Perceptual Learning and Decision-Making in Human Medial Frontal Cortex

Thorsten Kahnt,[1,2,6,*] Marcus Grueschow,[1,3,4,6] Oliver Speck,[5] and John-Dylan Haynes[1,2,3,4,*]

[1]Bernstein Center for Computational Neuroscience, Charité–Universitätsmedizin Berlin, Philippstrasse 13/House 6, D-10115 Berlin, Germany
[2]Berlin School of Mind and Brain, Humboldt Universität zu Berlin, 10099 Berlin, Germany
[3]Max Planck Institute for Human Cognitive and Brain Sciences, 04303 Leipzig, Germany
[4]Department of Neurology, Otto-von-Guericke University Magdeburg, 39120 Magdeburg, Germany
[5]Faculty of Natural Sciences, Institute for Experimental Physics, Otto-von-Guericke University Magdeburg, 39120 Magdeburg, Germany
[6]These authors contributed equally to this work
*Correspondence: thorsten.kahnt@bccn-berlin.de (T.K.), haynes@bccn-berlin.de (J.-D.H.)
DOI 10.1016/j.neuron.2011.02.054

SUMMARY

The dominant view that perceptual learning is accompanied by changes in early sensory representations has recently been challenged. Here we tested the idea that perceptual learning can be accounted for by reinforcement learning involving changes in higher decision-making areas. We trained subjects on an orientation discrimination task involving feedback over 4 days, acquiring fMRI data on the first and last day. Behavioral improvements were well explained by a reinforcement learning model in which learning leads to enhanced readout of sensory information, thereby establishing noise-robust representations of decision variables. We find stimulus orientation encoded in early visual and higher cortical regions such as lateral parietal cortex and anterior cingulate cortex (ACC). However, only activity patterns in the ACC tracked changes in decision variables during learning. These results provide strong evidence for perceptual learning-related changes in higher order areas and suggest that perceptual and reward learning are based on a common neurobiological mechanism.

INTRODUCTION

For decades the dominant view in visual perceptual learning has been that performance improvements on visual tasks are accompanied by changes in early visual areas (Sasaki et al., 2010; Seitz and Watanabe, 2005). However, this assumption was mainly based on psychophysical data (Goldstone, 1998; Karni and Sagi, 1991) and received only inconsistent support from neural recording studies (Crist et al., 2001; Ghose et al., 2002; Schoups et al., 2001). Recent studies suggest that perceptual improvements might rather be associated with changes outside the early visual cortices (Zhang and Li, 2010; Zhang et al., 2010b). Specifically, perceptual learning is thought to be related to an enhanced readout of sensory information by higher cortical areas that are directly involved in decision-making

(Chowdhury and DeAngelis, 2008; Law and Gold, 2008; Li et al., 2004, 2009). This idea has recently been supported by single-unit recordings in primates. More specifically, it has been shown that performance improvements in motion-direction discrimination are accompanied by changes in responses of lateral intraparietal area (LIP), but not middle temporal area (MT) neurons (Law and Gold, 2008). Moreover, this pattern of results is predicted by a reinforcement learning model in which perceptual learning is established by changes in connectivity between visual and decision areas leading to altered representations in higher cortical areas (Law and Gold, 2009).

Similar to this proposed mechanism, reward-based learning and decision-making is also accompanied by activity changes in decision-making areas such as LIP (Platt and Glimcher, 1999; Sugrue et al., 2004), dorsolateral prefrontal cortex (DLPFC) (Barraclough et al., 2004; Pasupathy and Miller, 2005), and the anterior cingulate cortex (ACC) (Kennerley et al., 2006; Matsumoto et al., 2007). Especially the ACC has been shown to be involved in flexibly updating and representing the value of actions leading to reward (Behrens et al., 2007; Hayden et al., 2009). In principle, the role of sensory evidence in forming a perceptual choice could be treated in the same way as the role of action values in forming a reward-based decision (Gold and Shadlen, 2007). Consequently, neural circuits that update and represent action values in reward-based tasks might be equally suited to integrate sensory information in the context of perceptual decision-making. However, a direct engagement of human prefrontal cortex in perceptual learning has not been shown so far.

Here we used a model-based neuroimaging approach to test the idea that human perceptual learning and decision-making can be accounted for by a reinforcement learning process involving higher cortical areas. We trained subjects on an orientation discrimination task with explicit performance feedback over the course of 4 days. Functional magnetic resonance imaging (fMRI) data were acquired on the first and last day of training. Behavioral improvements were well explained by a reinforcement learning model for perceptual learning. Learning in this model leads to enhanced readout of sensory information, thereby establishing noise-robust representations of decision variables that form the basis for perceptual choices. By using multivariate information mapping techniques (Haynes and Rees, 2006; Kriegeskorte et al., 2006), we find sensory evidence
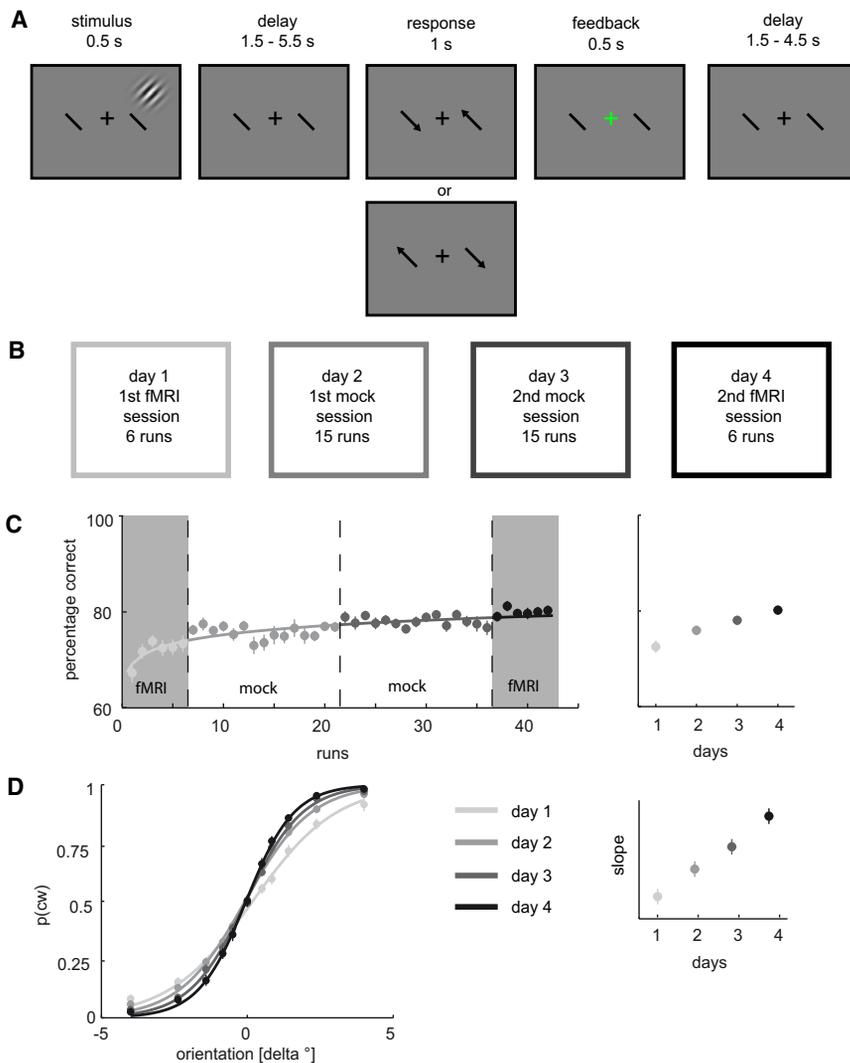
Figure 1. Experiment and Improvements in Perceptual Choices

(A) Sketch of the orientation discrimination task. Perceptual decisions and motor responses were dissociated by using a randomized response mapping screen 1.5–5.5 s after stimulus presentation. Feedback was provided by changing the color of the fixation cross for 500 ms to green for correct choices or red for incorrect choices.

(B) Training on the perceptual task took place on 4 days. During the first and the last day fMRI data were acquired. During the second and third day training took place in the environment of a mock scanner.

(C) Behavioral performance (percentage correct) as a function of training runs (left) and days (right). On the left, dashed vertical lines separate training days and gray shaded areas indicate training during fMRI data acquisition. Data from different training days are color coded from light gray (day 1) to black (day 4). Error bars = SEM for n = 20.

(D) Psychophysical functions are shown (left) relating stimulus orientation to the probability of a clockwise decision (p(cw)). Solid lines represent best-fitting sigmoidal function. Average slopes of the sigmoids are shown (right) as a function of training days. Error bars = SEM for n = 20.

encoded in early visual cortex as well as in higher order regions such as the putative LIP. However, only activity patterns in the ACC tracked changes in the model-derived decision variables during learning. Moreover, we find activity related to signed reward prediction errors, the teaching signal of the reinforcement learning model, in the ventral striatum and the same part of the ACC where learning-related changes were observed. These results provide strong evidence for perceptual learning-related changes in higher order brain regions. Furthermore, these results suggest that perceptual as well as reward learning and decision-making can be understood in the framework of reinforcement learning and that both forms of learning are based on a common neurobiological mechanism.

RESULTS

Improvements in Perceptual Decision-Making during Learning

During the course of 4 days 20 subjects (11 male, mean age ± SEM, 26.3 ± 0.74) participated in an orientation discrimination task involving explicit performance feedback (Figure 1A). In each trial subjects saw a low contrast Gabor in the right upper visual field for 500 ms while fixating on a central fixation cross. The orientation of the Gabor could deviate from 45° in both directions (counterclockwise and clockwise). Subjects were asked to indicate the perceived orientation (tilted toward counterclockwise versus tilted toward clockwise) on a response mapping screen. After the response, the fixation cross turned green given a correct decision or red given an erroneous response. Days 1 and 4 each involved six runs (110 trials each) of training while BOLD signals were acquired by using fMRI (Figure 1B). Days 2 and 3 each involved 15 behavioral training runs in a mock scanner.

Performance on the task (percentage of correct decisions) increased with training, demonstrating a robust effect of perceptual learning (Figure 1C). A one-way ANOVA with repeated measures on percentage correct revealed a significant main effect of run (F(41,779) = 6.49, p < 0.001). Furthermore, a more parsimonious one-way ANOVA comparing performance between training days revealed a significant effect of day (F(3,57) = 20.70, p < 0.001) with significant differences between all days (p < 0.05, Bonferroni corrected, Figure 1C, right).

Learning involved a steepening of the psychophysical function relating the stimulus to the perceptual decision (Figure 1D), i.e., subjects became increasingly sensitive to small deviations from 45°. To quantify this improvement in orientation discrimination, we fitted a sigmoidal function to the psychophysical data of
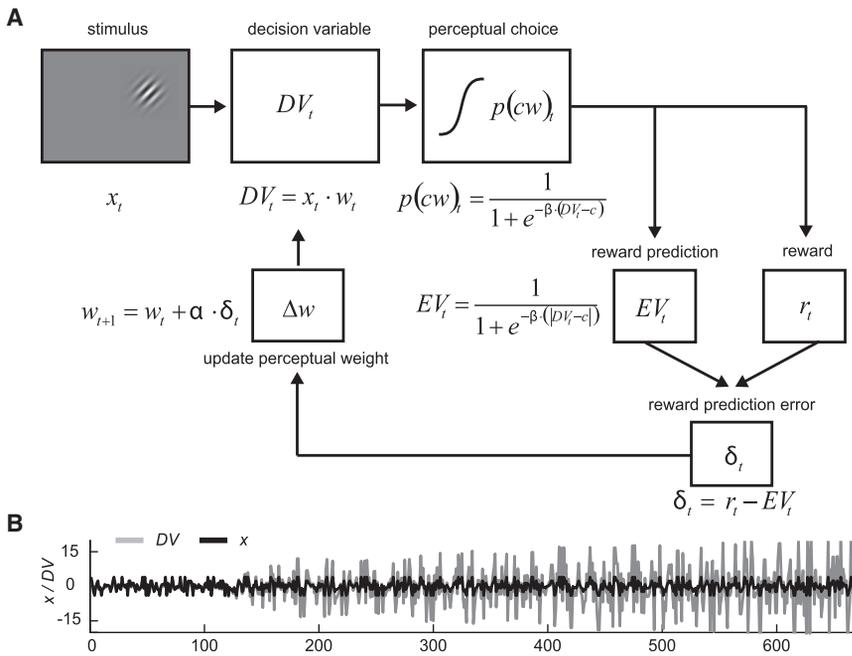
A

B

Figure 2. Reinforcement Learning Model for Perceptual Decision-Making

(A) Perceptual decisions are based probabilistically on a decision variable *DV* that is the sensory information *x* (orientation - 45°) scaled by a perceptual weight *w*. The more positive *DV* the more likely is a clockwise decision, and the more negative *DV* the more likely is a counterclockwise choice.

(B) Example time course of trial-wise values of *DV* derived from the reinforcement learning model (gray) for one subject on the first training day (six runs = 660 trials). For comparison the physical stimulus orientation is plotted on the same scale (black). It can be seen that learning is implemented by scaling the sensory evidence which thus becomes more robust to noise.

each subject and each day (Figure 1D, right). A one-way ANOVA with repeated measures on the slopes of this function revealed a significant main effect of day ($F_{(3,57)} = 31.97$, $p < 0.001$). Post hoc t test confirmed that the slope increased with every training day ($p < 0.05$, Bonferroni corrected). Taken together, these results provide strong evidence for improvements in perceptual decision-making over the course of learning.

A Reinforcement Learning Model for Perceptual Decision-Making

To account for improvements in perceptual decision-making during learning we designed a reinforcement learning model for perceptual decision-making (see Experimental Procedures and Figure 2A). A similar model has recently been applied to monkey behavioral and electrophysiological data (Law and Gold, 2009). In brief, the model makes perceptual choices *p(cw)* on the basis of a decision variable *DV*. Negative values of *DV* lead to counterclockwise decisions, whereas positive values of *DV* lead to clockwise decisions. The decision variable is computed as the product of the sensory stimulus *x* (stimulus orientation minus 45°) and a perceptual weight *w* accounting for the ability to read out sensory information provided by the stimulus *x*. Thus, the perceptual weight scales the stimulus representation; low values of *w* lead to small absolute values of *DV,* i.e., unreliable stimulus representations in the presence of noise, whereas high values of *w* lead to large absolute values of *DV*, i.e., noise-robust stimulus representations (Figure 2B). In essence, perceptual learning involves updating the perceptual weight by means of an error-driven reinforcement learning mechanism (i.e., Rescorla-Wagner updating). Specifically, *DV* forms not only the basis for the perceptual decision, but the absolute value of *DV* also provides the probability that the current trial will be rewarded (expected value *EV*). This expected value is then compared with the actual reward *r*, resulting in

a reward prediction error δ that is in turn used to update the perceptual weight in proportion to a learning rate α. Learning thus leads to an amplified representation of stimulus information that can be used to guide perceptual choices. It is important to note that the individual noise level is implicitly modeled as the slope of the sigmoidal function relating a given value of *DV* to the probability of a clockwise decision. The learning rate α and the other free model parameters were estimated for each subject individually (see Experimental Procedures).

Comparison of the Model and Behavioral Data

The estimated model parameters and the individual sequences of stimuli, choices, and feedback were used to construct decision variables for each subject (see Figure 2B for an example). In the following analyses we compare the behavior of the model with the behavior of the subjects to assess how well the model can characterize subjects' perceptual choices and perceptual improvements over the course of training.

Model performance was computed by using the probability of a correct decision, $p(correct) = p(cw) \cdot \kappa + (1 - p(cw)) \cdot (1 - \kappa)$, where $\kappa = 1$ if $x \geq 0$ and $\kappa = 0$ if $x < 0$. Similar to subjects' choice behavior, model performance improved with training (Figure 3A). A one-way ANOVA with repeated measures revealed a significant main effect of runs ($F_{(41,779)} = 19.89$, $p < 0.001$). Additionally, a one-way ANOVA on performance over training days revealed a significant main effect of day ($F_{(3,57)} = 36.53$, $p < 0.001$) with significant differences between all days ($p < 0.05$, one-tailed, Bonferroni corrected). We found a significant relationship ($r = 0.81$, $p < 0.001$) between the performance of models and subjects across individual runs (Figure 3B).

Psychophysical functions were estimated from the decision-making behavior of the model. Similar to subjects' behavior, learning was accompanied by a steepening of the psychophysical function (Figure 3C). The slope of the function changed significantly over the 4 training days ($F_{(3,57)} = 45.20$, $p < 0.001$, Figure 3C, inset). Post hoc t test revealed that the slope increased with every day of training ($p < 0.05$, one-tailed, Bonferroni corrected). Figure 3D depicts the relationship between the
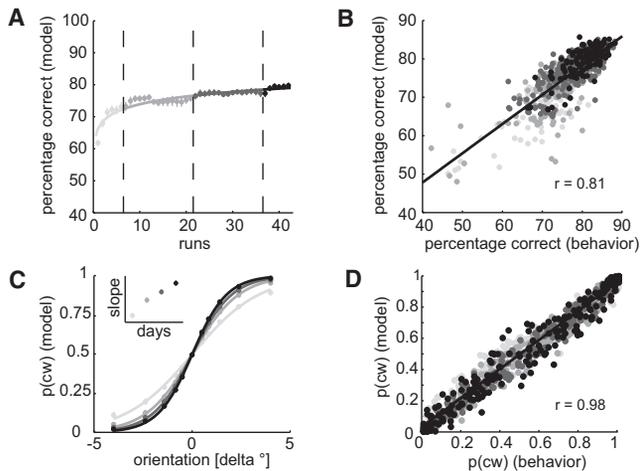
Figure 3. Comparison of the Model and Behavior

(A) Performance (percentage correct) improvements of model behavior with training. Error bars = SEM for n = 20.

(B) Scatterplot depicts the relationship between the run-wise performance of individual subjects and models.

(C) Psychophysical functions of model behavior for different training days. Inset depicts slope of the psychophysical functions across days. Error bars = SEM for n = 20.

(D) Scatterplot depicts the relationship between the day-wise behavior (*p(cw)*) for different orientations of individual subjects and models.

model's and subjects' psychophysical function. Both *p(cw)* values were highly correlated (r = 0.98, p < 0.001) across individual training days and orientations. Also the slopes of the psychophysical functions of the model and the subjects were highly correlated across individual training days (r = 0.97, p < 0.001). Taken together these results demonstrate that the reinforcement learning model accounted very well for subjects' perceptual improvements over training.

Neural Representation of Stimulus Orientation and Decision Variable

Having established the reinforcement learning model that accounts for perceptual learning and decision-making we proceeded to investigate the underlying neural mechanism. In a first step we identified brain regions that encode objective sensory evidence, that is, the orientation of the Gabor patch. Specifically, we used linear support vector regression (SVR) in combination with a searchlight approach (radius = 4 voxels) that allows information mapping without potentially biasing prior voxel selection (Haynes et al., 2007; Kahnt et al., 2010; Kriegeskorte et al., 2006). We used a leave-one-out cross-validation procedure by training the regression model on one part of the data (11 scanning runs) and predicted the orientation of the stimuli in the 12th scanning run. This was repeated 12 times, each time by using a different run as the independent test data set. Information about the orientation was defined as the average Fisher's z-transformed correlation coefficient between the orientation predicted by the SVR model and the actual orientation in the independent test data set (Kahnt et al., 2011).

During stimulus presentation orientation was significantly encoded (p < 0.0001, k = 20, corrected for multiple comparisons

at the cluster level, p < 0.001) in activity patterns in the lower left early visual cortex (BA 17, MNI coordinates [−12, −87, 0], t = 6.31, Figure 4A), the left lateral parietal cortex (putative lateral intraparietal area, LIP, BA 7 [−24, −69, 57], t = 6.01, Figure 4C), the precuneus (BA 23 [−3, −36, 36], t = 6.26), and the medial frontal gyrus (BA 9 [0, 48, 30], t = 6.75) (see Figure S1 and Table S1, available online, for complete results). Activity patterns in these regions can be used as a spatial filter to make linear predictions about the orientation of the Gabor (Figures 4A and 4C, right). In Figures 4B and 4D the idiosyncratic patterns of two subjects in the early visual cortex and the LIP are plotted along with their orientation predictions, respectively. It can be seen that these predictive patterns consist of small subregions in which activity increases and decreases with larger angles. Specifically, some voxels have higher responses for orientations >45° (yellow), whereas other voxels show higher responses for orientations <45° (blue).

We compared our multivariate results to a more conventional univariate whole-brain analysis searching for correlations between stimulus orientation and the BOLD signal in each voxel by using a parametric approach (Büchel et al., 1998). This analysis did not reveal any significant voxels (p < 0.0001, uncorrected, k = 5). Furthermore, a region of interest (ROI) analysis at a more liberal threshold of p < 0.05 revealed no univariate correlations with stimulus orientation in the early visual cortex (t = 1.29, p = 0.21), the lateral parietal cortex (t = 1.34, p = 0.20), and the medial frontal gyrus (t = 0.56, p = 0.58) as identified by our multivariate analysis (see above). This suggests that the results of the multivariate analysis are above and beyond what could have been obtained through univariate approaches.

Our results so far suggest that information about the physical properties of the stimulus, i.e., its orientation, is encoded in the early visual cortex as well as in higher brain regions such as the putative LIP. However, our model suggests that the orientation of the Gabor is not used directly to make the perceptual decision. What is used to make the choice is the decision variable *DV*. Thus, activity patterns in brain regions that are directly involved in perceptual decision-making should correlate with *DV*. We identified such brain regions by applying the same local information mapping procedure described above, but this time searching for representations of *DV* rather than orientation. We found significant information (p < 0.0001, k = 20, corrected for multiple comparisons at the cluster level, p < 0.001) about the model-derived decision variable in the left putative LIP (BA 7 [−24, −63, 48], t = 5.98, Figure 5A), the ACC (BA 32 [−3, 45, 24], t = 9.01, Figure 5C) and the precuneus (BA 23 [0, −39, 39], t = 6.57) but not the early visual cortex (see Figure S2 and Table S2 for complete results). In these regions distributed patterns of activity can be used to make linear predictions about the decision variable derived from the reinforcement learning model (Figures 5A and 5C, right). Again, a univariate whole-brain analysis searching for correlations with *DV* revealed no significant voxels (p < 0.0001, uncorrected, k = 5). Furthermore, an ROI analysis revealed no significant (p < 0.05) univariate correlations with *DV* in the lateral parietal cortex (t = 0.64, p = 0.53) or the ACC (t = 0.75, p = 0.46) as identified by our multivariate analysis (see above).
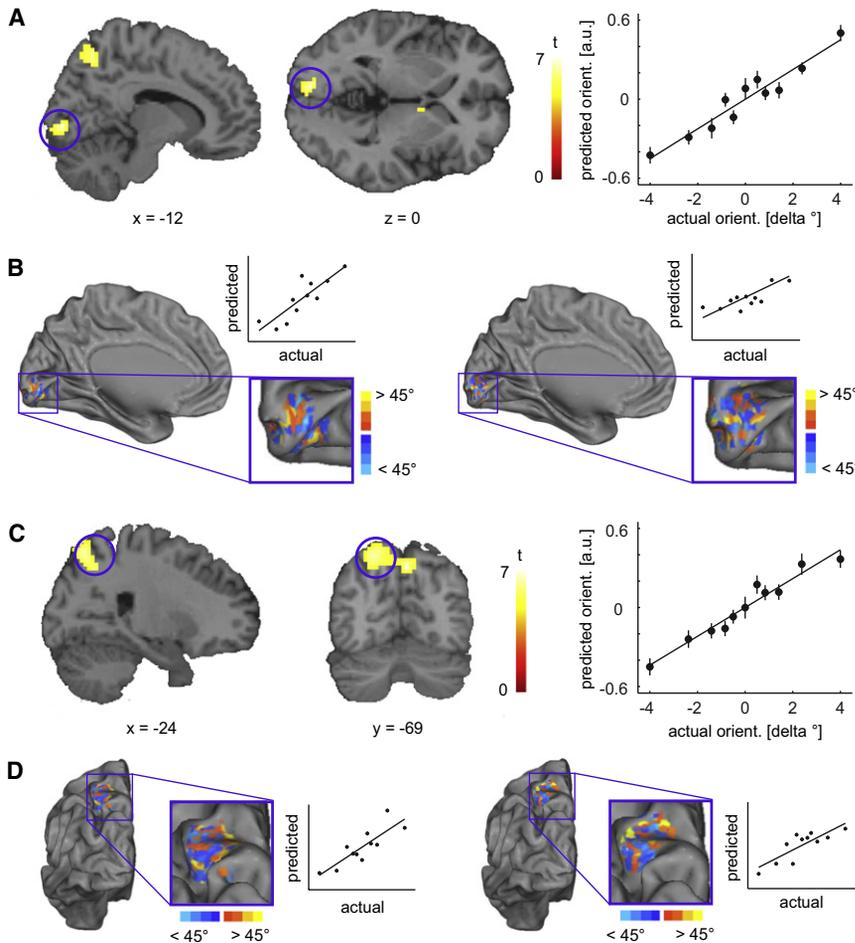
Figure 4. Encoding of Stimulus Orientation

(A) Information about stimulus orientation is encoded in activity patterns in the early visual cortex. T-maps on prediction accuracy (Fisher's z-transformed correlation coefficients) are thresholded at p < 0.0001, k = 20. For illustration purposes, scatterplot (right) visualizes the relationship between actual orientations and the orientation predicted by the SVR model (average of normalized predictions across cross-validation steps and subjects). Error bars = SEM for n = 20.

(B) Example of individual predictive maps (individual searchlight with peak prediction accuracy within the early visual cortex) from two subjects together with their orientation predictions. These idiosyncratic maps have subareas in which activity is differently correlated with stimulus orientation. Yellow indicates areas where activity increases with orientations >45°, whereas in blue areas activity increases with orientations <45°. These maps can be understood as optimal spatial filters to predict the stimulus orientation based on the activity in this region.

(C and D) Same as in (A) and (B) but for information about stimulus orientation in the lateral parietal cortex.

Perceptual Learning-Related Information in ACC

The physical stimulus orientation is correlated with the decision variable (*DV*) that is used by the model to make perceptual choices. This makes it difficult to decisively dissociate regions involved in representing sensory evidence from such regions involved in perceptual learning and decision-making. However, the decision variable used by the model changes over the course of learning and encoding in regions involved in perceptual learning should thus follow *DV* rather than the stimulus orientation. Accordingly, regions involved in perceptual leaning should have more information about *DV* than the stable stimulus orientation. We identified brain regions involved in perceptual learning by performing a voxel-wise comparison between information maps of *DV* and stimulus orientation by using paired t tests. This analysis revealed only one significant (p < 0.0001, k = 20, corrected for multiple comparisons at the cluster level, p < 0.001) cluster in the ACC (BA 32 [-9, 39, 24], t = 6.82, Figure 6). During stimulus presentation activity patterns in this region contain significantly more information about *DV* than stimulus orientation. Thus, this medial frontal region encodes a decision variable that changes during learning, suggesting that the ACC plays a key role for perceptual learning.

The discrepancy between the model-derived decision variable and stimulus orientation depends on the learning rate of the reinforcement learning model. The higher the learning rate the more *DV* deviates from the stimulus orientation. Therefore we reasoned that if the ACC encodes a decision variable which is shaped by a reinforcement learning mechanism, the contrast of information about *DV* > stimulus orientation in this region should be correlated with the individual learning rate of the model. Indeed, this correlation was significant (r = 0.50, p < 0.05), suggesting that subjects with higher learning rates have larger differences between encoding of *DV* and orientation in the ACC. This further strengthens our conclusion that ACC is critically involved in perceptual learning and decision-making.

No Changes in Early Visual Cortex with Training

One previous study suggested small changes in early visual stimulus representations during learning (Schoups et al., 2001). To investigate the possibility of such changes with training, we conducted an ROI analysis by using the cluster in the left lower early visual cortex in which significant information about orientation was encoded (see above). First we examined the orthogonal question whether stimulus representation in early visual cortex changes with training. The direct comparison between the information about stimulus orientation and the information about the decision variable in the early visual ROI revealed no significant differences (p = 0.24, t = 1.22). Thus, the dynamically changing *DV* does not provide a better account for early sensory representations than the static stimulus orientation. Importantly, we also did not find a significant difference between orientation encoding in the first and the second scanning session (p = 0.55, t = 0.61), suggesting that the representation of stimulus orientation did not
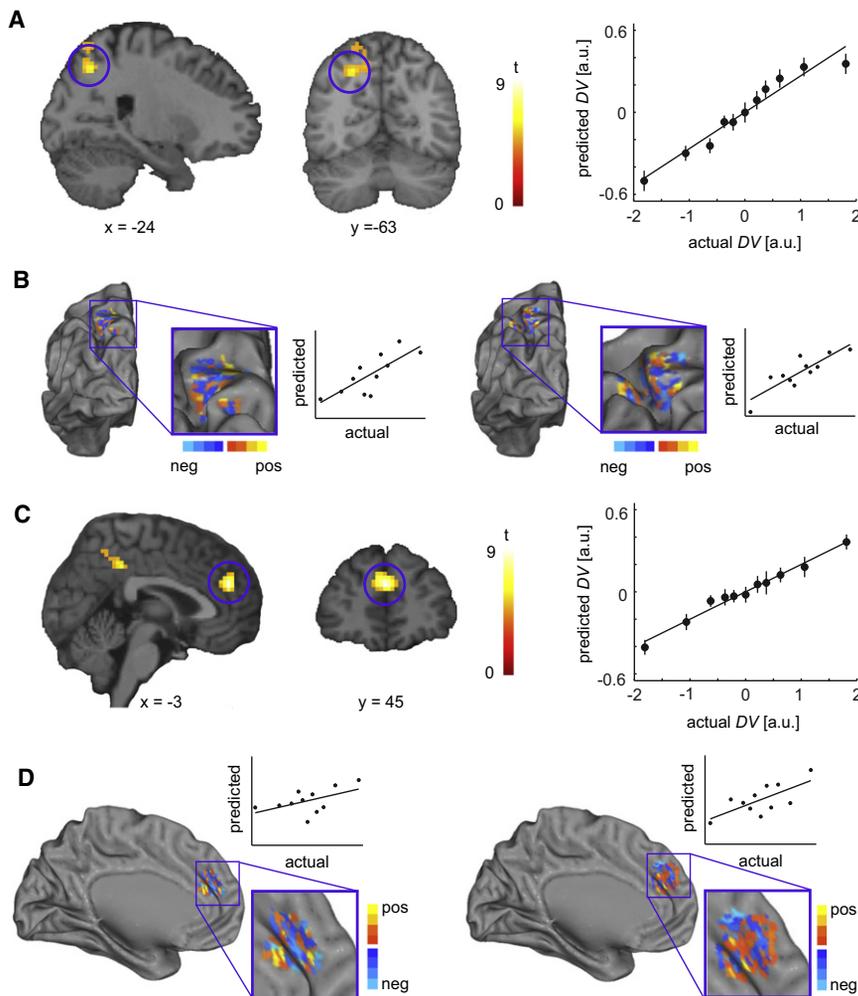
Figure 5. Encoding of Model-Derived Decision Variables

(A) Information about model-derived decision values is encoded in activity patterns in lateral parietal cortex. T-maps on prediction accuracy (Fisher's z-transformed correlation coefficients) are thresholded at p < 0.0001, k = 20. For illustration purposes, the scatterplot (right) visualizes the relationship between actual values of *DV* and values of *DV* predicted by the SVR (average of normalized predictions across cross-validation steps and subjects). Error bars = SEM for n = 20.
(B) Example of individual predictive maps (individual searchlight with peak prediction accuracy within the lateral parietal cortex) from two subjects together with their *DV* predictions. In yellow subregions activity increases with increasingly positive values of *DV*, whereas in blue subregions activity increases with increasingly negative *DV*.
(C and D) Same as in (A) and (B) but for information about model-derived *DV* in the ACC.

dopamine neurons (Bayer and Glimcher, 2005; Schultz et al., 1997). In human fMRI studies, however, prediction error signals have been identified in the ventral striatum, a target area of dopaminergic midbrain neurons (Kahnt et al., 2009; McClure et al., 2003; O'Doherty et al., 2003; Pessiglione et al., 2006). Thus, to provide further evidence for a reinforcement learning process in the current perceptual learning task, we regressed the signed prediction errors from the model against the feedback-locked BOLD signal in each voxel (see Experimental Procedures). We identified significant (p < 0.0001, k = 5) correlations between model-derived prediction errors and activity in the left ventral striatum ([-9, 0, −3], t = 4.77; Figure 7A), the bilateral anterior insular cortex extending into the lateral OFC (left BA 47 [-33, 21, −3], t = 5.56; right BA 47 [30, 21, −6], t = 6.49), the dorsolateral PFC (right BA 9 [54, 15, 36], t = 5.17), as well as the dorsomedial prefrontal cortex including the ACC (BA 32 [0, 27, 42], t = 5.81; Figure 7B; see Table S3 for complete results). This shows that the key learning variable of our computational model, namely the signed reward prediction error, is coded in the activity of reward-related regions such as the ventral striatum, providing further evidence for a reinforcement learning process in perceptual learning.

In a second step, we aimed to confirm that the learning-related changes in *DV* are indeed related to an updating mechanism that is based on signed prediction errors as proposed by our model. Thus, the same region in the ACC where activity patterns track perceptual learning-related changes in *DV* should also process reward prediction error signals. We performed a conjunction analysis searching for voxels which fulfill both of the following two criteria (both individually significant at p < 0.01): (1) voxels should contain more information about *DV* than orientation and

change with training. Moreover, to search for unspecific effects of perceptual learning in early visual cortex we compared the mean stimulus locked BOLD amplitude (estimated by a general linear model) between the two scanning days. Again, this comparison revealed no significant difference (p = 0.14, t = 1.55). Taken together, these analyses suggest that the representation of sensory evidence as well as unspecific BOLD responses in early sensory areas did not change significantly over the course of learning.

Reinforcement Processes in Perceptual Learning

So far we have shown that (1) the predictions of an adapted reinforcement learning model correlate with learning-related changes in orientation discrimination performance over time and (2) that the model-derived *DV*, which builds the basis for perceptual decisions, is coded in the medial frontal cortex. However, because alternative learning models would also predict similar increases in *DV* over learning, in the following analyses we provide further evidence for the proposed reinforcement learning mechanism. Evidence for Rescorla-Wagner-like updating in the reward-learning literature originally came from the observation of signed reward prediction error signals in
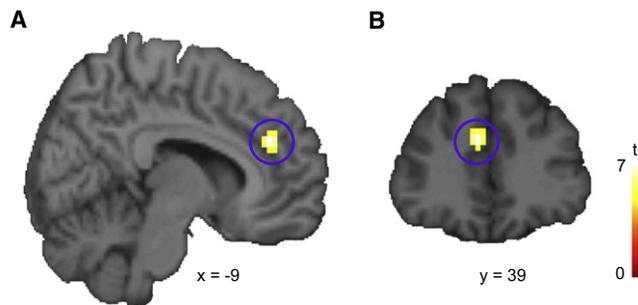
Figure 6. Perceptual Learning-Related Activity in the ACC
(A and B) Activity patterns in the ACC contain significantly more information about model-derived *DV* than stimulus orientation. T-maps on differential prediction accuracy (decision variable > stimulus orientation) are thresholded at p < 0.0001, k = 20.

(2) BOLD signals should correlate with signed prediction errors derived from the model. This conjunction analysis identified a cluster in the ACC (BA 24/32) in which voxels fulfilled both criteria (Figures 7C and 7D). This supports our conclusion that perceptual learning in the ACC is indeed driven by a Rescorla-Wagner-like updating mechanism, providing further and necessary support for a role of reinforcement processes in perceptual learning and decision-making.

DISCUSSION

Here we have shown that a reinforcement learning process can account for behavioral and neural changes during perceptual learning. Specifically, perceptual improvements over the course of 42 training runs were well explained by a reinforcement learning model. This model uses a simple delta rule (Sutton and Barto, 1998) to update a perceptual weight which is used to transform sensory information into a decision variable. In other words, perceptual learning in this model is established by an improved readout of sensory information leading to noise-robust representations of decision variables that build the basis for perceptual choices. By using multivariate information mapping techniques we found stimulus orientation to be encoded in the early visual cortex as well as higher cortical regions such as the LIP. However, learning-related changes in activity were found only in higher order brain regions. Specifically, we found activity patterns in the ACC that encoded learning-related changes in *DV* significantly better than the stimulus orientation. This provides direct evidence that perceptual learning is accompanied by changes in higher order brain regions. Furthermore, we show that our task involves reward prediction error signaling in reward-related brain regions but also higher decision-making areas, providing further evidence for reinforcement processes in perceptual learning.

Previous electrophysiological work in primates also showed that reinforcement learning models can account for perceptual learning (Law and Gold, 2009). Similar to our finding for the ACC, Law and Gold showed that decision variables represented in LIP neurons became more noise-robust during training. However, here we found such changes in the ACC but not the
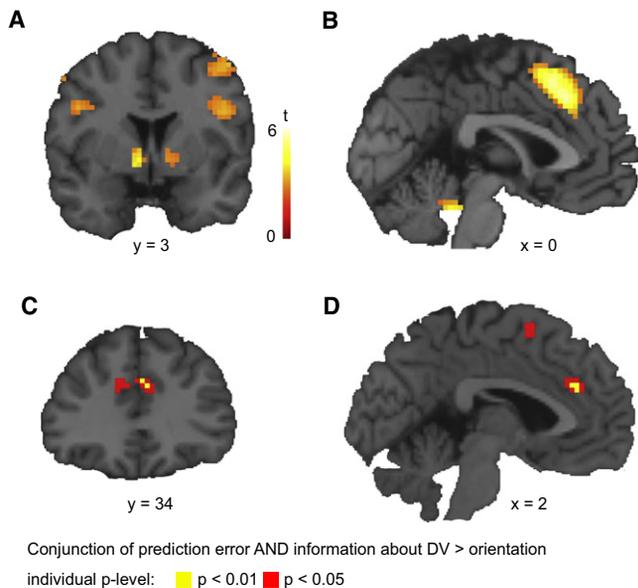

Conjunction of prediction error AND information about DV > orientation

individual p-level: ▨ p < 0.01 ■ p < 0.05

Figure 7. Activity Related to Signed Reward Prediction Errors
(A and B) Activity in the ventral striatum (A) and the medial frontal cortex (B) is significantly correlated with signed reward prediction errors derived from the reinforcement learning model. T-maps are thresholded at p < 0.005 for illustration purposes.
(C and D) Conjunction analysis identifying voxels in which BOLD activity correlates with model-derived signed prediction errors and which contain more information about *DV* than orientation. Voxels that fulfill both criteria (individually) at p < 0.01 and p < 0.05 are shown in yellow and red, respectively.

putative LIP. This discrepancy can be explained by differences in the experimental design. In their original study (Law and Gold, 2008), monkeys made saccades into and out of the response field of the recorded LIP neurons and single-unit responses were analyzed during stimulus presentation, which overlapped with saccade execution (i.e., decisions equal the ocular motor action). In contrast, in the current fMRI experiment human subjects made button presses by using a response mapping screen later in the trial that allowed the dissociation of the perceptual choice from preparatory end executive motor signals. However, in line with their results our data also demonstrate that activity changes in decision-making areas but not early visual areas account for perceptual learning.

We conducted several analyses to search for perceptual learning-related changes in early visual representations of stimulus orientation. None of these provided any evidence for perceptual learning. This is in line with findings that monkeys trained on similar visual tasks show only little (Schoups et al., 2001) if any change in the early visual cortex (Crist et al., 2001; Ghose et al., 2002). Nevertheless, our combination of fMRI, multivariate decoding, and computational modeling might not be sensitive enough to find any potentially subtle changes in early visual representations. However, our method is sensitive enough to decode the stimulus orientation itself in visual cortex. It is also sufficiently sensitive to find learning-related changes in medial frontal cortex. This could suggest an alternative account for perceptual learning which involves higher cortical

representations of decision variables. Importantly, this account is in line with results from monkey electrophysiology (Law and Gold, 2008) as well as with recent psychophysical and modeling work (Zhang and Li, 2010; Zhang et al., 2010a, 2010b). Furthermore, studies investigating perceptual decisions revealed a similar dissociation between early sensory regions and frontal areas (Heekeren et al., 2008; Romo and Salinas, 2003). Specifically, sensory areas have been shown to track the physical stimulus properties, whereas neural activity in frontal cortex tracks perceptual judgments and thus the subjective experience of the stimulus (de Lafuente and Romo, 2005, 2006; Heekeren et al., 2004; Hernández et al., 2010; Lemus et al., 2010; Salinas et al., 2000).

Our model suggests that reinforcement processes account for perceptual learning. This is in line with recent conceptual work that proposes a common mechanism for perceptual and reward-based decisions (Rushworth et al., 2009). It is also consistent with recent models of perceptual learning (Seitz and Watanabe, 2005; Seitz and Dinse, 2007) in which reinforcement signals drive perceptual learning, even if features are task-irrelevant, unattended (Dinse et al., 2003; Seitz and Watanabe, 2003), or invisible (Seitz et al., 2009). Moreover, besides the behavioral fit of our model, we show that prediction errors correlate with activity in reward-related regions such as the ventral striatum but also in the ACC where perceptual learning-related changes in DV were identified. The presence of activity that correlates with signed prediction errors, the teaching signal in reinforcement learning models (Kahnt et al., 2009; McClure et al., 2003; O'Doherty et al., 2003; Pessiglione et al., 2006), provides further evidence for a reinforcement process in perceptual learning. Thus, our results provide empirical evidence that perceptual learning indeed operates by means of a reinforcement learning process that involves reward prediction errors and is accompanied by activity changes in frontal decision-making areas. This medial frontal brain region is also critically involved in reward-based learning and decision-making (Behrens et al., 2007; Hayden et al., 2009; Holroyd and Coles, 2002; Ito et al., 2003; Kennerley et al., 2006; Matsumoto et al., 2007; Rushworth et al., 2007). Thus, our results suggest that perceptual as well as reward-based learning and decision-making share a common neurobiological basis and that both can be studied in the framework of reinforcement learning.

Our results were achieved by combining computational models of reinforcement learning with multivariate data analysis methods. Rather than searching for univariate representations of model variables as in conventional model-based fMRI (O'Doherty et al., 2007), we searched for multivariate representations by using pattern recognition techniques (Haynes and Rees, 2006; Kriegeskorte et al., 2006; Norman et al., 2006). Multivariate approaches have proven to be more sensitive than univariate approaches for revealing the distributed cortical patterns encoding sensory variables, such as stimulus orientation, motion direction, or color, which are known to be encoded in the joint activity of differentially tuned neurons (Brouwer and Heeger, 2009; Haynes and Rees, 2005; Kamitani and Tong, 2005; Seymour et al., 2009). These patterns have been hypothesized to reflect biased sampling of neural activity (Haynes and Rees, 2005; Kamitani and Tong, 2005), complex spatiotemporal

dynamics involving the vascular system (Kriegeskorte et al., 2010; Shmuel et al., 2010), or large-scale biases (Mannion et al., 2010; Sasaki et al., 2006). Moreover, recent studies suggest that cognitive and decision variables also are encoded in distributed cortical activity patterns (Hampton and O'Doherty, 2007; Haynes et al., 2007; Kahnt et al., 2010, 2011; Soon et al., 2008). Taken together, our current approach of decoding variables derived from computational models could provide a fruitful tool to study neurocomputational processes underlying learning and decision-making.

In conclusion, here we have shown that behavioral improvements in an orientation discrimination task are accompanied by activity changes in the ACC. Thus, our data provide strong evidence for perceptual learning-related changes in higher order areas. Furthermore, perceptual improvements were well explained by a reinforcement learning model in which learning leads to an enhanced readout of sensory information, which in turn leads to noise-robust representations of decision variables. This learning process involves an updating mechanism based on signed prediction errors, just like classical reward learning. Taken together, these findings support the notion that perceptual learning relies on reinforcement processes and that it engages the same neural processes as reward-based learning and decision-making. They advance our knowledge about the neurobiological basis of perceptual learning and suggest that the long-established distinction between perceptual and reward-based learning should be reconsidered.

EXPERIMENTAL PROCEDURES

Task and Experimental Setup
In each trial subjects saw a low contrast (10%) Gabor patch (∼1cycle per degree) on mean gray background in the right upper visual field for 500 ms while fixating on a central fixation cross (Figure 1A). Fixation was controlled by using eye tracking throughout the experiment. In each trial the orientation of the Gabor could deviate from 45° in five steps in both directions, counterclockwise (41°, 42.6°, 43.6°, 44.2°, and 44.5°) and clockwise (45.5°, 45.8°, 46.4°, 47.4°, and 49°). After a variable delay (1.5–5.5 s), subjects were asked to indicate the perceived orientation (tilted toward counterclockwise versus tilted toward clockwise) on a response mapping screen (randomly assigning counterclockwise and clockwise decisions to left and right button presses) with the index or middle finger of their right hand. This allowed us to disentangle the perceptual decision from planning and executing the behavioral response. Directly after the response, feedback was provided for 500 ms by changing the color of the fixation cross to green given a correct decision or to red given an erroneous response. In 45° trials positive and negative feedback was provided randomly and balanced. Trials were separated by a variable interval of 1.5–4.5 s.

Subjects were trained over the course of 4 days. The first and last day involved six runs of fMRI data acquisition, whereas days 2 and 3 consisted of 15 runs of training without scanning. However, to ensure a constant environment across the entire experiment, training during days 2 and 3 took place in a mock scanner, simulating body position, visual stimulation, and noise of the actual MRI system in great detail. The experimental procedure was approved by the local ethics review board of the University of Magdeburg.

Reinforcement Learning Model for Perceptual Decision-Making
In each trial t a decision variable DV_t is computed according to $DV_t = x_t \cdot w_t$, where x_t is the stimulus orientation (minus 45°) and w_t is the perceptual weight that changes during learning. The model makes perceptual choices $p(cw)$ on the basis of DV according to: $p(cw)_t = 1/1 + e^{-\beta \cdot (DV_t - c)}$, where c is a bias term accounting for unspecific biases and β is the slope of the sigmoidal

function accounting for individual levels of noise. An expected value EV is computed based on absolute values of DV ($|DV|$) which equal the probability that the current trial will be rewarded: $EV_t = 1/1 + e^{-\beta \cdot (|DV_t| - c)}$. During feedback the expected value is compared to the actual reward (coded as 1 and 0 for positive and negative feedback, respectively) resulting in a reward prediction error δ: $\delta_t = r_t - EV_t$. This error is then used to update the perceptual weight in proportion to a learning rate α: $w_{t+1} = w_t + \alpha \cdot \delta_t$. For each subject the free model parameters (α, $w_{t=1}$, β and c) were estimated by fitting the vector of trial-by-trial model predictions $p(cw)$ against a vector of subjects' actual trial-by-trial perceptual choices (coded as counterclockwise = 0 and clockwise = 1). It is important to note that DV is orthogonal to potentially confounding variables like expected value, uncertainty, choice confidence, or task difficulty. These variables are based on absolute values of DV (i.e., $|DV|$) and therefore, both highly negative (high evidence for counterclockwise gratings) and highly positive values of DV (high evidence for clockwise gratings) result in a high expected value (as well as high confidence, low difficulty, and low uncertainty), whereas values close to zero result in a low expected value (as well as low confidence, high difficulty, and high uncertainty). Thus these variables encompass a U- (expected value, choice confidence) or inverted U-shaped (uncertainty, difficulty) relationship with stimulus orientation and hence DV.

fMRI Data Acquisition and Preprocessing

Functional MRI data were acquired on a 3-Tesla Siemens Trio (Erlangen, Germany) scanner. In each scanning run 341 volumes were acquired (TR = 2 s, 24 slices, 4.4 mm thick, in plane resolution 2 × 2 mm). Preprocessing was performed by using SPM2 (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) and included slice-time correction, realignment, and spatial normalization to a standard template (resampling to 3 mm isotropic voxels). Spatial normalization was used to ensure that data from both scanning days are in a common reference space.

Multivariate Decoding of fMRI Data

We used a searchlight approach that allows whole-brain information mapping without potentially biasing voxel selection (Haynes et al., 2007; Kahnt et al., 2010; Kriegeskorte et al., 2006) in combination with linear SVR. In a first step, for each subject, general linear models (GLM) were applied to the preprocessed functional imaging data of each run. The GLM contained 11 regressors for different stimulus orientations (41°, 42.6°, 43.6°, 44.2°, 44.5°, 45°, 45.5°, 45.8°, 46.4°, 47.4°, and 49°) and four regressors accounting for left and right button presses and positive and negative feedback, respectively (all convolved with a canonical hemodynamic response function) as well as six regressors accounting for variance induced by head motion. The voxel-wise parameter estimates represent the response amplitudes to each of the 11 orientations in each of the 12 scanning runs. These parameter estimates were then used as input for the SVR and deviations from 45° were used as labels. SVR was performed by using the LIBSVM implementation (http://www.csie.ntu.edu.tw/~cjlin/libsvm/) with a linear kernel and a preselected cost parameter of c = 0.01. For each searchlight (all voxels within a radius of 4 voxels surrounding the central voxel) we performed a 12-fold leave-one-out cross-validation. In each fold, training was based on data from 11 scanning runs and prediction accuracy was obtained in the independent 12th scanning run. The prediction accuracy of the central searchlight voxel was defined as the average Fisher's z-transformed correlation coefficient between the actual labels of the independent test data set and the labels predicted by the SVR model. Please note that because correlation coefficients are computed based on model predictions in the independent test data and not on model fits in the training data, this cross-validation procedure is completely insensitive to potential noise fitting (i.e., overfitting) in the training data. This method resulted in a three-dimensional map of locally distributed information about stimulus orientation for each subject.

To map information about the model-derived decision variables, we sorted trials into 11 groups based on their value of DV instead of stimulus orientation. Here, parameter estimates of the GLM represent the response amplitudes to each of the 11 values of DV in each of the 12 scanning runs. The average decision variable in each group and run was used as label for the SVR. Searchlight-based information mapping was performed in the very same way as for

stimulus orientation (see above) allowing an unbiased comparison of both information maps.

To identify regions with significant information about orientation and DV, respectively, we performed second-level analyses by using voxel-wise one-sample t tests on smoothed accuracy maps (6 mm full width at half maximum). To identify regions where significantly more information about the decision variable than orientation was encoded, we used voxel-wise paired t tests. For all whole-brain tests we applied the same statistical threshold of p < 0.0001, uncorrected, together with a cluster-extend threshold of k = 20 continuous voxels that survive whole-brain correction for multiple comparisons on the cluster level (p < 0.001).

Prediction Error Analysis

To confirm the involvement of prediction error like updating in the context of the current perceptual learning task, we searched for activity that correlates with the trial-wise prediction errors derived from our reinforcement learning model for perceptual decision-making. For this we set up a GLM with a parametric design (Büchel et al., 1998) in which the onset regressors for positive and negative feedback were trial-wise parametrically modulated by the model-derived signed reward prediction errors (δ). These modulated regressors were orthogonalized with respect to the regressors for the onset of positive and negative feedback and simultaneously regressed against the BOLD signal in each voxel. Activity that correlates with signed prediction errors was identified by using voxel-wise t tests on the parameter estimates of the parametrically modulated regressors.

REFERENCES

Barraclough, D.J., Conroy, M.L., and Lee, D. (2004). Prefrontal cortex and decision making in a mixed-strategy game. Nat. Neurosci. 7, 404–410.

Bayer, H.M., and Glimcher, P.W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47, 129–141.

Behrens, T.E., Woolrich, M.W., Walton, M.E., and Rushworth, M.F. (2007). Learning the value of information in an uncertain world. Nat. Neurosci. 10, 1214–1221.

Brouwer, G.J., and Heeger, D.J. (2009). Decoding and reconstructing color from responses in human visual cortex. J. Neurosci. 29, 13992–14003.

Büchel, C., Holmes, A.P., Rees, G., and Friston, K.J. (1998). Characterizing stimulus-response functions using nonlinear regressors in parametric fMRI experiments. Neuroimage 8, 140–148.

Chowdhury, S.A., and DeAngelis, G.C. (2008). Fine discrimination training alters the causal contribution of macaque area MT to depth perception. Neuron 60, 367–377.

Crist, R.E., Li, W., and Gilbert, C.D. (2001). Learning to see: Experience and attention in primary visual cortex. Nat. Neurosci. 4, 519–525.

de Lafuente, V., and Romo, R. (2005). Neuronal correlates of subjective sensory experience. Nat. Neurosci. 8, 1698–1703.

de Lafuente, V., and Romo, R. (2006). Neural correlate of subjective sensory experience gradually builds up across cortical areas. Proc. Natl. Acad. Sci. USA *103*, 14266–14271.

Dinse, H.R., Ragert, P., Pleger, B., Schwenkreis, P., and Tegenthoff, M. (2003). Pharmacological modulation of perceptual learning and associated cortical reorganization. Science *301*, 91–94.

Ghose, G.M., Yang, T., and Maunsell, J.H. (2002). Physiological correlates of perceptual learning in monkey V1 and V2. J. Neurophysiol. *87*, 1867–1888.

Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

Goldstone, R.L. (1998). Perceptual learning. Annu. Rev. Psychol. *49*, 585–612.

Hampton, A.N., and O'Doherty, J.P. (2007). Decoding the neural substrates of reward-related decision making with functional MRI. Proc. Natl. Acad. Sci. USA *104*, 1377–1382.

Hayden, B.Y., Pearson, J.M., and Platt, M.L. (2009). Fictive reward signals in the anterior cingulate cortex. Science *324*, 948–950.

Haynes, J.D., and Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat. Neurosci. *8*, 686–691.

Haynes, J.D., and Rees, G. (2006). Decoding mental states from brain activity in humans. Nat. Rev. Neurosci. *7*, 523–534.

Haynes, J.D., Sakai, K., Rees, G., Gilbert, S., Frith, C., and Passingham, R.E. (2007). Reading hidden intentions in the human brain. Curr. Biol. *17*, 323–328.

Heekeren, H.R., Marrett, S., Bandettini, P.A., and Ungerleider, L.G. (2004). A general mechanism for perceptual decision-making in the human brain. Nature *431*, 859–862.

Heekeren, H.R., Marrett, S., and Ungerleider, L.G. (2008). The neural systems that mediate human perceptual decision making. Nat. Rev. Neurosci. *9*, 467–479.

Hernández, A., Nácher, V., Luna, R., Zainos, A., Lemus, L., Alvarez, M., Vázquez, Y., Camarillo, L., and Romo, R. (2010). Decoding a perceptual decision process across cortex. Neuron *66*, 300–314.

Holroyd, C.B., and Coles, M.G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. Psychol. Rev. *109*, 679–709.

Ito, S., Stuphorn, V., Brown, J.W., and Schall, J.D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science *302*, 120–122.

Kahnt, T., Park, S.Q., Cohen, M.X., Beck, A., Heinz, A., and Wrase, J. (2009). Dorsal striatal-midbrain connectivity in humans predicts how reinforcements are used to guide decisions. J. Cogn. Neurosci. *21*, 1332–1345.

Kahnt, T., Heinzle, J., Park, S.Q., and Haynes, J.D. (2010). The neural code of reward anticipation in human orbitofrontal cortex. Proc. Natl. Acad. Sci. USA *107*, 6010–6015.

Kahnt, T., Heinzle, J., Park, S.Q., and Haynes, J.D. (2011). Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. Neuroimage *56*, 709–715.

Kamitani, Y., and Tong, F. (2005). Decoding the visual and subjective contents of the human brain. Nat. Neurosci. *8*, 679–685.

Karni, A., and Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. Proc. Natl. Acad. Sci. USA *88*, 4966–4970.

Kennerley, S.W., Walton, M.E., Behrens, T.E.J., Buckley, M.J., and Rushworth, M.F.S. (2006). Optimal decision making and the anterior cingulate cortex. Nat. Neurosci. *9*, 940–947.

Kriegeskorte, N., Goebel, R., and Bandettini, P. (2006). Information-based functional brain mapping. Proc. Natl. Acad. Sci. USA *103*, 3863–3868.

Kriegeskorte, N., Cusack, R., and Bandettini, P. (2010). How does an fMRI voxel sample the neuronal activity pattern: Compact-kernel or complex spatiotemporal filter? Neuroimage *49*, 1965–1976.

Law, C.T., and Gold, J.I. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. Nat. Neurosci. *11*, 505–513.

Law, C.T., and Gold, J.I. (2009). Reinforcement learning can account for associative and perceptual learning on a visual-decision task. Nat. Neurosci. *12*, 655–663.

Lemus, L., Hernández, A., Luna, R., Zainos, A., and Romo, R. (2010). Do sensory cortices process more than one sensory modality during perceptual judgments? Neuron *67*, 335–348.

Li, W., Piëch, V., and Gilbert, C.D. (2004). Perceptual learning and top-down influences in primary visual cortex. Nat. Neurosci. *7*, 651–657.

Li, S., Mayhew, S.D., and Kourtzi, Z. (2009). Learning shapes the representation of behavioral choice in the human brain. Neuron *62*, 441–452.

Mannion, D.J., McDonald, J.S., and Clifford, C.W. (2010). Orientation anisotropies in human visual cortex. J. Neurophysiol. *103*, 3465–3471.

Matsumoto, M., Matsumoto, K., Abe, H., and Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. Nat. Neurosci. *10*, 647–656.

McClure, S.M., Berns, G.S., and Montague, P.R. (2003). Temporal prediction errors in a passive learning task activate human striatum. Neuron *38*, 339–346.

Norman, K.A., Polyn, S.M., Detre, G.J., and Haxby, J.V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. Trends Cogn. Sci. (Regul. Ed.) *10*, 424–430.

O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H., and Dolan, R.J. (2003). Temporal difference models and reward-related learning in the human brain. Neuron *38*, 329–337.

O'Doherty, J.P., Hampton, A., and Kim, H. (2007). Model-based fMRI and its application to reward learning and decision making. Ann. N Y Acad. Sci. *1104*, 35–53.

Pasupathy, A., and Miller, E.K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. Nature *433*, 873–876.

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., and Frith, C.D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature *442*, 1042–1045.

Platt, M.L., and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. Nature *400*, 233–238.

Romo, R., and Salinas, E. (2003). Flutter discrimination: Neural codes, perception, memory and decision making. Nat. Rev. Neurosci. *4*, 203–218.

Rushworth, M.F., Buckley, M.J., Behrens, T.E., Walton, M.E., and Bannerman, D.M. (2007). Functional organization of the medial frontal cortex. Curr. Opin. Neurobiol. *17*, 220–227.

Rushworth, M.F., Mars, R.B., and Summerfield, C. (2009). General mechanisms for making decisions? Curr. Opin. Neurobiol. *19*, 75–83.

Salinas, E., Hernandez, A., Zainos, A., and Romo, R. (2000). Periodicity and firing rate as candidate neural codes for the frequency of vibrotactile stimuli. J. Neurosci. *20*, 5503–5515.

Sasaki, Y., Rajimehr, R., Kim, B.W., Ekstrom, L.B., Vanduffel, W., and Tootell, R.B. (2006). The radial bias: A different slant on visual orientation sensitivity in human and nonhuman primates. Neuron *51*, 661–670.

Sasaki, Y., Nanez, J.E., and Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. Nat. Rev. Neurosci. *11*, 53–60.

Schoups, A., Vogels, R., Qian, N., and Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. Nature *412*, 549–553.

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

Seitz, A.R., and Dinse, H.R. (2007). A common framework for perceptual learning. Curr. Opin. Neurobiol. *17*, 148–153.

Seitz, A.R., and Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive? Nature *422*, 36.

Seitz, A., and Watanabe, T. (2005). A unified model for perceptual learning. Trends Cogn. Sci. (Regul. Ed.) *9*, 329–334.

Seitz, A.R., Kim, D., and Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. Neuron *61*, 700–707.

Seymour, K., Clifford, C.W., Logothetis, N.K., and Bartels, A. (2009). The coding of color, motion, and their conjunction in the human visual cortex. Curr. Biol. *19*, 177–183.

Shmuel, A., Chaimow, D., Raddatz, G., Ugurbil, K., and Yacoub, E. (2010). Mechanisms underlying decoding at 7 T: Ocular dominance columns, broad structures, and macroscopic blood vessels in V1 convey information on the stimulated eye. Neuroimage *49*, 1957–1964.

Soon, C.S., Brass, M., Heinze, H.J., and Haynes, J.D. (2008). Unconscious determinants of free decisions in the human brain. Nat. Neurosci. *11*, 543–545.

Sugrue, L.P., Corrado, G.S., and Newsome, W.T. (2004). Matching behavior and the representation of value in the parietal cortex. Science *304*, 1782–1787.

Sutton, R., and Barto, A. (1998). Reinforcement Learning: An Introduction (Cambridge, MA: MIT Press).

Zhang, E., and Li, W. (2010). Perceptual learning beyond retinotopic reference frame. Proc. Natl. Acad. Sci. USA *107*, 15969–15974.

Zhang, J.Y., Zhang, G.L., Xiao, L.Q., Klein, S.A., Levi, D.M., and Yu, C. (2010a). Rule-based learning explains visual perceptual learning and its specificity and transfer. J. Neurosci. *30*, 12323–12328.

Zhang, T., Xiao, L.Q., Klein, S.A., Levi, D.M., and Yu, C. (2010b). Decoupling location specificity from perceptual learning of orientation discrimination. Vision Res. *50*, 368–374.