analyzed by predicting their amplitudes with the fixed effects 'trial number' and 'feedback (gain and loss)', as well as the interaction between 'trial number' and 'feedback'.

Time-frequency analysis

In order to test whether alpha power was lateralized towards the upcoming location of choice, we performed a cluster-based permutation test (Maris and Oostenveld, 2007). First, we calculated the contralateral-minus-ipsilateral difference in power of all frequency bands included in our time-frequency analysis (i.e., between 1 and 40 Hz) relative to the next-trial choice per subject. Then, t-tests were performed on

the average lateralized power within the alpha frequency band (8–14 Hz) for our N2pc ROI, on each time-point between 600 and 2000 ms post-feedback presentation (the interval following P3 activity and preceding the next trial; replicating our previous work in van den Berg et al., 2019). If a test statistic reached significance at p = 0.05, that time point was included into a cluster made of multiple adjacent significant time points. Cluster statistics were obtained by summing all t-values within a cluster. Each cluster was then compared to a permutation distribution (created of 1000 iterations by randomly switching labels of conditions at subject level) and considered significant at p < 0.05. Model assumptions were tested via visual inspection.

For our model analysis of alpha power and N2pc, we first extracted alpha power from our ROIs and TOI and applied a t-test to confirm that the averaged lateralized alpha power was significantly different from zero. Then, we fitted a linear mixed model with the N2pc amplitude as the dependent variable, and lateralized alpha power and trial number as the predictors. Random slopes were allowed to vary per trial number oversubjects.