

UNIVERSITÄTSKLINIKUM HAMBURG-EPPENDORF

Institut für Systemische Neurowissenschaften

Prof. Dr. med. Christian Büchel

Changes in neural patterns following error-induced learning

Dissertation

zur Erlangung des Grades eines Doktors der Medizin
an der Medizinischen Fakultät der Universität Hamburg

vorgelegt von:

Léon Ferdinand Möhring
aus Bremen

Hamburg 2024

(wird von der Medizinischen Fakultät ausgefüllt)

Angenommen von der

Medizinischen Fakultät der Universität Hamburg am: 06.02.2025

Veröffentlicht mit Genehmigung der

Medizinischen Fakultät der Universität Hamburg.

Prüfungsausschuss, der/die Vorsitzende: Prof. Dr. Jürgen Gallinat

**Prüfungsausschuss, zweite/r Gutachter/in: Prof. Dr. Christian Büchel
Prof. Dr. Dr. Andras Heinz**

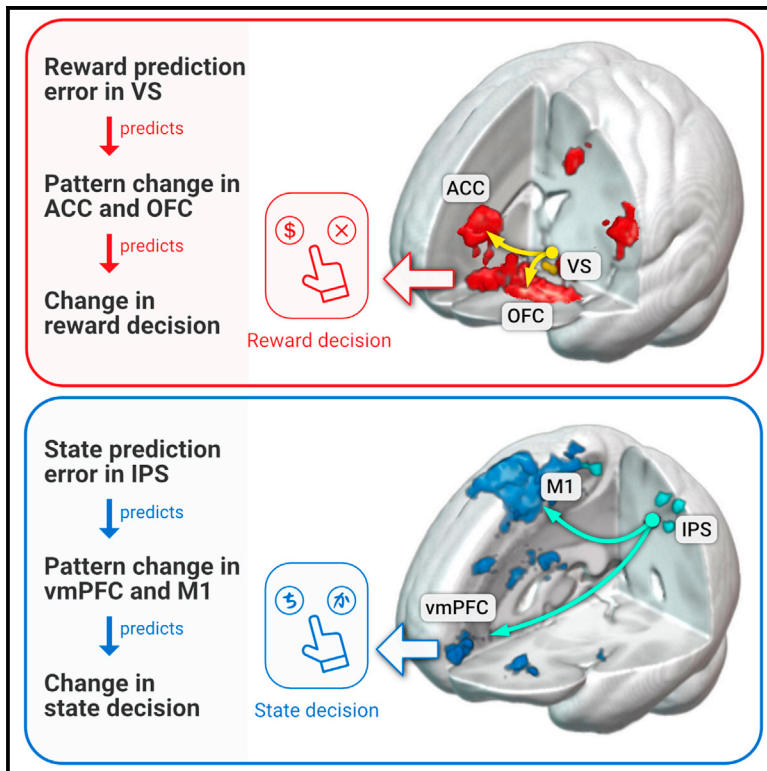
1. Original manuscript	1
1.1 Prediction errors drive dynamic changes in neural patterns that guide behavior (<i>Cell Reports, 2023, Möhring and Gläscher</i>)	1
1.2 Protocol for predicting multivariate change of brain patterns using model-informed fMRI activations (<i>STAR Protocols, 2024, Möhring and Gläscher</i>)	32
2. Description of the publication with references	51
2.1 Introduction	51
2.2 Model-free and model-based Reinforcement learning	51
2.3 Neural correlates of the reward prediction error	53
2.4 The BOLD response and model-informed fMRI analyses	54
2.5 Predicting multivariate change of brain patterns using PEs	55
2.6 Results of the publication	57
2.7 Discussion	59
2.8 References	61
3. Summary	64
3.1 English version	64
3.2 German version	65
4. Declaration of self-contribution	66
5. Acknowledgements	66
6. Curriculum vitae / Lebenslauf	67
7. Statutory declaration / Eidesstattliche Versicherung	68

1. Original manuscript

1.1 Prediction errors drive dynamic changes in neural patterns that guide behavior
(*Cell Reports*, 2023, Möhring and Gläscher)

Prediction errors drive dynamic changes in neural patterns that guide behavior

Graphical abstract



Authors

Leon Möhring, Jan Gläscher

Correspondence

leon.moehring@stud.uke.uni-hamburg.de (L.M.),
glaescher@uke.de (J.G.)

In brief

Möhring and Gläscher discovered that distinct prediction error signals in the human brain modulate short-term reconfigurations of neural patterns in dissociable brain networks. Additionally, these pattern reconfigurations predict changes in the behavioral policy. These findings provide a comprehensive account of the interaction between prediction errors, neural patterns, and behavioral adaptations.

Highlights

- RPEs and SPEs elicit BOLD responses in distinct areas, replicating earlier findings
- RPEs and SPEs predict short-term pattern changes in dissociable brain networks
- Pattern change in vmPFC, ACC, and OFC predicts adaptations of the behavioral policy



Article

Prediction errors drive dynamic changes in neural patterns that guide behavior

Leon Möhring^{1,*} and Jan Gläscher^{1,2,*}¹Institute for Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany²Lead contact*Correspondence: leon.moehring@stud.uke.uni-hamburg.de (L.M.), glaescher@uke.de (J.G.)<https://doi.org/10.1016/j.celrep.2023.112931>**SUMMARY**

Learning describes the process by which our internal expectation models of the world are updated by surprising outcomes (prediction errors [PEs]) to improve predictions of future events. However, the mechanisms through which error signals dynamically influence existing neural representations are unknown. Here, we use functional magnetic resonance imaging (fMRI) in humans solving a two-step Markov decision task to investigate changes in neural activation patterns following PEs. Using a dynamic multivariate pattern analysis, we can show that PE-related fMRI responses in error-coding regions predict trial-by-trial changes in multivariate neural patterns in the orbitofrontal cortex, the precuneus, and the ventromedial prefrontal cortex (vmPFC). Importantly, the dynamics of these pattern changes in the vmPFC also predicted upcoming changes in choice strategies and thus highlight the importance of these pattern changes for behavior.

INTRODUCTION

Living in a multifaceted world requires humans and other organisms to be responsive to the environment. The brain's internal model of this world encodes our expectations, enabling us to predict upcoming events and to respond efficiently to them. However, when the world changes dynamically, these internal models have to be updated to allow for adaptive changes in the behavioral policy. Minimizing the discrepancy between internal expectations encoded in our models and real-world outcomes is an essential concept of learning. Prediction errors (PEs) are the computational equivalent of these expectancy violations and act as a teaching signal in reinforcement learning (RL).¹ Throughout the learning process, both PEs and expected values change over time, ideally to the point where our expectations perfectly match reality, and the PE reduces to zero. However, the process through which error signals can update our internal model and how this is implemented at the neural level is an outstanding question. In RL, PEs together with the continuously updated expected values are sufficient to drive learning. Though, an implementation of this process in a biological system (e.g., a human brain) entails not only changes in representations of these core computational variables but may also evoke changes in other cognitive systems that facilitate the adjustment of expectations.^{2,3} Attentional resources, working memory, and motor planning are likely to be involved in reshaping our internal model following a PE.

Multivariate methods have introduced a new window into stimulus representations in the human brain by analyzing the information contained in spatially distributed patterns of neural activity as measured by functional magnetic resonance imaging (fMRI).^{4–6} While some studies have used machine-learning algorithms to

decode information about stimuli and their properties,^{7–9} others have characterized stimulus representations through the (dis)similarity of activation patterns and compared them with the perceptual or categorical similarity.^{10–12} It has been demonstrated that even hidden states, such as subjective value of outcomes, have a multivariate representation within the brain.^{13,14} In addition, the multivariate approach has also been extended to learning-induced changes in the similarity of activation patterns.^{15,16}

PEs are a quantitative signal of how much expectations deviate from the experienced stimuli in the environment. The occurrence of a large PE signals the need for substantial adaptations of our internal predictive model, whereas small PEs only call for minor adjustments. Consequently, patterns of neural activity encoding these internal representations need to undergo greater changes following a large PE than in the case of a small PE. These theoretical considerations suggest that activation patterns of two adjacent trials should change proportionally to the size of the PE between them. Importantly, to detect error-specific changes in neural patterns, we utilize two different types of PEs simultaneously in this project: reward PEs (RPEs) and state PEs (SPEs). Hereby, we can reveal regions in which neural patterns are only modulated by one distinct PE type.

Over the past two decades, neural correlates of PEs have been studied in great detail. RPEs are the key teaching signal of model-free RL,^{17–19} a domain of RL algorithms that learn in a pure trial-and-error fashion and only focus on the obtained outcomes without considering the transition probabilities between states. RPEs have been robustly associated with the phasic activity of dopaminergic neurons in the ventral tegmental area (VTA),^{20,21} an area projecting to a wide range of subcortical and cortical regions,^{22,23} including the ventral striatum (VS), where RPEs are commonly detected in human fMRI studies.^{24–26} In addition,



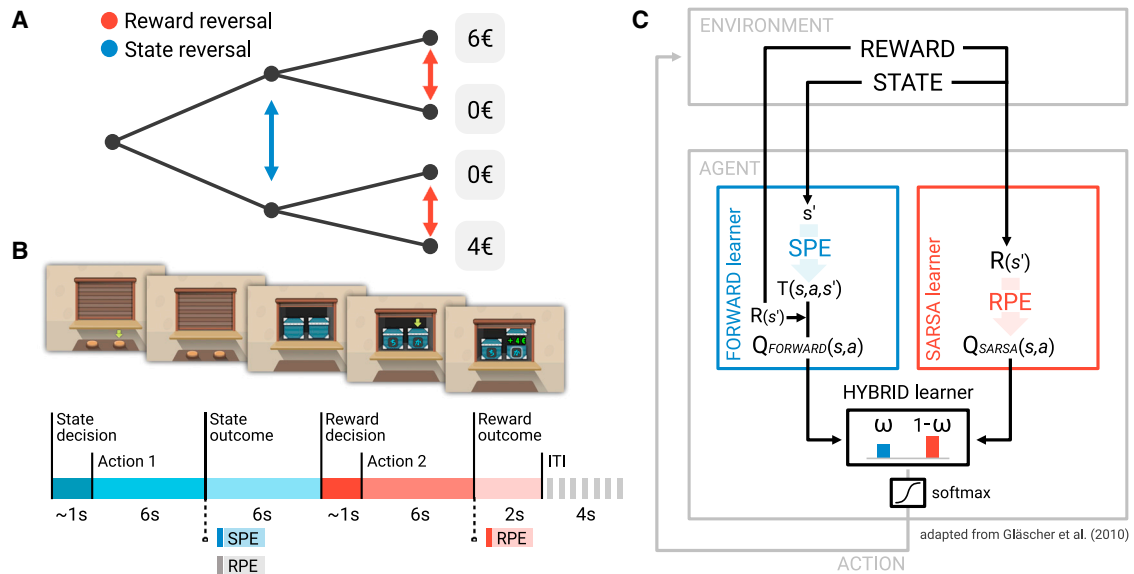


Figure 1. Two-step task and hybrid-learning model

(A and B) Participants ($n = 44$) performed a modified two-step Markov decision task that produces orthogonal SPEs and RPEs (correlation of SPE and RPE: $r = 0.0635$) (ITI, inter-trial interval). Detailed information about the task can be found in the [STAR Methods](#).

(C) Schematic of the hybrid-learning model combining model-based and model-free RL using a constant weighting parameter. s' , the state the agent transitioned to; a , action; $R(s')$, reward obtained; $T(s,a,s')$, state transition probability estimates of the agent; $Q(s,a)$, reward value estimates of the agent; ω , constant weighting parameter.

dopaminergic VTA neurons have also been associated with reward probability and value during reward anticipation.^{27,28}

On the contrary, model-based RL algorithms aim to estimate state-transition probabilities and build a model of the existing state space of the task. Following a theoretical distinction between model-free and model-based RL,²⁹ we identified distinct brain regions exhibiting a significant correlation of their blood oxygenation level-dependent (BOLD) activity with the specific PE signal derived for these two types of learning.³⁰ While RPEs are represented in the VS, we have shown correlates of model-based SPEs in the intraparietal sulcus (IPS) and the dorsolateral prefrontal cortex (dlPFC). These findings have been replicated multiple times using variants of the original task³¹ as well as another two-step Markov decision task.³² Critically, all studies demonstrated a correlation of brain activity with the magnitude of these two error signals, suggesting that errors of various sizes express themselves in corresponding strengths of the BOLD response. However, none of these studies investigated the ensuing changes in activity patterns of associated brain regions that facilitate learning in these two domains.

Here, using a modified two-step Markov decision task that elicits orthogonal SPEs and RPEs, we investigate representational changes following two types of PEs during learning. In particular, we utilize these orthogonal PEs to investigate how expectancy violations change neural representations that encode these expectations in our internal model. By using these two orthogonal error signals, we can further test for the specificity of these changes in neural patterns of one error-related brain network but not in the other.

We develop an analytical approach, which combines model-derived univariate and multivariate fMRI analyses and reveals

the influence of PEs on subsequent changes in patterns of neural activity. In particular, this method is capable of quantifying the magnitude of multivariate changes in neural patterns over time and allows us to dynamically predict these pattern changes using trial-by-trial BOLD responses in error-coding regions. We demonstrate that the PE-evoked BOLD activity predicts the magnitude of subsequent multivariate changes in activation patterns of cognitive systems involved in working memory, attention, and motor planning. Additionally, pattern changes in the orbitofrontal cortex (OFC) and the ventromedial prefrontal cortex (vmPFC) also predict upcoming adjustments of the behavioral policy. In summary, our findings point to an immediate effect of PE signals on the reconfiguration of neural representations in distinct brain regions for two different domains of learning.

RESULTS

Experimental design and behavioral modeling

Each participant ($n = 44$) performed 5 runs of a two-step Markov decision task (Figure 1). The first-stage decision (state decision) determined to which of two possible states (blue, red) the participants would transition to. In the second stage (reward decision), participants chose between one of two boxes to obtain a reward of either 0, 4, or 6 €. PEs derived from model-based and model-free learning occurred at the time of state outcome (SPE) and reward outcome (RPE). The task included reversal trials (permanently swapping states or rewards) and catch trials (onetime unexpected state or reward in a trial) to subvert expectations and trigger the participants to react to changes and update their internal model of the task. Importantly, this task produces orthogonal SPEs and RPEs (correlation of all model-derived

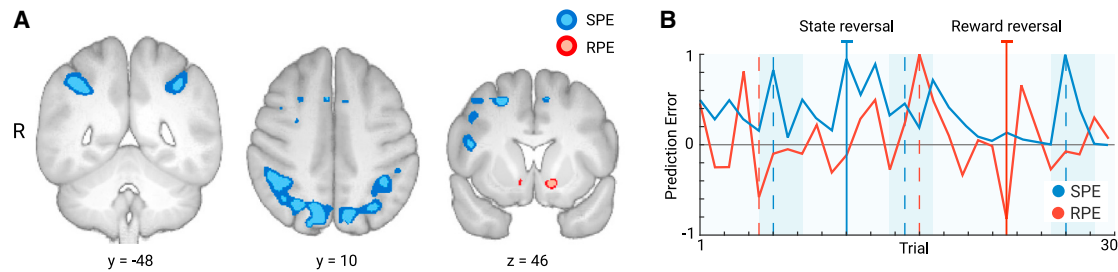


Figure 2. Error-coding regions of SPE and RPE

(A) Results of a univariate regression analysis showing neural correlates of the model-derived SPE and RPE signals. All results are presented at a threshold of $p < 0.005$ uncorrected (dark area) and $p < 0.001$ uncorrected (light area) for visualization. These and all other images are in radiological orientation. (B) Exemplary visualization of model-derived SPE and RPE signals throughout a single run. Light blue background represents free-choice trials, and dark blue background represents forced-choice trials. Solid vertical lines represent reversal trials, and dashed lines represent catch trials. The modified two-step task produces orthogonal SPEs and RPEs: state-reversal trials elicit a large SPE but almost no RPE, and vice versa.

SPEs and RPEs: $r = 0.0635$; see Figure S1), enabling us to study the effects of both PE types independently. We achieve this by ensuring that both possible states (blue, red) always allow for similar amounts of rewards³³ so that a state reversal elicits a large SPE but almost no RPE (see Figure 2B for an example of both error signals).

Based on previous work in this field,^{30,32} we assumed that participants utilize a mixture of model-based and model-free learning during the execution of this task. Behavioral modeling was conducted with a hybrid-learning algorithm combining model-based and model-free RL using a constant weighting parameter to predict the behavioral choice data and derive SPEs and RPEs.³⁰ This hybrid-learning algorithm provides a significantly better model fit than degenerate versions of the model using either strategy (model-free or model-based) alone (see model comparison and group-level mean values for model parameters in Table S1).

in the VS: $x, y, z = -8, 15, -8$, $p_{\text{FWE,SVC}} = 0.005$ [small-volume-corrected]). We then defined ROIs of error-coding regions for the SPE and RPE by combining the current results with the ones of our previous study³⁰ in a conjunction³⁵ at an uncorrected threshold of $p < 0.001$.

Measuring short-term changes in neural activation patterns

To investigate the hypothesis that PEs modulate changes in neural activation patterns, it is necessary to quantify the magnitude of multivariate changes in neural patterns over time. Here, we introduce a metric called pattern change, which is based on the Pearson correlation coefficient of two patterns. The pattern change occurring between corresponding patterns of two adjacent trials is defined as the Fisher-transformed correlation distance of the pattern in the current and upcoming trial:

$$\text{pattern change}[t, t + 1] = -\text{arctanh}(\text{corr}(\text{pattern}[t], \text{pattern}[t + 1])).$$

Model-based and model-free PE signals

A number of studies have established that BOLD responses in the VS convey a model-free RPE signal.^{24–26} Moreover, in our previous work, we were able to detect the existence of an SPE signal in the IPS.³⁰ To define regions of interest (ROIs) for subsequent analyses, we searched for neural correlates of SPEs and RPEs derived from the hybrid-learning model. For this and all other neuroimaging analyses, we decided to present statistical maps at uncorrected thresholds of $p < 0.001$ and < 0.005 for visualization purposes. However, in the main text, we report family-wise-error-corrected (FWE) effects using threshold-free cluster enhancement.³⁴ We were able to replicate the earlier findings in this project (Figure 2A). Specifically, we found significant clusters resembling the SPE in the right IPS and the precuneus (SPE in the IPS: $x, y, z = 40, -45, 45$, $p_{\text{FWE}} = 0.019$; SPE in the precuneus [PCu]: $x, y, z = -10, -70, 43$, $p_{\text{FWE}} = 0.018$), while the RPE was located in the VS (RPE

The resulting pattern change is minimal when two patterns are perfectly correlated and is maximal when there is a perfect inverse correlation.

We used a searchlight-based algorithm (searchlight radius = 10 mm) to calculate pattern change throughout the whole brain for all trials (Figure 3). The input data for the searchlight algorithm were trial-wise beta estimates at the time of state and reward decision. If in alignment with our hypothesis, pattern change at state decision would correlate with the preceding SPE, and pattern change at reward decision would correlate with the preceding RPE (i.e., the pattern change is modeled after the occurrence of the PE). This procedure resulted in a set of trial-by-trial images encoding the magnitude of changes in neural activation patterns from one trial to the next in every searchlight. Critically, this approach compresses the multivariate information contained in the transitions of corresponding patterns from the current to the next trial into a univariate signal of change. As a

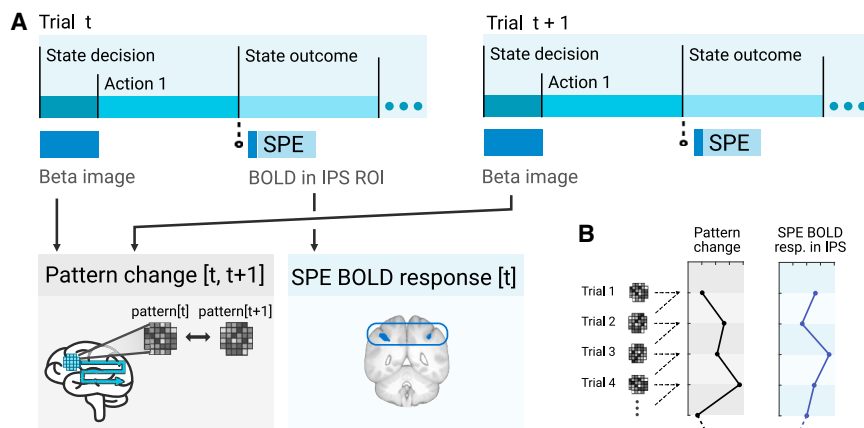


Figure 3. Measuring changes in neural activation patterns

(A) Schematic of detecting SPE-modulated pattern change at the time of state decision. For RPE-modulated pattern change at the time of reward decision, the procedure was performed in the same way. For each searchlight sphere (radius = 10 mm), pattern change was calculated using the patterns of corresponding events in the current and upcoming trials. We hypothesized the pattern change at the time of state decision to be modulated by SPE-evoked BOLD responses in the IPS ROI. Importantly, the PE occurred between the two patterns that define the pattern change and is therefore suited to drive the process of updating these neural representations.

(B) Maps of pattern change were created for every trial. Critically, the use of a searchlight-based analysis compresses the multivariate changes in neural activation patterns into a univariate signal of change. This allowed for directly relating pattern changes to univariate PE-evoked BOLD responses in a linear model.

consequence, this enables us to directly relate trial-by-trial changes in neural activation patterns to univariate PE responses in a general linear model (GLM).

To visualize pattern change throughout the brain, we constructed a mean pattern change map of all collected trials (Figure 4). Additionally, we calculated the mean change in BOLD signal strength of a given searchlight sphere (i.e., pure univariate information of each pattern) of all trials for comparison purposes. This revealed stark differences between these two measures and their distribution throughout the brain: while mean BOLD change is strongest in the vicinity of large vessels (e.g., superior sagittal sinus, medial cerebral artery), mean pattern change was maximal in the superior parietal lobule (SPL), the OFC, and the vmPFC. These findings underline that pattern change and change in mean BOLD responses are indeed different measures with distinct distributions throughout the brain.

PE-evoked BOLD responses modulate subsequent pattern change

In a next step, we extracted trial-wise BOLD responses induced by the SPE and RPE from the error-coding ROIs. These PE-evoked BOLD responses were then used as predictors for subsequent pattern change at every voxel in the brain in a GLM:

$$\text{pattern change} \sim b_0 + b_1 * \text{BOLD response to PE in ROI},$$

where pattern change is a set of trial-by-trial images as created in the previous section and the BOLD response to PE is a set of trial-by-trial estimates of the BOLD response to PEs in the VS ROI (RPE) or in the IPS ROI (SPE). It is important to note that the PE event $\text{PE_response}[t]$ predicting $\text{pattern_change}[t, t+1]$ occurs between the two decision-making events used to calculate the pattern change (Figure 3A). We hypothesized that a large PE should lead to greater adjustments of our internal model, which in turn go along with greater changes in neural activation patterns. In case of the VS, we therefore used absolute estimates of the BOLD response (i.e., both negative and positive RPEs of high magnitude should

lead to an increase in pattern change). The SPE is an unsigned signal and hence needed no correction. We validated the results of this analysis by performing a permutation test in which the real SPE and RPE regressors were compared against 10,000 random permutations (see STAR Methods and Figure S2).

The resulting brain maps of this linear model (Figure 5) reveal areas in which pattern change at the time of decision-making was significantly modulated by the preceding PE-evoked BOLD response in the corresponding error-coding ROI. The RPE-evoked BOLD response in VS modulated pattern change at the time of reward decision in the OFC, the anterior cingulate cortex (ACC), the bilateral dIPFC, and the bilateral insula (OFC: $x, y, z = -28, 28, -20$, $p_{FWE} = 0.002$; ACC: $x, y, z = 13, 38, 10$, $p_{FWE} = 0.003$; dIPFC: $x, y, z = 47, 20, 15$, $p_{FWE} = 0.001$; insula: $x, y, z = -35, -2, -15$, $p_{FWE} = 0.003$), whereas the SPE-evoked BOLD response in IPS predicted pattern change at the time of state decision in the SPL, the PCu, the supplementary motor area (SMA), and the IPS (SPL: $x, y, z = -23, -42, 45$, $p_{FWE} = 0.002$; PCu: $x, y, z = 8, -77, 50$, $p_{FWE} = 0.004$; SMA: $x, y, z = 10, 0, 62$, $p_{FWE} = 0.003$; IPS: $x, y, z = -55, -42, 42$, $p_{FWE} = 0.011$). These results point to an immediate effect (over the course of a single trial, duration approx. 30 s) of PEs in the IPS and the VS on the amount of change in subsequent neural activation patterns in distinct areas of the brain, underlining that these different PEs are not only represented in different brain regions but that they also evoke pattern changes in dissociable brain networks.

Error-specific dissociation of pattern change

Additionally, we visualized the evolution of pattern change at the time of state and reward reversals (Figure 6). The OFC exhibited a sharp increase in the amount of change in neural activation patterns at the time of reward decision following the reversal of reward contingencies, whereas pattern change at state decision did not respond to reward reversals (paired t test of pattern change at reward decision vs. pattern change at state decision: 1 trial after reward reversal $p = 0.0128$, 2 trials after reward reversal $p = 0.0002$). Interestingly, the increase in pattern change following

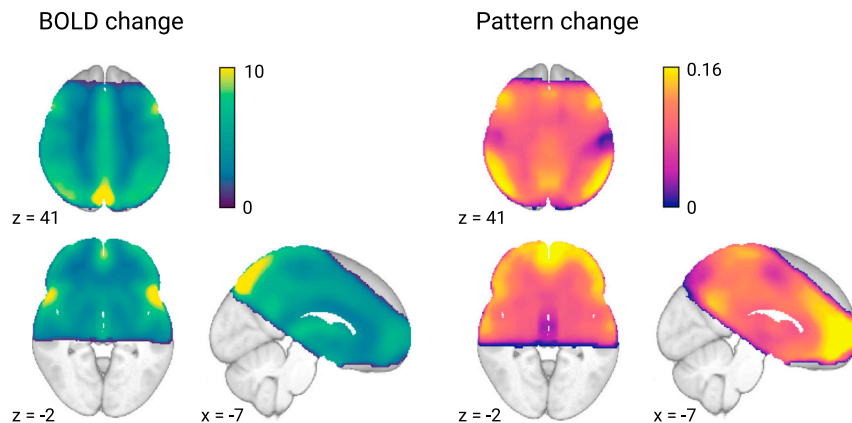


Figure 4. Change in mean BOLD signal vs. pattern change

Whole-brain maps averaged across all trials (145 trials per participant) and participants ($n = 44$) were created for comparison between pattern change (metric of changes in multivariate neural patterns over time) and change in mean BOLD signal of a given searchlight (i.e., pure univariate information of each pattern).

Pattern change predicts ensuing adaptations of behavioral policy

PEs in IPS and VS modulate the amount of change in neural activation patterns during learning. Computationally, PEs are

reward reversals was also apparent in the trial after reversal, indicating sustained changes in neural patterns before decreasing again. In the PCu, we observed the opposite pattern, albeit not as sustained as in the OFC: following a state reversal, pattern change in the PCu at the time of state decision exhibited an increase, while pattern change at the time of reward decision decreased (paired t test of pattern change at state decision vs. pattern change at reward decision: 1 trial after state reversal $p = 0.0132$). These findings also emphasize the dissociation of increased pattern change induced by the two different PE types.

Information content of observed patterns

To investigate the information content of these evolving neural patterns, we utilized additional behavioral decision variables derived from the hybrid-learning model. Specifically, we searched for voxels coding for the value of the chosen action at the time of reward decision (i.e., the same time at which reward-related patterns were acquired). The results revealed significant clusters coding for the value of the chosen action in the ACC and the vmPFC (see Figure S3; ACC: $x, y, z = 1, 12, -10, p_{FWE} = 0.023$; vmPFC: $x, y, z = 1, 52, -7, p_{FWE} = 0.036$). A conjunction analysis (Figure 5E) found evidence that the BOLD response of a cluster of voxels in the subgenual ACC ($x, y, z = 2, 10, -12, p_{unc.} = 0.00037$) both encoded the value of the chosen action and featured RPE-modulated pattern change at the time of reward decision, suggesting that ACC pattern change is involved in configuring the value signal that guides action selection.

Furthermore, to detect state-related information, we performed a searchlight-based decoding analysis of participants' expectations of the next upcoming state at the time of state decision. Neural patterns in the PCu contained significant information about the next expected state (see Figure S4; state expectations in the PCu: $x, y, z = 2, -61, 37, p_{FWE} = 0.025$). However, these results did not exhibit a meaningful overlap with patterns modulated by SPE-evoked BOLD responses (Figure 5E). In summary, our custom-made GLM analyses detected several distinct brain regions in which pattern change was predicted by the PE-evoked BOLD response in the IPS and the VS, respectively, but only a subset of those were also involved in representing expectation signals for states and rewards calculated by the computational model.

a teaching signal used to inform and update our internal representations of the world. Consequently, adjustments made to our internal model should also lead to altered behavior in order to maximize rewards. Therefore, we tested whether the observed changes in neural patterns over time do also predict changes in the participants' actions. Due to the binary nature of choices in this task, we utilized the continuous action probabilities derived from the hybrid-learning model to separately calculate changes in the behavioral policy ($\Delta policy$) of state and reward decisions (A1 being the participants' action following the state decision and A2 being that following the reward decision). Indeed, relating pattern change at state decision to trial-by-trial changes in action probabilities of A1 revealed significant clusters in the primary motor cortex (M1), the vmPFC, the insula, and the SMA (see Figure S5; M1: $x, y, z = 35, -15, 53, p_{FWE} = 0.001$; vmPFC: $x, y, z = 12, 57, -6, p_{FWE} = 0.001$; insula: $x, y, z = 35, -11, -10, p_{FWE} = 0.001$; SMA: $x, y, z = -12, 10, 52, p_{FWE} = 0.002$). Furthermore, we detected clusters in the vmPFC, the OFC, the M1, and the dlPFC in which pattern change at reward decision significantly predicted changes in the behavioral policy of A2 (vmPFC: $x, y, z = -15, 43, 20, p_{FWE} = 0.001$; OFC: $x, y, z = -25, 25, -18, p_{FWE} = 0.001$; M1: $x, y, z = 30, -16, 49, p_{FWE} = 0.001$; dlPFC: $x, y, z = -57, 15, 12, p_{FWE} = 0.002$).

Importantly, repeating this analysis using the change in mean BOLD signal of a given searchlight (cf. Figure 4) did not result in any brain regions showing a significant correlation to changes in the participants' actions. This suggests that, indeed, the multivariate information contained in pattern changes (and not the univariate change in mean BOLD signal of a given pattern) holds information relevant for updating the behavioral policy.

To pinpoint areas featuring changes in neural activation patterns that are both modulated by PE-evoked BOLD responses and predict ensuing behavioral adaptations, we performed conjunction analyses.³⁵ Results of these conjunctions are presented separately for SPE- and RPE-related pattern change (Figure 7). We found a significant cluster of voxels in the OFC in which pattern change at reward decision was both modulated by RPE-evoked BOLD in the VS and predicted subsequent behavioral adjustments of A2 (OFC: $x, y, z = -27, 24, -21, p_{FWE} = 0.032$). In addition, we found clusters in the ACC and the dlPFC exhibiting the same conjunction, but they did not surpass an FWE-corrected threshold (ACC: $x, y, z = 3, 40,$

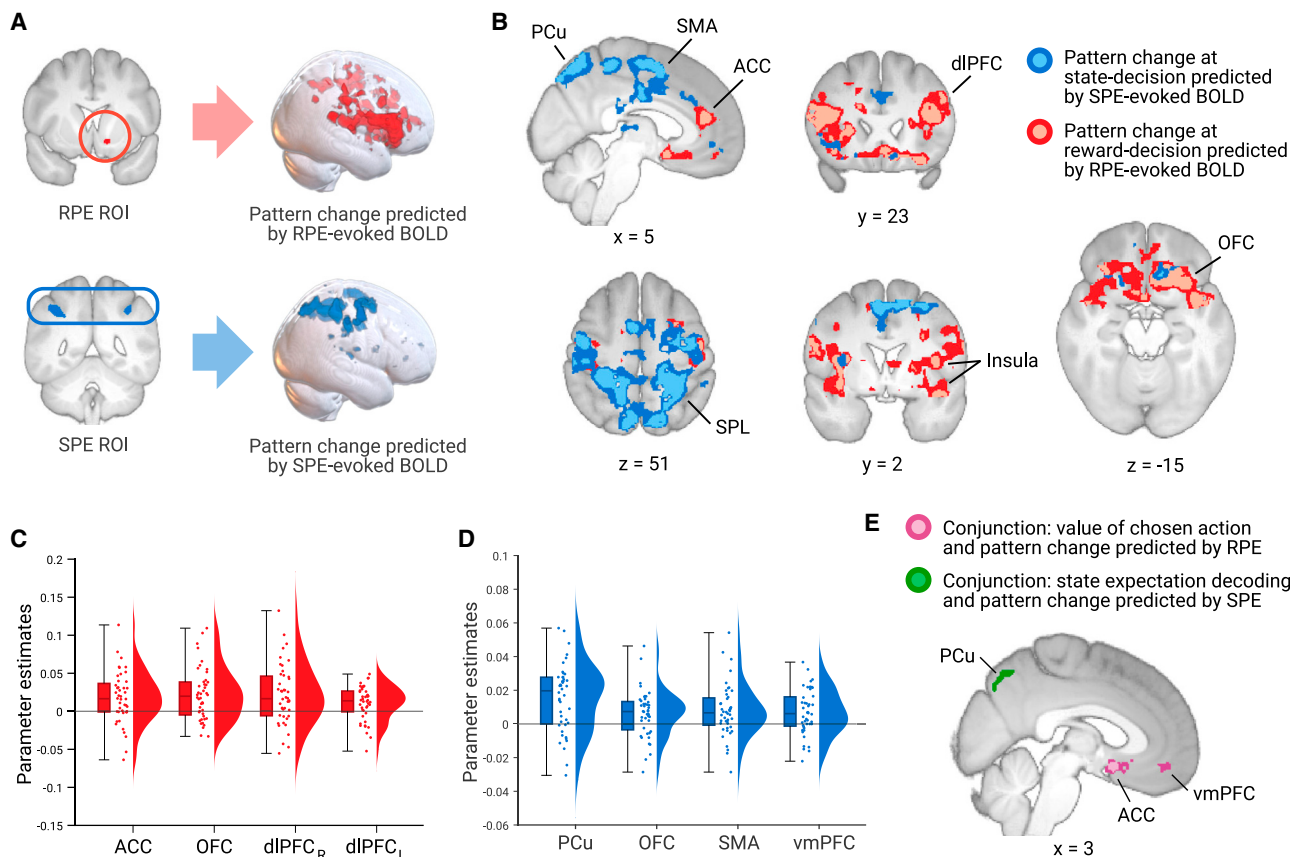


Figure 5. PE-evoked BOLD responses in error-coding ROIs modulate pattern change

(A) PE-evoked BOLD responses were extracted from error-coding ROIs and served as predictors for subsequent pattern change throughout the whole brain in a linear regression (statistical maps are presented at a threshold of $p < 0.001$ uncorrected).

(B) Results of a trial-by-trial regression of pattern change on PE-evoked BOLD responses in error-coding ROIs. Results are presented at a threshold of $p < 0.005$ uncorrected (dark area) and $p < 0.001$ uncorrected (light area).

(C and D) Parameter estimates of peak voxels derived from the trial-by-trial regression model (as in B). Boxplots visualize the median, 25th, and 75th percentiles as well as minimum and maximum values of the data. Data points represent parameter estimates of individual participants. Half-violin plots visualize the shape of data distributions.

(E) Violet: conjunction results combining the values of chosen action at reward decision and RPE-modulated pattern change at the time of reward decision. Green: conjunction results combining the decoding of state expectations and SPE-modulated pattern change at the time of state decision. Results are presented at a threshold of $p < 0.005$ uncorrected (dark area) and $p < 0.001$ uncorrected (light area).

19, $p_{\text{unc.}} = 0.00086$; dIPFC: $x, y, z = 49, 15, 17$, $p_{\text{unc.}} = 0.00071$). In the conjunction of pattern change at state decision correlating with both SPE-evoked BOLD in the IPS and subsequent changes in participants' behavioral policy of A1, we observed clusters in the vmPFC and the bilateral motor cortex (vmPFC: $x, y, z = 10, 48, -9$, $p_{\text{unc.}} = 0.00093$; M1: $x, y, z = 39, -19, 48$, $p_{\text{unc.}} = 0.00031$).

In summary, these results suggest that high amounts of pattern change are induced by strong PE-evoked BOLD responses in error-coding ROIs and that they lead to adaptations in behavior. To underline this finding, we divided trials into bins of low (<10th percentile) and high (>90th percentile) pattern change at peak voxels in the OFC and the vmPFC (Figure 7). For trials with high pattern change at reward decision in OFC, we observed significantly greater preceding RPE-evoked absolute BOLD responses in the VS and more pronounced changes in behavioral policy of A2 (paired t test of RPE-evoked absolute

BOLD response [abs. BOLD resp.] in the VS for high vs. low pattern change: $p = 0.0005$; paired t test of $\Delta policy$ of A2 for high vs. low pattern change: $p = 0.0009$). When comparing trials of high vs. low pattern change at state decision in the vmPFC, we found increased preceding SPE-evoked BOLD responses in the IPS as well as greater adaptations of participants' behavior in A1 (paired t test of SPE-evoked BOLD resp. in the IPS for high vs. low pattern change: $p = 0.0094$; paired t test of $\Delta policy$ of A1 for high vs. low pattern change: $p = 0.0292$).

DISCUSSION

An implementation of RL in the human brain requires the existence of neural representations that are being updated by PEs. Although being a crucial part of learning, the mechanisms of transforming neural representations have been sparsely studied.

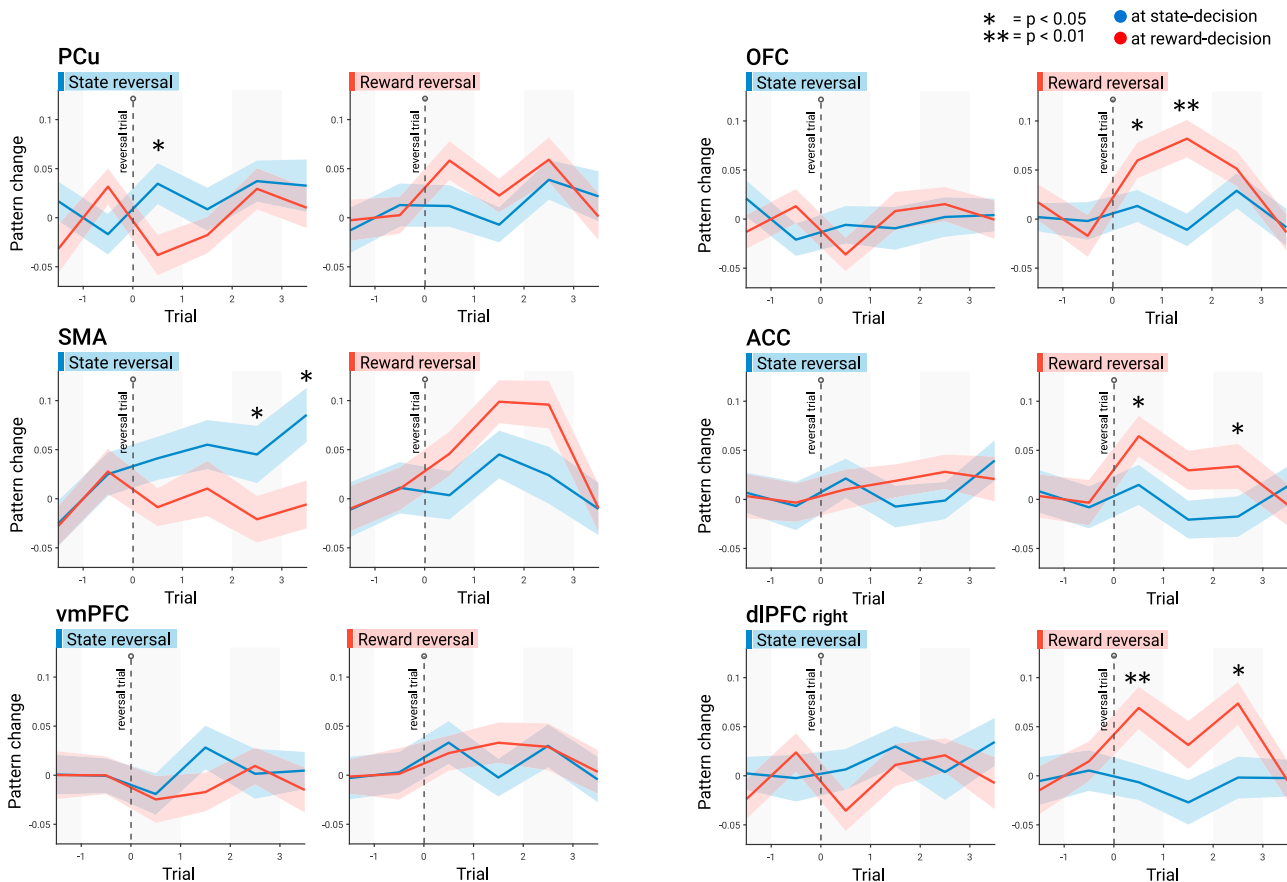


Figure 6. Pattern change progression surrounding state and reward reversals

For illustration, pattern change magnitude at peak voxels of the ROIs is plotted for trials surrounding state and reward reversals. Blue lines represent pattern change of adjacent trials at state decision and red lines at reward decision. Shaded areas depict standard error of the mean. Pattern change occurring between corresponding events of trial t and $t + 1$ is plotted between those trials. Dashed lines represent reversal trials. Pattern change values were baseline corrected using the mean of two trials leading up to state and reward reversals (the task did not include catch trials in an interval of 2 trials surrounding reversal trials, to achieve more stable representations before and after reversals). Stars indicate significant differences in a paired t test (* $p < 0.05$ and ** $p < 0.01$) of pattern change between the blue curve (at state decision) vs. the red curve (at reward decision).

Here, we show that the magnitude of short-term changes in neural activation patterns is driven by PEs. Importantly, we utilized two distinct PE types, model-based SPEs and model-free RPEs, and showed specific effects in two learning systems. Furthermore, we identified brain regions in which the amount of pattern change also predicted ensuing changes in participants' behavioral policy for the corresponding decision that was triggered by the specific PE. In summary, our results suggest that PEs drive the reconfiguration of dynamic neural activation patterns that encode our internal expectation models and guide future behavior.

Our findings reveal that the two different PE types were not only represented in different regions (VS and IPS) but also that they influenced distinct areas of the brain. We demonstrated that pattern change in the OFC, the dlPFC, and the ACC was modulated by RPEs. These brain regions are consistent with the current understanding of reward-related processes. The OFC encodes representations of expected rewards in patterns of neural activity,^{12,36–38} and the dlPFC has been shown to

code for the value of individual attributes of a stimulus.³⁹ The activity of the dlPFC also increases when context changes require a reweighting of values.^{40,41} Furthermore, the ACC holds reward predictions during decision-making,^{42,43} reacts to reward outcome, and is relevant for behavioral adjustments.⁴⁴ Our observations of the ACC encoding the value of the chosen action endorse the central role of the ACC for the decision-making process.

RPE signals are being distributed throughout the brain via fluctuations in the concentration of dopamine.²⁰ Recent work in the field suggests that mesolimbic dopamine might even convey causal associations during learning,⁴⁵ and another study observed heterogeneous dopamine responses even at the evidence accumulation phase (before a decision is made).⁴⁶ This has prompted a differentiated extension of the classic RPE hypothesis:⁴⁷ whereas the standard model posits a scalar RPE at the time of outcome, this new model favors a vector-valued RPE that encodes feature-specific expectancy violations. However, the authors also observed in their deep RL model that the

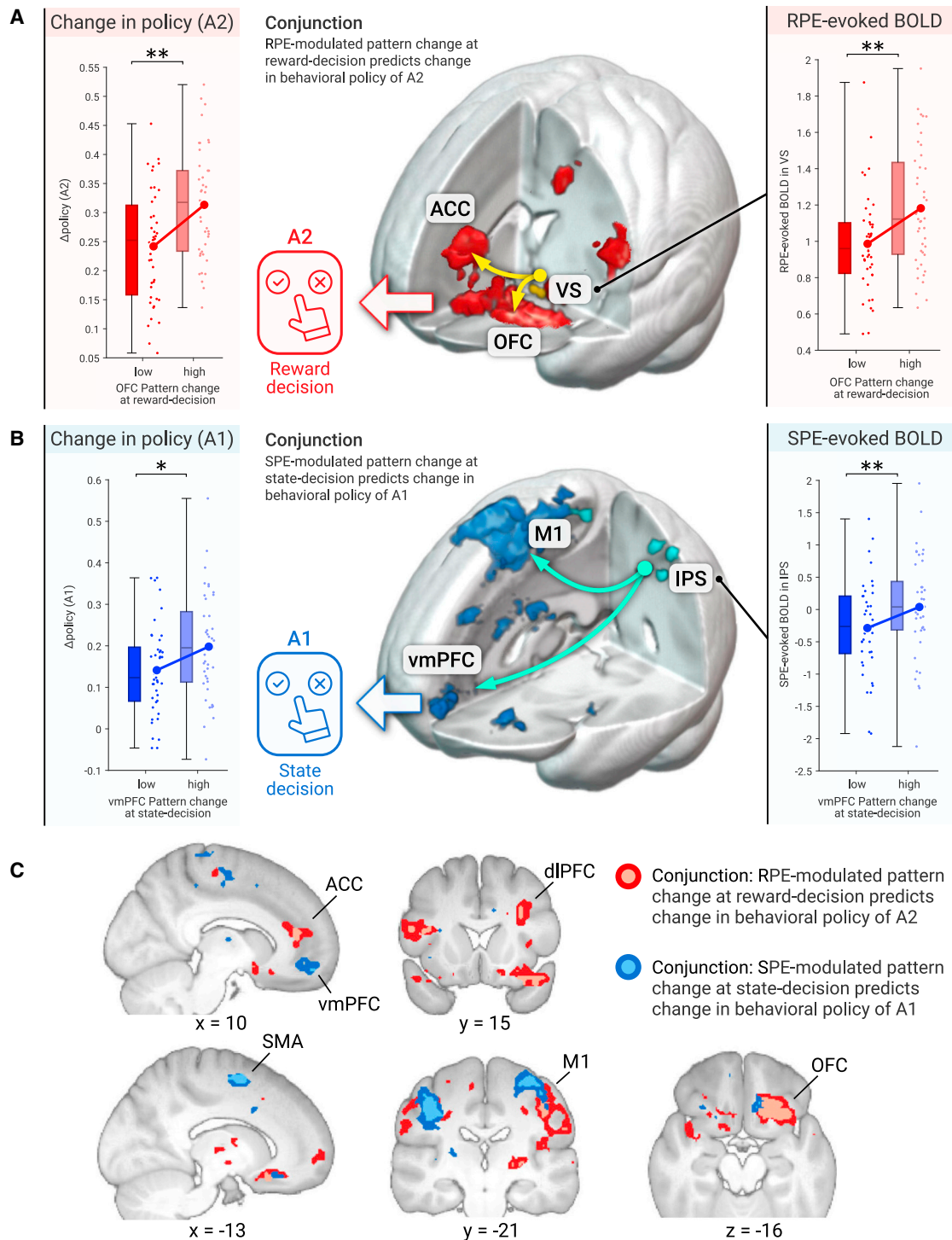


Figure 7. Pattern change modulated by PE-evoked BOLD responses predicts ensuing adjustments of the behavioral policy

Results generated by a conjunction combining pattern change modulated by PE-evoked BOLD responses in error-coding ROIs and pattern change predicting ensuing changes in behavioral policy.

(A and B) Yellow and cyan arrows connect error-coding ROIs with clusters identified in the conjunction. Trials were divided into bins of low (<10th percentile) and high (>90th percentile) pattern change in peak voxels of the OFC and the vmPFC. For trials with low vs. high pattern change, we plotted RPE- and SPE-evoked BOLD responses in the VS and the IPS. Due to RPE being represented as a signed prediction error we used the absolute parameter estimates of VS. Additionally,

(legend continued on next page)

vector-value RPE at the time of outcome resembles a scalar RPE. Therefore, the (scalar) RPE used to detect seed ROIs for the pattern change analysis are not invalidated by these recent theoretical developments. A possible biochemical substrate of the observed effects of RPEs on changes in neural activity patterns could be dopamine's ability to modulate short-term spike-timing-dependent plasticity (STDP).^{48–50} Our findings support the idea that PEs are involved in priming brain regions for subsequent functional reconfigurations.

In contrast, we identified SPE-modulated pattern change in the PCu and the SPL. Recent work established that the PCu is involved in conscious processing of external information based on internal beliefs and that it also updates our internal model of the environment.⁵¹ Additionally, the PCu contributes to attentional shifting between different internal representations.⁵² We found that the pattern change in a subset of OFC voxels was also modulated by the SPE, which is consistent with the idea that the OFC holds a cognitive map about the state space of a task and that information about the current state can be decoded from patterns of neural activity in the OFC.¹³ Our findings also broadly agree with the cortico-basal ganglia-thalamus-cortex (CBGTC) loop, which instantiates a direct pathway commonly associated with the Go response (i.e., action selection) following positive PEs and an indirect pathway that is usually associated with NoGo responses (i.e., response inhibition) following negative PEs.^{53,54} The network defined by these regions is organized into several parallel pathways that connect the basal ganglia with the sensorimotor, associative, and limbic cortices, the latter including the OFC and the ACC,^{55,56} which are also involved in flexible learning of expected values of stimuli and/or states.⁵⁷ One of the primary findings of this study, namely the prediction of pattern change in the OFC and the ACC that represents an expected reward value by BOLD activity in the VS, which encodes an RPE, is in support of the CBGTC loop and the associated cortico-basal ganglia GoNoGo model of decision-making.⁵³

In RL, expectancy violations elicit modifications of our internal model of the environment, leading to altered behavior. Here, we demonstrated that the amount of pattern change in the OFC and the vmPFC predicts subsequent adaptations of the behavioral policy. These results are in alignment with recent studies, which revealed that value representations in the OFC inform the general value signal constructed in the vmPFC to guide goal-directed behavior.^{58–63} Moreover, pattern change in the M1 and the SMA also correlated with changes in behavioral policy. These effects were presumably caused by alterations in action planning and execution and were therefore also accompanied by changes in neural activity patterns.

To our knowledge, there are a few other reports of investigating evolving multivoxel patterns of BOLD activity during learning. Visser and colleagues reported an increase in pattern similarity in the ACC, the amygdala, and the superior frontal gyrus during fear conditioning, when formerly neutral stimuli

were negatively reinforced.⁶⁴ Nassar and colleagues observed changes in multivariate activity patterns during periods of rapid behavioral change and found that these “network resets” can be explained by the degree of uncertainty.¹⁶ Howard and Kahnt identified reward-identity representations in the OFC and demonstrated that transitions of one representation into another were directly related to midbrain identity-based error signals.¹⁵ In the present study, we present a comprehensive account of the interactions among PE signals, neural activity patterns, and behavioral adaptations. Moreover, by leveraging a two-step task eliciting two types of PEs, we were able to dissociate the effects of each PE on subsequent pattern change at the time of decision-making.

The use of multivariate approaches has facilitated the detection of neural representations encoding state-^{13,65} and reward-related information.^{14,58} In the present study, we primarily focused on the temporal changes these neural activation patterns undergo during learning. Multivariate patterns in fMRI data can be considered complex spatiotemporal filters sampling spatially distributed neural activity through the fine-grained architecture of vessels inside the tissue.⁶⁶ When examining the magnitude of change in patterns of the BOLD response to the same event in adjacent trials, we effectively capture how much the underlying distributed neural activity differs. Combined with the general assumption that a great portion of neural information is encoded in the multivariate interactions of neurons,^{67–69} our proposed pattern change metric is capable of quantifying the differences in neural information content over time. Other complementary analyses (e.g., representational similarity analyses or decoding analyses as conducted here) performed on the same pattern data can then infer what the observed neural information represents.

We present an approach of directly relating multivariate and univariate features of fMRI data through the use of trial-by-trial searchlight-algorithms. By compressing multivariate features of fMRI data into univariate signals and incorporating those in whole-brain GLMs, we were able to dynamically regress trial-by-trial changes in neural activation patterns against preceding trial-by-trial, PE-evoked BOLD responses in error-coding regions. Additionally, we show that the sole univariate information of patterns (i.e., increase in mean neural activity of a pattern) is not sufficient to predict behavioral adaptations but that, instead, the multivariate information (i.e., spatially distributed pattern of neural activity) is required.

Limitations of the study

An alternative interpretation of our findings could attribute the observed changes in neural patterns to an enhanced allocation of attentional resources. Surprising events, such as PEs, often trigger an increase in activation within the fronto-parietal attention network.⁷⁰ The greater demand for attentional resources may arise from both positive and negative PEs, irrespective of their sign. This correlation with unsigned RPEs suggests the

we display the adaptations in behavioral policy $\Delta policy$ of A1 and A2 for low vs. high pattern change trials. Boxplots visualize the median, 25th, and 75th percentiles as well as minimum and maximum values of the data. Data points represent parameter estimates of individual participants. Connecting lines are drawn between the group-level means. Stars indicate significant differences ($p < 0.05$ and $**p < 0.01$) of data in low vs. high pattern change trials assessed using paired t tests. (C) Two-dimensional slices of the same conjunction results. All results are presented at a threshold of $p < 0.005$ uncorrected (dark area) and $p < 0.001$ uncorrected (light area).

possibility of confounding attentional effects. However, by employing a signed RPE in our model-informed univariate analyses to identify relevant brain regions, we have taken steps to partially mitigate the confounding influence of attention. Future studies could also employ an independent task specifically designed to measure attentional allocation. By doing so, researchers could then estimate the effects of attention on neural pattern change and behavior across different participants, providing additional insights into the role of attention in the observed findings.

While the study reveals correlations between PEs, neural pattern changes, and subsequent behavioral adaptations, it does not establish a definitive causal relationship. The design and experimental methods of the study does not allow for causal conclusions to be drawn. Further studies incorporating interventions (like brain stimulation) are necessary to establish a more robust causal link among PEs, neural patterns, and behavior.

Conclusion

In summary, our work provides insights on the mechanisms through which PEs facilitate the short-term reconfiguration of neural representations and, consequently, provoke a shift in behavioral policy. In addition to the detailed characterization of the temporal evolution of pattern change, we could also show that parts of these neural representations encode information about expectancy signals. This dual perspective on neural patterns—the temporal changes during reconfigurations of these representations and the information that they encode—also casts new light on the nature of neural activation patterns in the context of a coherent theoretical framework that describes learning as the process of updating internal models to enable flexible and more accurate predictions of future events.

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 STUDY PARTICIPANT DETAILS**
- **METHOD DETAILS**
 - Experimental task
 - Behavioral data acquisition
 - fMRI data acquisition
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Computational modeling
 - fMRI data preprocessing
 - Univariate analyses of prediction error signals
 - Pattern change searchlight analysis
 - Predicting pattern change using PE-evoked BOLD
 - Permutation test of pattern change regression
 - Relating pattern change to behavioral policy
 - Detecting the value of the chosen action

- Decoding of state expectations
- Group-level statistical analyses

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.celrep.2023.112931>.

ACKNOWLEDGMENTS

We thank Christian Büchel and Tobias Sommer for helpful comments on earlier versions of this paper. This work is funded by the German Research Foundation (SFB TRR 169 “Cross-modal Learning”).

AUTHOR CONTRIBUTIONS

J.G. and L.M. designed the experimental task, developed analytical protocols, and wrote the paper. L.M. acquired the data and conducted the statistical analyses.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 31, 2023

Revised: June 13, 2023

Accepted: July 18, 2023

Published: August 3, 2023

REFERENCES

1. Sutton, R.S., and Barto, A.G. (2018). *Reinforcement Learning: An Introduction, Second Edition* (MIT Press).
2. Dayan, P., and Daw, N.D. (2008). Decision theory, reinforcement learning, and the brain. *Cognit. Affect Behav. Neurosci.* 8, 429–453. <https://doi.org/10.3758/CABN.8.4.429>.
3. Niv, Y. (2009). Reinforcement learning in the brain. *J. Math. Psychol.* 53, 139–154. <https://doi.org/10.1016/j.jmp.2008.12.005>.
4. Kahnt, T. (2018). A decade of decoding reward-related fMRI signals and where we go from here. *Neuroimage* 180, 324–333. <https://doi.org/10.1016/j.neuroimage.2017.03.067>.
5. Haynes, J.-D. (2015). A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. *Neuron* 87, 257–270. <https://doi.org/10.1016/j.neuron.2015.05.025>.
6. Mahmoudi, A., Takerkart, S., Regragui, F., Boussaoud, D., and Brovelli, A. (2012). Multivoxel pattern analysis for fMRI data: a review. *Comput. Math. Methods Med.* 2012, 961257. <https://doi.org/10.1155/2012/961257>.
7. Liu, Y., Nour, M.M., Schuck, N.W., Behrens, T.E.J., and Dolan, R.J. (2022). Decoding cognition from spontaneous neural activity. *Nat. Rev. Neurosci.* 23, 204–214. <https://doi.org/10.1038/s41583-022-00570-z>.
8. Wittkuhn, L., and Schuck, N.W. (2021). Dynamics of fMRI patterns reflect sub-second activation sequences and reveal replay in human visual cortex. *Nat. Commun.* 12, 1795. <https://doi.org/10.1038/s41467-021-21970-2>.
9. Wilbertz, G., van Kemenade, B.M., Schmack, K., and Sterzer, P. (2017). fMRI-based decoding of reward effects in binocular rivalry. *Neurosci. Conscious.* 2017, nix013. <https://doi.org/10.1093/nc/nix013>.
10. Glitz, L., Juechems, K., Summerfield, C., and Garrett, N. (2022). Model Sharing in the Human Medial Temporal Lobe. *J. Neurosci.* 42, 5410–5426. <https://doi.org/10.1523/JNEUROSCI.1978-21.2022>.
11. Freund, M.C., Etzel, J.A., and Braver, T.S. (2021). Neural Coding of Cognitive Control: The Representational Similarity Analysis Approach. *Trends Cognit. Sci.* 25, 622–638. <https://doi.org/10.1016/j.tics.2021.03.011>.

12. Yan, C., Su, L., Wang, Y., Xu, T., Yin, D.-Z., Fan, M.-X., Deng, C.-P., Hu, Y., Wang, Z.-X., Cheung, E.F.C., et al. (2016). Multivariate Neural Representations of Value during Reward Anticipation and Consummation in the Human Orbitofrontal Cortex. *Sci. Rep.* 6, 29079. <https://doi.org/10.1038/srep29079>.
13. Schuck, N.W., Cai, M.B., Wilson, R.C., and Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron* 91, 1402–1412. <https://doi.org/10.1016/j.neuron.2016.08.019>.
14. Wisniewski, D., Reverberi, C., Momennejad, I., Kahnt, T., and Haynes, J.-D. (2015). The Role of the Parietal Cortex in the Representation of Task-Reward Associations. *J. Neurosci.* 35, 12355–12365. <https://doi.org/10.1523/JNEUROSCI.4882-14.2015>.
15. Howard, J.D., and Kahnt, T. (2018). Identity prediction errors in the human midbrain update reward-identity expectations in the orbitofrontal cortex. *Nat. Commun.* 9, 1611. <https://doi.org/10.1038/s41467-018-04055-5>.
16. Nassar, M.R., McGuire, J.T., Ritz, H., and Kable, J.W. (2019). Dissociable Forms of Uncertainty-Driven Representational Change Across the Human Brain. *J. Neurosci.* 39, 1688–1698. <https://doi.org/10.1523/JNEUROSCI.1713-18.2018>.
17. Tobler, P.N., O’Doherty, J.P., Dolan, R.J., and Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *J. Neurophysiol.* 95, 301–310. <https://doi.org/10.1152/jn.00762.2005>.
18. Schultz, W. (2019). Recent advances in understanding the role of phasic dopamine activity. *F1000Res.* 8. <https://doi.org/10.12688/f1000research.19793.1>.
19. Lerner, T.N., Holloway, A.L., and Seiler, J.L. (2021). Dopamine, Updated: Reward Prediction Error and Beyond. *Curr. Opin. Neurobiol.* 67, 123–130. <https://doi.org/10.1016/j.conb.2020.10.012>.
20. Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. <https://doi.org/10.1126/science.275.5306.1593>.
21. Waelti, P., Dickinson, A., and Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412, 43–48. <https://doi.org/10.1038/35083500>.
22. Morales, M., and Margolis, E.B. (2017). Ventral tegmental area: cellular heterogeneity, connectivity and behaviour. *Nat. Rev. Neurosci.* 18, 73–85. <https://doi.org/10.1038/nrn.2016.165>.
23. Breton, J.M., Charbit, A.R., Snyder, B.J., Fong, P.T.K., Dias, E.V., Himmels, P., Lock, H., and Margolis, E.B. (2019). Relative contributions and mapping of ventral tegmental area dopamine and GABA neurons by projection target in the rat. *J. Comp. Neurol.* 527, 916–941. <https://doi.org/10.1002/cne.24572>.
24. 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. [https://doi.org/10.1016/s0896-6273\(03\)00169-7](https://doi.org/10.1016/s0896-6273(03)00169-7).
25. Jocham, G., Klein, T.A., and Ullsperger, M. (2011). Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *J. Neurosci.* 31, 1606–1613. <https://doi.org/10.1523/JNEUROSCI.3904-10.2011>.
26. Garrison, J., Erdeniz, B., and Done, J. (2013). Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neurosci. Biobehav. Rev.* 37, 1297–1310. <https://doi.org/10.1016/j.neubiorev.2013.03.023>.
27. Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645. <https://doi.org/10.1126/science.1105370>.
28. Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902. <https://doi.org/10.1126/science.1077349>.
29. Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711. <https://doi.org/10.1038/nn1560>.
30. Gläscher, J., Daw, N., Dayan, P., and O’Doherty, J.P. (2010). States versus Rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66, 585–595. <https://doi.org/10.1016/j.neuron.2010.04.016>.
31. Lee, S.W., Shimojo, S., and O’Doherty, J.P. (2014). Neural computations underlying arbitration between model-based and model-free learning. *Neuron* 81, 687–699. <https://doi.org/10.1016/j.neuron.2013.11.028>.
32. Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-based influences on humans’ choices and striatal prediction errors. *Neuron* 69, 1204–1215. <https://doi.org/10.1016/j.neuron.2011.02.027>.
33. Feher da Silva, C., and Hare, T.A. (2018). A note on the analysis of two-stage task results: How changes in task structure affect what model-free and model-based strategies predict about the effects of reward and transition on the stay probability. *PLoS One* 13, e0195328. <https://doi.org/10.1371/journal.pone.0195328>.
34. Smith, S.M., and Nichols, T.E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44, 83–98. <https://doi.org/10.1016/j.neuroimage.2008.03.061>.
35. Nichols, T., Brett, M., Andersson, J., Wager, T., and Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage* 25, 653–660. <https://doi.org/10.1016/j.neuroimage.2004.12.005>.
36. Li, Y., Vanni-Mercier, G., Isnard, J., Mauguière, F., and Dreher, J.-C. (2016). The neural dynamics of reward value and risk coding in the human orbitofrontal cortex. *Brain* 139, 1295–1309. <https://doi.org/10.1093/brain/awv409>.
37. Klein-Flügge, M.C., Barron, H.C., Brodersen, K.H., Dolan, R.J., and Behrens, T.E.J. (2013). Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. *J. Neurosci.* 33, 3202–3211. <https://doi.org/10.1523/JNEUROSCI.2532-12.2013>.
38. Hare, T.A., O’Doherty, J., Camerer, C.F., Schultz, W., and Rangel, A. (2008). Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J. Neurosci.* 28, 5623–5630. <https://doi.org/10.1523/JNEUROSCI.1309-08.2008>.
39. 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. <https://doi.org/10.1016/j.neuroimage.2010.05.058>.
40. Klüen, L.M., Dandolo, L.C., Jocham, G., and Schwabe, L. (2019). Dorsolateral Prefrontal Cortex Enables Updating of Established Memories. *Cerebr. Cortex* 29, 4154–4168. <https://doi.org/10.1093/cercor/bhy298>.
41. Rudolf, S., and Hare, T.A. (2014). Interactions between dorsolateral and ventromedial prefrontal cortex underlie context-dependent stimulus valuation in goal-directed choice. *J. Neurosci.* 34, 15988–15996. <https://doi.org/10.1523/JNEUROSCI.3192-14.2014>.
42. Vassena, E., Krebs, R.M., Silvetti, M., Fias, W., and Verguts, T. (2014). Dissociating contributions of ACC and vmPFC in reward prediction, outcome, and choice. *Neuropsychologia* 59, 112–123. <https://doi.org/10.1016/j.neuropsychologia.2014.04.019>.
43. Silvetti, M., Seurinck, R., and Verguts, T. (2013). Value and prediction error estimation account for volatility effects in ACC: a model-based fMRI study. *Cortex* 49, 1627–1635. <https://doi.org/10.1016/j.cortex.2012.05.008>.
44. Becker, M.P.I., Nitsch, A.M., Hewig, J., Miltner, W.H.R., and Straube, T. (2016). Parametric modulation of reward sequences during a reversal task in ACC and VMPFC but not amygdala and striatum. *Neuroimage* 143, 50–57. <https://doi.org/10.1016/j.neuroimage.2016.09.022>.
45. Jeong, H., Taylor, A., Floeder, J.R., Lohmann, M., Mihalas, S., Wu, B., Zhou, M., Burke, D.A., and Namboodiri, V.M.K. (2022). Mesolimbic

- dopamine release conveys causal associations. *Science* 378, eabq6740. <https://doi.org/10.1126/science.abq6740>.
46. Engelhard, B., Finkelstein, J., Cox, J., Fleming, W., Jang, H.J., Ornelas, S., Koay, S.A., Thiberge, S.Y., Daw, N.D., Tank, D.W., and Witten, I.B. (2019). Specialized coding of sensory, motor, and cognitive variables in VTA dopamine neurons. *Nature* 570, 509–513. <https://doi.org/10.1038/s41586-019-1261-9>.
 47. Lee, R.S., Engelhard, B., Witten, I.B., and Daw, N.D. (2022). A Vector Reward Prediction Error Model Explains Dopaminergic Heterogeneity. <https://doi.org/10.1101/2022.02.28.482379>.
 48. Condon, M.D., Platt, N.J., Zhang, Y.-F., Roberts, B.M., Clements, M.A., Vietti-Michelina, S., Tseu, M.-Y., Brimblecombe, K.R., Threlfell, S., Mann, E.O., and Cragg, S.J. (2019). Plasticity in striatal dopamine release is governed by release-independent depression and the dopamine transporter. *Nat. Commun.* 10, 4263. <https://doi.org/10.1038/s41467-019-12264-9>.
 49. Brzosko, Z., Zannone, S., Schultz, W., Clopath, C., and Paulsen, O. (2017). Sequential neuromodulation of Hebbian plasticity offers mechanism for effective reward-based navigation. *Elife* 6, e27756. <https://doi.org/10.7554/eLife.27756>.
 50. Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G.C.R., Urakubo, H., Ishii, S., and Kasai, H. (2014). A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science* 345, 1616–1620. <https://doi.org/10.1126/science.1255514>.
 51. Lyu, D., Pappas, I., Menon, D.K., and Stamatakis, E.A. (2021). A Precuneal Causal Loop Mediates External and Internal Information Integration in the Human Brain. *J. Neurosci.* 41, 9944–9956. <https://doi.org/10.1523/JNEUROSCI.0647-21.2021>.
 52. Bzdok, D., Heeger, A., Langner, R., Laird, A.R., Fox, P.T., Palomero-Gallagher, N., Vogt, B.A., Zilles, K., and Eickhoff, S.B. (2015). Subspecialization in the human posterior medial cortex. *Neuroimage* 106, 55–71. <https://doi.org/10.1016/j.neuroimage.2014.11.009>.
 53. Frank, M.J. (2005). Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and non-medicated Parkinsonism. *J. Cognit. Neurosci.* 17, 51–72. <https://doi.org/10.1162/0898929052880093>.
 54. Maia, T.V., and Frank, M.J. (2011). From reinforcement learning models to psychiatric and neurological disorders. *Nat. Neurosci.* 14, 154–162. <https://doi.org/10.1038/nn.2723>.
 55. Haber, S.N. (2003). The primate basal ganglia: parallel and integrative networks. *J. Chem. Neuroanat.* 26, 317–330. <https://doi.org/10.1016/j.jchemneu.2003.10.003>.
 56. Postuma, R.B., and Dagher, A. (2006). Basal ganglia functional connectivity based on a meta-analysis of 126 positron emission tomography and functional magnetic resonance imaging publications. *Cerebr. Cortex* 16, 1508–1521. <https://doi.org/10.1093/cercor/bhj088>.
 57. Knudsen, E.B., and Wallis, J.D. (2022). Taking stock of value in the orbitofrontal cortex. *Nat. Rev. Neurosci.* 23, 428–438. <https://doi.org/10.1038/s41583-022-00589-2>.
 58. Howard, J.D., and Kahnt, T. (2017). Identity-Specific Reward Representations in Orbitofrontal Cortex Are Modulated by Selective Devaluation. *J. Neurosci.* 37, 2627–2638. <https://doi.org/10.1523/JNEUROSCI.3473-16.2017>.
 59. Algermissen, J., Swart, J.C., Scheeringa, R., Cools, R., and den Ouden, H.E.M. (2022). Striatal BOLD and Midfrontal Theta Power Express Motivation for Action. *Cerebr. Cortex* 32, 2924–2942. <https://doi.org/10.1093/cercor/bhab391>.
 60. Vaidya, A.R., and Badre, D. (2020). Neural Systems for Memory-based Value Judgment and Decision-making. *J. Cognit. Neurosci.* 32, 1896–1923. https://doi.org/10.1162/jocn_a_01595.
 61. Lee, S., Yu, L.Q., Lerman, C., and Kable, J.W. (2021). Subjective value, not a gridlike code, describes neural activity in ventromedial prefrontal cortex during value-based decision-making. *Neuroimage* 237, 118159. <https://doi.org/10.1016/j.neuroimage.2021.118159>.
 62. Shapiro, A.D., and Grafton, S.T. (2020). Subjective value then confidence in human ventromedial prefrontal cortex. *PLoS One* 15, e0225617. <https://doi.org/10.1371/journal.pone.0225617>.
 63. Dreher, J.-C. (2013). Neural coding of computational factors affecting decision making. *Prog. Brain Res.* 202, 289–320. <https://doi.org/10.1016/B978-0-444-62604-2.00016-2>.
 64. Visser, R.M., Scholte, H.S., Beemsterboer, T., and Kindt, M. (2013). Neural pattern similarity predicts long-term fear memory. *Nat. Neurosci.* 16, 388–390. <https://doi.org/10.1038/nn.3345>.
 65. Schuck, N.W., and Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. *Science* 364, eaaw5181. <https://doi.org/10.1126/science.aaw5181>.
 66. 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. <https://doi.org/10.1016/j.neuroimage.2009.09.059>.
 67. Rissman, J., and Wagner, A.D. (2012). Distributed representations in memory: insights from functional brain imaging. *Annu. Rev. Psychol.* 63, 101–128. <https://doi.org/10.1146/annurev-psych-120710-100344>.
 68. Schlegel, A., Alexander, P., and Tse, P.U. (2016). Information Processing in the Mental Workspace Is Fundamentally Distributed. *J. Cognit. Neurosci.* 28, 295–307. https://doi.org/10.1162/jocn_a_00899.
 69. Anderson, A.J., Kiela, D., Binder, J.R., Feraudino, L., Humphries, C.J., Conant, L.L., Raizada, R.D.S., Grimm, S., and Lalor, E.C. (2021). Deep Artificial Neural Networks Reveal a Distributed Cortical Network Encoding Propositional Sentence-Level Meaning. *J. Neurosci.* 41, 4100–4119. <https://doi.org/10.1523/JNEUROSCI.1152-20.2021>.
 70. Corbetta, M., and Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. <https://doi.org/10.1038/nrn755>.
 71. Piray, P., Dezfouli, A., Heskes, T., Frank, M.J., and Daw, N.D. (2019). Hierarchical Bayesian inference for concurrent model fitting and comparison for group studies. *PLoS Comput. Biol.* 15, e1007043. <https://doi.org/10.1371/journal.pcbi.1007043>.
 72. Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., and Turner, R. (2002). Image distortion correction in fMRI: A quantitative evaluation. *Neuroimage* 16, 217–240. <https://doi.org/10.1006/nimg.2001.1054>.
 73. Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage* 38, 95–113. <https://doi.org/10.1016/j.neuroimage.2007.07.007>.
 74. Mumford, J.A., Turner, B.O., Ashby, F.G., and Poldrack, R.A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage* 59, 2636–2643. <https://doi.org/10.1016/j.neuroimage.2011.08.076>.
 75. Hebart, M.N., Gorgen, K., and Haynes, J.-D. (2014). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Front. Neuroinf.* 8, 88. <https://doi.org/10.3389/fninf.2014.00088>.
 76. Winkler, A.M., Ridgway, G.R., Webster, M.A., Smith, S.M., and Nichols, T.E. (2014). Permutation inference for the general linear model. *Neuroimage* 92, 381–397. <https://doi.org/10.1016/j.neuroimage.2014.01.060>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Behavioral data	This paper	https://doi.org/10.5281/zenodo.8015563
Code	This paper	https://doi.org/10.5281/zenodo.8015561
Software and algorithms		
MATLAB	MathWorks	https://de.mathworks.com/products/MATLAB.html , RRID:SCR_001622
SPM12	The Wellcome Center for Human Neuroimaging	https://www.fil.ion.ucl.ac.uk/spm/ , RRID:SCR_007037
The Decoding Toolbox	Martin Hebart	https://doi.org/10.3389/fninf.2014.00088
cbm toolbox	Payam Piray	https://github.com/payampiray/cbm
MRlcroGL	Chris Rorden	https://www.nitrc.org/projects/mricrogl , RRID:SCR_002403

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jan Gläscher (glaescher@uke.de).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Behavioral data have been deposited at Zenodo and are publicly available as of the date of publication. DOIs are listed in the key resources table. Imaging data reported in this paper will be shared by the [lead contact](#) upon request.
- All original code has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Forty-four healthy participants (23 female; age range, 18–38 years; mean age \pm SD, 24.1 \pm 4.3 years) with no history of psychiatric diseases gave informed written consent to participate in this study. Participants were recruited via advertisement in the university and through participants' referral. No participant was excluded from any analyses and all results presented here are from these 44 participants. One subject stopped the experiment after run 4 due to headache. The study protocol was approved by the Ethics Committee of the Medical Association of Hamburg (2020-10102-B0-ff) and conducted in accordance with the Declaration of Helsinki.

METHOD DETAILS

Experimental task

The experiment was built around a modified two-stage Markov decision task. The task consisted of 5 runs with 29 trials each. A thorough explanation about the task structure and the existence of state- and reward-reversals throughout the task was given to the participants before the experiment. Additionally, participants were instructed that they would receive a bonus of up to 5 € in addition to their hourly compensation depending on their performance during the task. Prior to the start of the 5 main runs, participants performed a short training sequence consisting of 5 example trials to get familiar with the controls and visual task presentation inside the fMRI scanner.

In the first stage of a trial (state-decision, Figure 1), subjects were presented with the choice of pressing either the left or right button (depicted by 2 buttons on the screen and a closed door). Following a delay of 6 s the door opened and, depending on the first-stage action, subjects were taken to either the blue or red state, represented by two blue or red boxes on the screen. After a delay of 6 s symbols appeared on both boxes indicating that subjects were now allowed to input the second action. At the second stage (reward-decision), subjects chose between opening the left or right box and their action was rewarded with money or not (depicted by green text rising above the opened box e.g., “6 €” or yellow text reading “0 €”, the possible reward quantities were 4 €, 6 € or 0 €). After a delay of 6 s the rewards were shown for a duration of 2 s followed by a 4 s inter-trial-interval (depicted by a closed door). Response times by participants were not limited, but they were asked to make prompt decisions (which they did: mean reaction time for state-decision was 1.23 s and for reward-decision 1.01 s). We decided to use unlimited input timings so that we did not have to discard parts of the trial-by-trial evolving neural representations in case of delayed actions.

Every run commenced with a new configuration of state transitions and reward distributions, which had to be learned by the participants. During each run we implemented 2 reversal-trials, one of them being a reversal of state transitions and the other one a reversal of rewards, forcing participants to adapt and update their knowledge. For instance, at the state-reversal trial, if the left button in stage 1 was linked to the blue state, it will from now on be linked to the red state for the rest of this run. At the reward-reversal trial, the content of the two blue boxes got switched among each other and the content of the two red boxes also got switched among those. In a given state, one box was always unrewarded, and the other box contained a reward. Importantly, to elicit orthogonal reward and state prediction errors in subjects, both states had similar rewards associated with them. This way, when a state-reversal occurred and a subject transitioned to the unanticipated state eliciting a large state prediction error, there would be only a small difference in achievable rewards and only small reward prediction errors would be induced. We counterbalanced the order in which state- or reward-reversals occurred.

State transitions as well as rewards were deterministic, but occasional catch-trials were implemented. In catch-trials participants are confronted with a onetime unexpected state or reward outcome only in this trial. Catch-trials made the task more demanding and were also used to facilitate the model-informed detection of prediction error correlates by eliciting a more balanced distribution of prediction error magnitudes. We opted for catch-trials instead of a probabilistic design to retain control over when participants were surprised by unexpected outcomes and to not confuse them on the 2 trials leading up to or following a reversal-trial in the task. Hereby, we obtained more stable neural representations before and after a reversal occurred. Per run we included ~ 3 state catch-trials and ~ 2 reward catch-trials.

The task consisted of free- and forced-choice trials to collect a more balanced set of trials in which participants chose also unfavorable states and rewards. This enabled later decoding of state and reward expectations of our participants. Each run consists of 17 free-choice and 12 forced-choice trials (see Figure 1B). We refrained from placing forced-choice trials in the vicinity of a reversal-trial to prevent interference with the individual decision-making of participants. In forced-choice trials, a green arrow was used to indicate which action had to be performed (free-choice reaction-times: state-decision, 1.58 ± 0.77 s [mean \pm SEM] and reward-decision, 1.07 ± 0.66 s; forced-choice reaction-times: state-decision, 0.87 ± 0.37 s and reward-decision, 0.95 ± 0.40 s).

Behavioral data acquisition

Stimulus presentation, MRI pulse-triggering, and response recording were implemented using MATLAB R2014b (www.mathworks.com) and Psychtoolbox-3 (www.psychtoolbox.org). Participants held a four-button MRI-compatible button box in each hand, of which only one button per hand was active, to choose between the left and right choice options presented on the screen with the respective hand.

fMRI data acquisition

MRI data collection was conducted on a Siemens Prisma 3T scanner (Siemens, Erlangen, Germany) equipped with a 64-channel head coil. Each brain volume consisted of 24 axial slices (voxel size, $2.5 \times 2.5 \times 2.5$ mm³, with 0.5 mm spacing between slices) acquired using a T2*-weighted echoplanar imaging (EPI) protocol (repetition time, TR = 627 ms; echo time, TE = 30 ms; flip angle = 45°; field of view = 224 mm, GRAPPA 2, Multiband 2, Phase encoding direction P >> A). The functional imaging orientation was tilted approximately 30° to the anterior commissure–posterior commissure axis to minimize susceptibility artifacts in OFC. Data for each participant were collected in five runs resulting in a total of ~ 6000 volumes, and the first three volumes of each run were discarded to obtain a steady-state magnetization. In addition, a gradient echo field-map was acquired before EPI scanning to measure the magnetic field inhomogeneity (TE1 = 5.29 ms, TE2 = 7.75 ms), and a high-resolution anatomical image (voxel size, $1 \times 1 \times 1$ mm³) was acquired after the experiment using a T1-weighted MPRAGE protocol.

QUANTIFICATION AND STATISTICAL ANALYSIS

Computational modeling

We modeled participants' behavior using an HYBRID learner, which combines state-action value estimates from a model-based FORWARD learner and a model-free SARSA temporal difference learner.³⁰ Both the FORWARD and SARSA learner had their own distinct learning rates $\alpha_{FORWARD}$ and α_{SARSA} , which remained constant per participant. We derived a first-stage and second-stage RPE using the SARSA learner and a first-stage SPE using the FORWARD learner. Even though the different stages are intended

to dissociate state and reward prediction errors, this does not exclude the possibility that reward predictions change subtly at the first stage (see Figure 1B). In fact, as the agent moves to one of the 2 first-stage states, the overall expected reward at that state changes slightly because of the different rewards available at the subsequent 2nd stage state. However, at the 2nd stage state reward and state coincide because rewards are deterministically linked to one of the 2nd stage states. Therefore, the state expectations are only updated at the 1st stage state. The state-action values $Q_{SARSA}(s, a)$ of the SARSA learner were uniformly initialized with 0.2 at the start of each run. The first-stage RPE $\delta_{RPE,1}$ is computed as:

$$\delta_{RPE,1} = \gamma Q_{SARSA}(s', a') - Q_{SARSA}(s, a)$$

where γ is the temporal discount factor. $Q_{SARSA}(s, a)$ and $Q_{SARSA}(s', a')$ are the state-action values of the current and next state. A reward $r(s)$ could only be obtained in the second-stage (reward-decision) and there were no rewards in the first-stage (state-decision). In the second-stage, due to being the terminal state in the task, there is no further state to be anticipated, so that $\delta_{RPE,2}$ reduces to a delta-rule as:

$$\delta_{RPE,2} = r(s) - Q_{SARSA}(s, a)$$

The RPE is used to update existing state-action values as:

$$Q_{SARSA}(s, a) = Q_{SARSA}(s, a) + \alpha_{SARSA} * \delta_{RPE}$$

The FORWARD learner uses an SPE to update an estimated state transition matrix $T(s, a, s')$ of probabilities to transition from the current state s to s' given the action a . The transition probabilities in T were uniformly initialized with 0.5 at the start of each run. The SPE is calculated after the observed transition to s' and defined as:

$$\delta_{SPE} = 1 - T(s, a, s')$$

The SPE updates the probability of the observed transition via:

$$T(s, a, s') = T(s, a, s') + \alpha_{FORWARD} * \delta_{SPE}$$

and the other possible transition probability is reduced by the same amount to ensure the sum of all transition probabilities given state s and action a remains 1. Estimated transition probabilities are used in combination with the acquired knowledge about rewards stored in Q_{SARSA} to calculate state-action values of the FORWARD learner:

$$Q_{FORWARD}(s, a) = \sum_{s'} T(s, a, s') * \max_{a'} (Q_{SARSA}(s', a'))$$

A free parameter w , which we assumed to be constant across trials, controlled the relative weighting of both models for action selection, resulting in combined state-action values Q_{HYBRID} calculated as:

$$Q_{HYBRID}(s, a) = w * Q_{FORWARD}(s, a) + (1 - w) * Q_{SARSA}(s, a)$$

Actions were selected stochastically according to probabilities determined by transforming Q_{HYBRID} using a softmax function with a constant temperature β , determining how deterministic the choices are. Furthermore, we implemented a free parameter p modeling choice repetition during first-stage actions as proposed by Daw and colleagues³² where $\text{rep}(a)$ indicates the action performed in the previous trial. Thus, action probabilities P are defined as:

$$P(s, a) = \frac{\exp(\beta[Q_{HYBRID}(s, a) + p * \text{rep}(a)])}{\sum_a \exp(\beta[Q_{HYBRID}(s, a_i) + p * \text{rep}(a_i)])}$$

In total the hybrid-learning model contained 6 free parameters (α_{SARSA} , $\alpha_{FORWARD}$, β , γ , w , p). Free- and forced-choice trials were modeled together, because forced-choice trials also contribute to the learning process. Parameter estimation was conducted using hierarchical Bayesian inference as implemented in the computational and behavioral modeling toolbox (cbm).⁷¹

fMRI data preprocessing

All fMRI image preprocessing was performed using SPM12 (Statistical Parametric Mapping; Wellcome Trust Center for Neuroimaging, University College London, London, UK). The raw Digital Imaging and Communications in Medicine (DICOM) images were converted to NIfTI (Neuroimaging Informatics Technology Initiative) format. First, a voxel displacement map (VDM) was calculated using the field-map to account for spatial distortion resulting from the magnetic field inhomogeneity.⁷² Incorporating this VDM, the EPI images were corrected for motion and spatial distortions through realignment and unwarping using the first EPI image as reference. The anatomical image was then co-registered to the first image in the EPI series, skull-stripped and segmented into gray matter, white matter, and cerebrospinal fluid (CSF), using the ‘‘Segment’’ tool in SPM12. Individual flow-fields were computed by applying the SPM12 DARTEL toolbox on these gray and white matter images.⁷³ EPI images were then normalized to MNI space using the respective flow-fields through the DARTEL normalization tool. We applied a Gaussian kernel of 6 mm full width at half maximum (FWHM) to smooth the EPI images for use in upcoming univariate analyses. For all multivariate analyses we applied a smoothing kernel of only 2 mm FWHM to retain potential spatially distributed information.

Univariate analyses of prediction error signals

General linear modeling was done using SPM12. To detect neural correlates of state and reward prediction errors we performed model-informed fMRI analyses incorporating the computational signals of the hybrid-learning model. For each participant, using the hierarchically fitted individual parameter estimates, we derived trial-by-trial SPE and RPE variables, as well as the univariate state value (blue state = 0, red state = 1) and reward value. The PE and value signals were entered in the first-level GLMs as unorthogonalized parametric modulators to the state and reward outcome events in the task. This ensures that the PE and value regressor only capture their unique variance respectively thus to rule out confounding value signatures in the detected error-coding ROIs.

The first-level design matrix of the GLM for the model-informed univariate analyses identifying error-coding regions consisted of 9 regressors: an onset regressor for the start of a trial; 2 onset regressors for the left and right button presses; 2 regressors for the state- and reward-outcome modeled with a duration of 3 s; 4 parametric modulators using the model-derived SPE, RPE, state value and reward value variables (SPE and state value modulating state-outcome; second-stage RPE and reward value modulating reward-outcome). All GLMs in this study were specified using the canonical HRF as a basis function. Data were high-pass filtered using the “FAST” algorithm implemented in SPM12.

The 2nd-level analyses were performed as voxel-wise one-sample t-tests corrected for multiple comparisons using family wise error rate (FWE). Error-coding ROIs in the IPS (SPE) and VS (RPE) were defined using conjunctions (threshold of $p < 0.001$ uncorrected) of the univariate model-informed analysis locating SPE and RPE correlates in this project and results of the same analysis conducted in our earlier study.³⁰ The ROIs served as seed regions in subsequent analyses.

Pattern change searchlight analysis

To investigate the evolution of neural representations on a trial-by-trial basis, we implemented a searchlight-based correlation analysis (Figure 3). The aim of this analysis was to construct for each trial a whole-brain map containing information about the intensity of change in the spatially distributed shape of neural activity to the next upcoming trial. In order to quantify this change we defined the pattern change between two trials t and $t + 1$ as:

$$\text{pattern change}[t, t + 1] = \text{arctanh}(-\text{corr}(\text{pattern}[t], \text{pattern}[t + 1]))$$

where $\text{corr}()$ denotes to the Pearson correlation, $\text{pattern}[t]$ resembles the searchlight voxel data of an event in trial t and $\text{pattern}[t + 1]$ the searchlight voxel data of the corresponding event in the next upcoming trial $t + 1$. By using the inverse Pearson correlation coefficient r between these two patterns the pattern change measure reached its maximum, when patterns were maximally apart (i.e., a correlation of -1). The resulting value was then Fisher’s z transformed ($z = \text{arctanh}(r)$) to improve the distributional properties and allow for statistical testing. Defining the pattern change as the change to the next upcoming trial facilitated subsequent analyses, because hereby the pattern change of a trial t was already in alignment with the corresponding prediction error of trial t , that occurred between those two pattern events. The pattern change of the last trial in every run remains undefined.

Trial specific beta images served as input patterns to the searchlight analysis. Beta images were calculated using the ‘Least Squares - Separate’ approach⁷⁴ by estimating trial-specific first-level GLMs containing 4 regressors each: 2 onset regressors for the SPE-related pattern (at state-decision) and RPE-related pattern (at reward-decision) of the current trial of interest; 2 onset regressors containing SPE-related and RPE-related patterns of all other trials of the current run. This procedure was repeated for every trial in each run. Thus, we obtained independent and uniquely identifiable estimates of BOLD activations for each trial.⁷⁴ Within each searchlight (spherical searchlight, radius = 10 mm), we calculated the pattern change for each pair of adjacent trials and mapped values back to the center voxel of each searchlight, resulting in unique brain maps of SPE- and RPE-related pattern change for each trial in each participant. The Decoding Toolbox⁷⁵ served as a customizable framework for implementing this searchlight-based trial-wise correlation analysis.

Using the same searchlight algorithm, we also calculated maps of the change in mean BOLD signal of each searchlight sphere (i.e., pure univariate change of patterns) for use in confirmatory analyses. The change in mean BOLD signal is computed by calculating the difference of the average BOLD signal across all voxels in a given searchlight sphere between the current and upcoming trial (i.e., trial t and $t+1$). This procedure is perfectly analogous to the way we calculate the multivariate pattern change between the current and upcoming trial in a given searchlight sphere. Whole-brain images averaged across all trials and participants of pattern change and change in mean BOLD signal were computed for comparison purposes (Figure 4).

Predicting pattern change using PE-evoked BOLD

To examine the effects of prediction errors on evolving neural representations we constructed an analysis which aims to utilize PE-evoked BOLD responses in error-coding ROIs to predict subsequent pattern changes throughout the whole brain. To model trial-specific SPE- and RPE-evoked BOLD responses we set up trial-wise GLMs using the ‘Least Squares - Separate’ approach⁷⁴ consisting of 4 regressors each: 2 regressors with a duration of 3 s for the BOLD response to SPE and RPE events in the current trial; 2 additional regressors with the same temporal configuration for all other trials of the current run. From these trial-specific beta images we calculated the mean effect over voxels in each ROI in every trial (Figure 3). In the case of the RPE we used the absolute beta values, because we expected that large PEs (irrespective of their sign) should result in an increase of pattern change in modulated brain regions.

We set up a whole-brain univariate first-level GLM to regress the trial-by-trial pattern change against the trial-by-trial PE-evoked BOLD response in error-coding ROIs:

$$\text{pattern change} \sim b_0 + b_1 * \text{BOLD response to PE in ROI}$$

Parameter estimates of this GLM were tested for statistical significance in a 2nd-level voxel-wise one-sample t test FWE-corrected for multiple comparisons.

Permutation test of pattern change regression

We performed a permutation test to check the validity of the results obtained from the previous analysis in which we used PE-induced BOLD to predict upcoming multivariate pattern changes in other brain regions. In this permutation test we randomly shuffled the trial-wise IPS or VS, PE-induced BOLD regressors within each subject (and within the respective PE-type, e.g., not shuffling SPE with RPE regressors). We performed 10000 permutations for the regions identified in Figure 5 and then compared the 2nd level t-statistic of all permuted analyses with the real unpermuted 2nd level t-statistic.

Relating pattern change to behavioral policy

We hypothesized that updates of participants' internal model of the task should lead to adaptations in behavioral policy, which are captured in a $\Delta policy$ term (see below). To investigate this effect we set up a whole-brain univariate GLM to regress the trial-by-trial pattern change against the subsequent trial-by-trial changes in behavior:

$$\Delta policy \sim b_0 + b_1 * \text{pattern change}$$

Due to the binary nature of choices in this task we defined $\Delta policy$ using the continuous model-derived action probabilities. The adaptations in behavioral policy were computed separately for state- and reward-decisions (A1 and A2 respectively) using the change in action probabilities P_{A1} or P_{A2} of the chosen action in a given state for each trial via:

$$\Delta policy_{A1}[t, t + 1] = P_{A1_chosen_on_trial_t}[t, s] - P_{A1_chosen_on_trial_t}[t + 1, s]$$

The construction of the behavioral policy change $\Delta policy$ followed the logic that a different choice on the next trial should result in a larger policy change index, than choosing the same action again (i.e., reinforcing the current policy). For instance, if the probability of the chosen action on trial t is 0.5 and on the next trial the probability for the same action at trial $t+1$ decreases to 0.2 the resulting $\Delta policy$ value will be 0.3 and thus positive. This also implies that the probability for the non-chosen action in trial t increases from 0.5 to 0.8 suggesting that a change in the behavioral response is likely. This is captured in the positive $\Delta policy$. Conversely, if the probability of the chosen action would increase from 0.5 on trial t to 0.9 on the following trial $t+1$ the computed $\Delta policy$ will take on a negative value of -0.4 . This also implies that the non-chosen action on trial t decreases from 0.5 to 0.1, thus making a change in the behavioral response at trial $t+1$ less likely. In summary, $\Delta policy$ describes the trial-by-trial amount of divergence from the current behavioral policy and can be directly related to the amount of pattern change. Parameter estimates of this GLM were tested for statistical significance in a 2nd-level voxel-wise one-sample t test FWE-corrected for multiple comparisons. To ensure that multivariate information of the pattern change is required for predicting changes in behavioral policy we repeated this analysis using the change in mean BOLD signal strength of each searchlight as the predictor, resulting in no significant effects.

Detecting the value of the chosen action

In a next step, we wanted to investigate if brain regions showing RPE-modulated pattern change also hold information about the participants' expectations regarding upcoming rewards. To identify brain regions coding for reward expectations we conducted a univariate model-informed analysis detecting the value of the chosen second-stage action A2 during the reward-decision (the same timepoint where RPE-modulated pattern change was computed). The first-level GLM was set up as follows: 1 onset regressor at state-decision, 1 onset regressor at reward-decision being parametrically modulated by $Q_{HYBRID}(s, a_{chosen})$, 2 onset regressors for state and reward outcomes. Parameter estimates of this GLM were tested for statistical significance in a 2nd-level voxel-wise one-sample t test FWE-corrected for multiple comparisons.

Decoding of state expectations

In contrast to the univariate reward-related action values, we assumed that participants' expectations about upcoming states are represented as multivariate information in the BOLD signal. To detect brain regions encoding state expectations we performed a searchlight decoding analysis (spherical searchlight, radius = 15 mm) using support vector machine (SVM) classifiers as implemented in The Decoding Toolbox.⁷⁵ First, we estimated participants' expectations by deriving the trial-by-trial estimated transition probabilities T from the hybrid learner. Trials were then sorted in one of three different groups: Label 'blue', Label 'red' and 'uncertain' expectation. 'Uncertain' trials were excluded from the classification process. A trial was considered Label 'blue' or Label 'red' when these states had the highest estimated transition probability for the chosen action in the state-decision. Expectations were assumed to be 'uncertain', if:

$$T(a_{1chosen}, blue) = = T(a_{1other}, red)$$

The GLM generating the input data for classification was set up as follows: 1 of 3 possible onset regressors (either Label ‘blue’, Label ‘red’, or ‘uncertain’) modeled the state-decision, 1 onset regressor for all reward decisions. We grouped all trials of a label in a single regressor instead of estimating single trial regressors to preempt problems regarding unbalanced training data. This technique makes more use of all collected data compared to obtaining balanced training data by bootstrapping single trial regressors and thus having to exclude trials from training. We applied a leave-one-out cross-validation scheme using 4 runs as training and the fifth run as test data, repeated for all 5 runs. Mean area under the curve (AUC, an indicator of the distinguishability of all classified labels) over all cross-validations was group-level tested for being significantly above chance-level in a 2nd-level voxel-wise one-sample t test FWE-corrected for multiple comparisons.

Group-level statistical analyses

All group-level analyses were performed using voxel-wise t-tests FWE-corrected for multiple comparisons (results were considered as group-level significant at a threshold of $p_{FWE} < 0.05$). We also applied threshold free cluster enhancement (TFCE, $H = 2$, $E = 0.5$)³⁴ with 5000 permutations per contrast as implemented in the PALM toolbox⁷⁶ to increase sensitivity for spatially distributed effects. Conjunctions combining two hypotheses (intersection-union test) were created using the minimal T-statistic³⁵ of FWE-corrected results (conjunctions of uncorrected p values were created for visualization purposes). A full list of results and significant sub-peaks can be found in the supplementary material.

Cell Reports, Volume 42

Supplemental information

**Prediction errors drive dynamic changes
in neural patterns that guide behavior**

Leon Möhring and Jan Gläscher

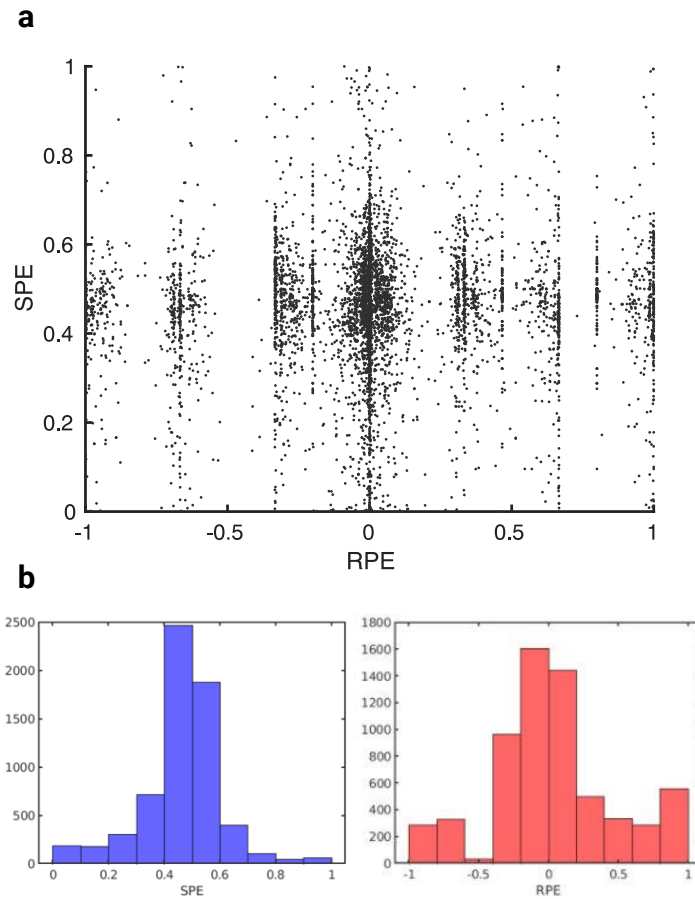


Figure S1 | SPE and RPE distributions, Related to Figure 1. **a** The modified two-step Markov decision task elicits orthogonal SPEs and RPEs (correlation of all SPEs and RPEs: $r = 0.0635$). **b** The task also produces balanced distributions of SPE and RPE magnitudes.

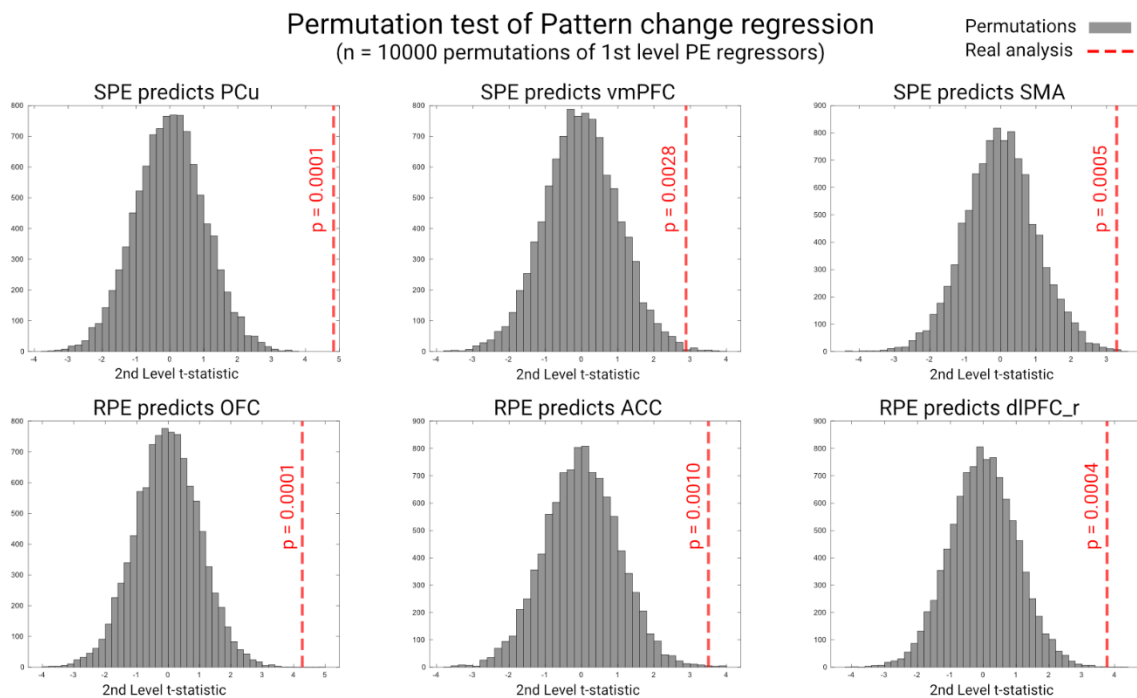


Figure S2 | Permutation test of the pattern change regression, Related to STAR Methods. Results of a permutation test in which 2nd level t-statistics of the real SPE and RPE induced BOLD responses (red dashed line) were compared against 10000 random permutations (gray histogram). Permutation tests were performed for relevant ROIs which have been identified in Figure 5. P-values are reported for the hypothesis that the real SPE and RPE induced BOLD responses better predict subsequent pattern change than the distribution of permutations.

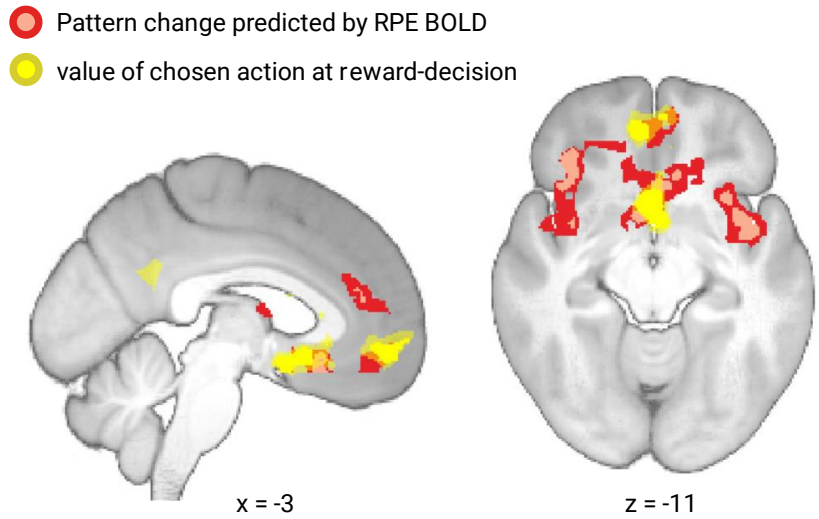


Figure S3 | Chosen action value at the time of reward-decision, Related to Figure 5. Yellow: value of the chosen action at the time of reward-decision. Red: pattern change at the time of reward-decision predicted by RPE-evoked BOLD responses in VS. All results are presented at a threshold of $p < 0.005$ uncorrected (dark area) and $p < 0.001$ uncorrected (light area).

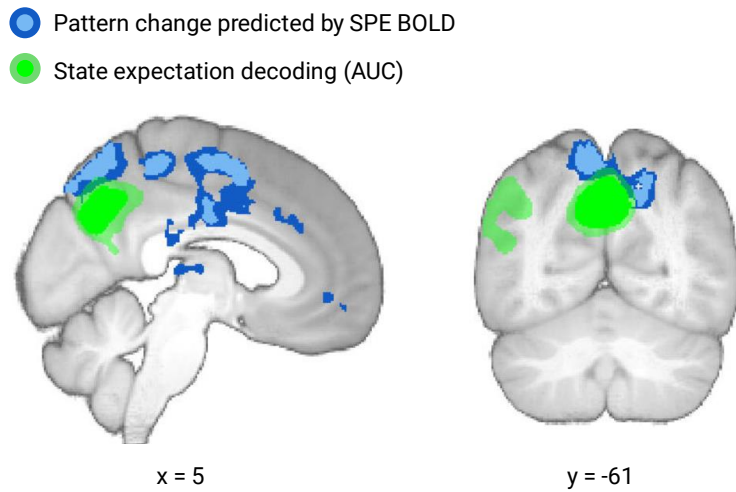


Figure S4 | State expectation decoding at the time of state-decision, Related to Figure 5. Green: decoding of participant's state expectations at the time of state-decision. Blue: pattern change at the time of state-decision predicted by SPE-evoked BOLD responses in IPS. All results are presented at a threshold of $p < 0.005$ uncorrected (dark area) and $p < 0.001$ uncorrected (light area).

- Pattern change at state decision predicting changes in behavioral policy of state decision
- Pattern change at reward decision predicting changes in behavioral policy of reward decision

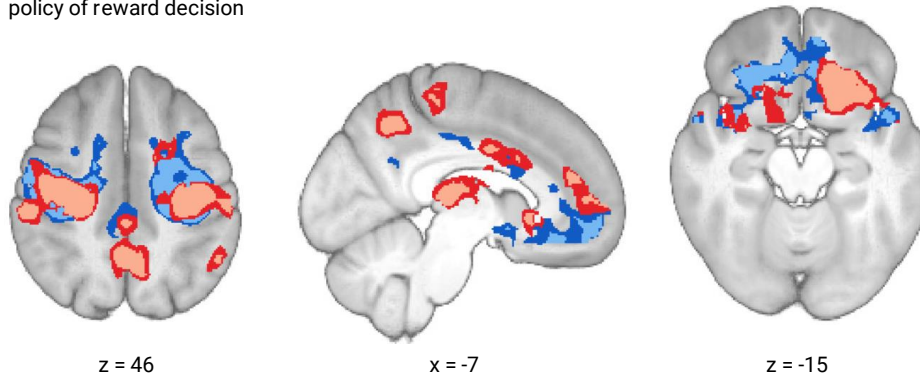


Figure S5 | Pattern change predicts subsequent adaptations of behavioral policy, Related to Figure7. Results of a trial-by-trial regression of pattern change on subsequent changes in behavioral policy ($\Delta policy$). All results are presented at a threshold of $p < 0.005$ uncorrected (dark area) and $p < 0.001$ uncorrected (light area).

Table S1 | Model comparison and hybrid model parameters, Related to STAR Methods. The HYBRID learner outperforms both the pure model-based (FORWARD) and model-free (SARSA) RL algorithms. Group-level mean parameters of the hybrid learner model (*lrSARSA* = learning-rate of SARSA learner; *lrFORWARD* = learning-rate of FORWARD-learner; temperature = temperature of softmax action selection; *discount factor* = temporal discount factor of rewards; *p* = probability to repeat the same action during the first-stage decision; *w* = constant weighting factor combining model-based and model-free algorithms, *w* = 1 is full model-based learning).

Model	Description	# Par.	AIC	BIC
HYBRID	combination of model-based and model-free	6	7620	7629
MB	pure model-based	5	8530	8540
MF	pure model-free	4	9055	9061

Hybrid model parameters (group-level mean \pm SEM)					
<i>lrSARSA</i>	<i>lrFORWARD</i>	<i>temperature</i>	<i>discount factor</i>	<i>p</i>	<i>w</i>
0.874	0.251	3.589	0.094	0.176	0.333
[0.860, 0.889]	[0.230, 0.272]	[3.341, 3.838]	[0.081, 0.108]	[0.161, 0.191]	[0.313, 0.353]

Table S2 | Univariate GLM and decoding results, Related to STAR Methods. Results of univariate model-informed analyses to detect SPE and RPE signals as well as the value of the chosen action in the brain. Furthermore, decoding results of participant's state expectations at the time of state-decision. Minimum distance of sub-peaks of 30 mm.

H	Brain region	x	y	z	Size (mm ³)	p FWE	TFCE
(a) Univariate SPE							
L	IPS, PCu	-10	-70	43	15585	0.018	850
R	IPS	40	-45	45	6419	0.019	829
R	MFG	35	5	53	1900	0.031	733
R	dIPFC	47	29	29	2707	0.037	697
R	Insula	34	24	-3	36	0.048	642
(b) Univariate RPE (small volume corrected using the VS ROI of Gläscher et al., 2010)							
L	VS	-8	15	-8	156	0.005	
(c) Univariate value of chosen action (at reward-decision, i.e. A2)							
	ACC subgenual	1	12	-10	1211	0.023	608
	vmPFC	1	52	-7	2852	0.036	532
(d) Decoding of state expectations (AUC minus chance-level)							
	PCu	2	-61	37	9520	0.025	979

Table S3 | Pattern change GLM results, Related to STAR Methods. Results of the GLMs combining pattern change, PE-evoked BOLD responses and changes in behavioral policy. Minimum distance of sub-peaks of 30 mm.

H	Brain region	x	y	z	Size (mm ³)	p FWE	TFCE
(a) SPE-evoked BOLD response in IPS predicts subsequent pattern change at state-decision							
L	SPL	-23	-42	45	186098	0.002	1392
R	↳ SMA	10	0	62	x	0.003	1286
R	↳ PCu	8	-77	50	x	0.004	1260
	↳ MCC	0	-9	33	x	0.007	1126
L	↳ PCu	-13	-47	74	x	0.011	1030
L	↳ IPS	-55	-42	42	x	0.016	947
R	↳ Insula	32	10	17	x	0.029	773
R	MFG	22	48	30	14979	0.023	863
R	↳ vmPFC	13	53	-13	x	0.036	769
L	OFC	-30	40	-10	10046	0.032	792
L	↳ MFG	-38	43	30	x	0.036	767
(b) RPE-evoked BOLD response in VS predicts subsequent pattern change at reward-decision							
R	dIPFC	47	20	15	395224	0.001	1637
L	↳ OFC	-28	28	-20	x	0.002	1468
R	↳ OFC	25	42	-17	x	0.003	1395
R	↳ M1	42	-25	52	x	0.003	1381
L	↳ M1	-48	-18	43	x	0.003	1357
L	↳ Insula	-35	-2	-15	x	0.003	1344
L	↳ IPS	-33	-63	38	x	0.003	1341
	↳ ACC pregenual	13	38	10	x	0.003	1338
R	↳ Insula	47	2	-15	x	0.004	1308
	↳ ACC subgenual	7	7	-10	x	0.004	1299
L	↳ dIPFC	-38	22	17	x	0.004	1278
R	↳ IPS	33	-55	38	x	0.028	832
R	↳ MFG	35	45	30	x	0.032	801
(c) Pattern change at state-decision predicts subsequent change in behavioral policy of A1							
R	M1	35	-15	53	378594	< 0.001	2127

L	↳ M1	-30	-23	51	x	< 0.001	1866
R	↳ vmPFC	12	57	-6	x	0.001	1662
R	↳ Insula	35	-11	-10	x	0.001	1508
R	↳ OFC	22	25	-18	x	0.001	1499
L	↳ SMA	-12	10	52	x	0.002	1442
L	↳ vmPFC	-7	57	19	x	0.003	1324
L	↳ OFC	-15	22	-18	x	0.003	1287
R	↳ PCu	17	-52	20	x	0.011	982
L	↳ dlPFC	-53	27	9	x	0.017	893
R	↳ dlPFC	52	25	7	x	0.037	725
R	↳ MFG	47	20	42	x	0.037	722
(d) Pattern change at reward-decision predicts subsequent change in behavioral policy of A2							
L	vmPFC	-15	43	20	421564	0.001	1766
L	↳ OFC	-25	25	-18	x	0.001	1753
R	↳ M1	30	-16	49	x	0.001	1677
L	↳ M1	-37	-19	42	x	0.001	1617
L	↳ MCC	-10	10	34	x	0.002	1397
L	↳ dlPFC	-57	15	12	x	0.002	1312
R	↳ OFC	20	21	-15	x	0.005	1185
L	↳ MFG	-42	34	37	x	0.110	1014
L	↳ PCu	-10	-64	67	x	0.018	902
	↳ SMA	0	7	65	x	0.023	835
R	↳ MFG	44	36	37	x	0.044	715
(e) Conjunction: Pattern change at reward-decision correlated with RPE-evoked BOLD response in VS AND with change in behavioral policy of A2 (conjunction performed on FWE-corrected results without TFCE, because conjunctions may not be performed on TFCE enhanced data)							
L	OFC	-27	24	-21	121	0.032	

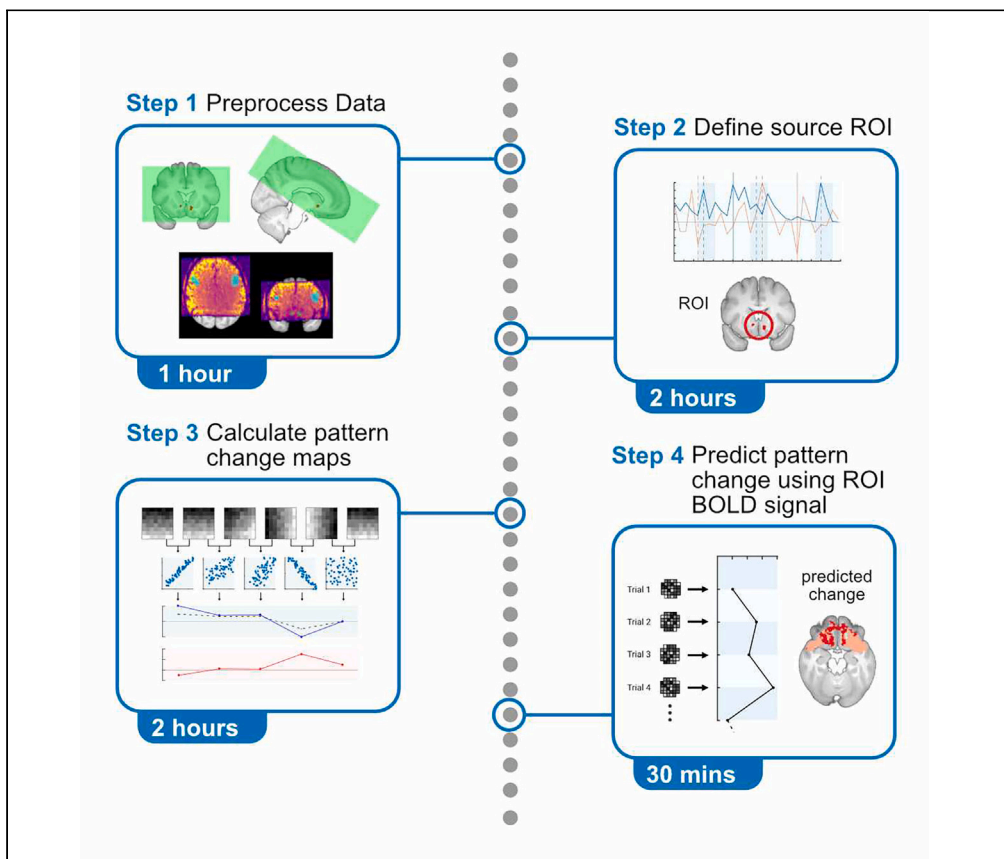
Table S4 | Uncorrected conjunction results, Related to STAR Methods. Results of conjunction analyses on uncorrected results to test joint hypotheses. Minimum distance of sub-peaks of 30 mm.

H	Brain region	x	y	z	Size (mm ³)	Peak p unc.
(a) Conjunction: Pattern change at reward-decision correlated with RPE-evoked BOLD response in VS AND with change in behavioral policy of A2						
L	OFC	-26	23	-17	1895	0.00028
R	Insula	38	-5	16	1318	0.00046
R	dIPFC	49	15	17	722	0.00071
L	M1	-44	-20	44	405	0.00054
	ACC pregenual	3	40	19	178	0.00086
(b) Conjunction: Pattern change at state-decision correlated with SPE-evoked BOLD response in IPS AND with change in behavioral policy of A1						
R	M1	39	-19	48	3322	0.00031
L	M1	-33	-15	56	2135	0.00048
R	vmPFC	10	48	-9	739	0.00093
(c) Conjunction: Pattern change at reward-decision correlated with RPE-evoked BOLD response in VS AND value of chosen action A2 at the time of reward-decision						
	ACC subgenual	2	10	-12	191	0.00037
(d) Conjunction: Pattern change at state-decision correlated with SPE-evoked BOLD response in IPS AND decoding of state expectations at the time of state-decision						
	PCu	3	-65	48	549	0.00186

1.2 Protocol for predicting multivariate change of brain patterns using model-informed fMRI activations (*STAR Protocols, 2024, Möhring and Gläscher*)

Protocol

Protocol for predicting multivariate change of brain patterns using model-informed fMRI activations



Leon Möhring, Jan Gläscher

glaescher@uke.de

Highlights
Quantify changes in multivariate fMRI patterns using a searchlight algorithm

Dynamically predict pattern change using trial-by-trial BOLD activity of a seed region

Investigate the temporal dynamics of changes in multivariate fMRI patterns

Investigating the spatially distributed information contained in fMRI data is essential for understanding brain functions. Here, we present a protocol to dynamically predict short-term changes in neural patterns using trial-by-trial blood-oxygen-level-dependent (BOLD) activity of a seed region. We describe steps for setting fMRI data acquisition parameters and quantification of changes in multivariate patterns. We then detail procedures for defining seed regions and identifying brain areas in which changes in multivariate patterns can be predicted by BOLD activity of the seed region.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Möhring & Gläscher, STAR Protocols 5, 102978
June 21, 2024 © 2024 The Author(s).
<https://doi.org/10.1016/j.xpro.2024.102978>



Protocol

Protocol for predicting multivariate change of brain patterns using model-informed fMRI activations

Leon Möhring^{1,2} and Jan Gläscher^{1,3,*}

¹Institute for Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany

²Technical contact

³Lead contact

*Correspondence: glaescher@uke.de
<https://doi.org/10.1016/j.xpro.2024.102978>

SUMMARY

Investigating the spatially distributed information contained in fMRI data is essential for understanding brain functions. Here, we present a protocol to dynamically predict short-term changes in neural patterns using trial-by-trial blood-oxygen-level-dependent (BOLD) activity of a seed region. We describe steps for setting fMRI data acquisition parameters and quantification of changes in multivariate patterns. We then detail procedures for defining seed regions and identifying brain areas in which changes in multivariate patterns can be predicted by BOLD activity of the seed region.

For complete details on the use and execution of this protocol, please refer to Möhring et al.¹

BEFORE YOU BEGIN

Institutional permissions

This study protocol was approved by the Ethics Committee of the Medical Association of Hamburg (2020-10102-B0-ff) and conducted in accordance with the Declaration of Helsinki. Any following fMRI studies will need to acquire permissions from the relevant institutions.

Predicting multivariate pattern change – Description of analysis strategy

Multivariate methods provide novel perspectives on stimulus representations in the human brain, unveiling the information embedded in spatially distributed patterns of neural activity, as measured by functional magnetic resonance imaging (fMRI). A plethora of studies have utilized machine learning algorithms to decode information about stimuli or investigate the (dis)similarity of activation patterns in relation to perceptual or categorical similarity. However, limited methods exist to explore the correlation between multivariate changes and potentially causal signals in the brain that give rise to these alterations in patterns. Our strategy of predicting multivariate pattern changes integrates both univariate and multivariate techniques. It enables the prediction of pattern alterations by leveraging preceding BOLD responses within specified regions of interest (ROIs) on a trial-by-trial basis. The protocol involves the following 4 major parts, each of them consisting of several analysis steps.

1. Define a seed ROI that is hypothesized to elicit pattern change in some other brain region(s).
2. Extract the trial-by-trial BOLD activations in the seed ROI.
3. Calculate multivariate pattern change across the entire brain using a searchlight algorithm. This condenses the multivariate patterns into a univariate signal of change.
4. Regress the pattern change images series (dependent variable) onto the BOLD activations from the seed ROI (independent variable, predictor) using a univariate GLM.



We describe these parts in more detail in the sections below.

Note: This protocol describes the specific steps to investigate pattern change following prediction errors (PE) during a reinforcement learning (RL) task. In the original publication we utilize two distinct seed ROIs exhibiting two different types of PEs to predict subsequent changes in neural patterns of other brain areas. These seed ROIs are the ventral striatum (VS), a region well known to show a reward prediction error (RPE) signal, as well as the intraparietal sulcus (IPS) which exhibits a state prediction error (SPE) signal. This enables us to detect brain regions in which pattern changes can be predicted by only one of two prediction error types. Though, the analysis does not require two seed ROIs and can also be used for hypotheses involving only a single seed ROI.

Note: The methodological approaches we present are applicable to many other types of tasks and brain regions. To apply this protocol to your studies the requirement is a seed ROI exhibiting a trial-wise explanatory BOLD response in combination with changing neural representations in other parts of the brain.

Design experimental task

⌚ Timing: 4 weeks

5. Choose an experimental task for the project.

Note: There are a number of things to consider when designing a task for multivariate analyses in fMRI. Some of the points below apply to multivariate analyses in general; others are specific to this protocol.

- a. Dynamics and effect strength of the seed ROI activity.
 - i. Estimate the required sample size of the project to achieve a significant effect in the seed ROI.

Note: Such power analysis can be carried out with ready-made tools for fMRI, e.g. neuropowertools.org by Durnez et al.² or [fmripower](https://fmripower.org) by Mumford et al. (although this tool is no longer developed and updated).³

- ii. Aim for a Gaussian distribution of BOLD response magnitudes in the seed ROI to facilitate estimation of the final GLM.
- b. Time interval between adjacent pattern events (i.e., the events at which multivariate patterns are extracted and used to calculate the pattern change, see [Figure 1](#)).
 - i. Slightly shift the timing of pattern events in each run to reduce temporal autocorrelation.

Note: see [Figure 2](#), the time shift is introduced by participant's input reaction times during decision 1 (state-decision) 1.58 ± 0.77 s [mean \pm SEM] and decision 2 (reward-decision) 1.07 ± 0.66 s.

- ii. The time interval between two adjacent patterns, which will be compared in the analysis, should not be too short (temporal autocorrelation) or too long (not enough datapoints/trials, possibly greater changes in neuronal background activity masking the real pattern change of interest).

Note: We were able to produce good results with an interval of around 25–30 seconds between two corresponding pattern events (see [Figure 2](#) for event timings).

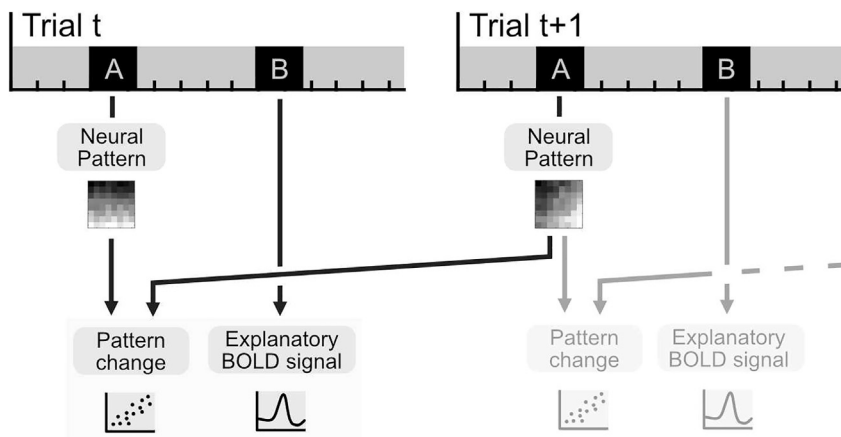


Figure 1. Analysis strategy and trial structure

The explanatory BOLD signal at event B (explanatory event) of the seed ROI, which will later be used to predict the change in neural patterns of other brain regions, occurs between both of the events A (pattern event), which elicit the patterns that are used to calculate the pattern change metric. In the final GLM the pattern change between two adjacent events A is regressed onto the explanatory BOLD responses at event B. Notably, the pattern change of the last trial is undefined, because there is no next pattern to compare the last trial's pattern against.

c. Temporal relation of seed region event to pattern event.

Note: The event eliciting the BOLD response which is hypothesized to predict pattern changes should occur between the two pattern events used to calculate the pattern change, see [Figure 1](#).

Optional: When using multiple model-informed computational signals as predictors of BOLD activity, ensure that these signals are mostly uncorrelated, which helps with the identifiability of the neural correlates of the computational variables.

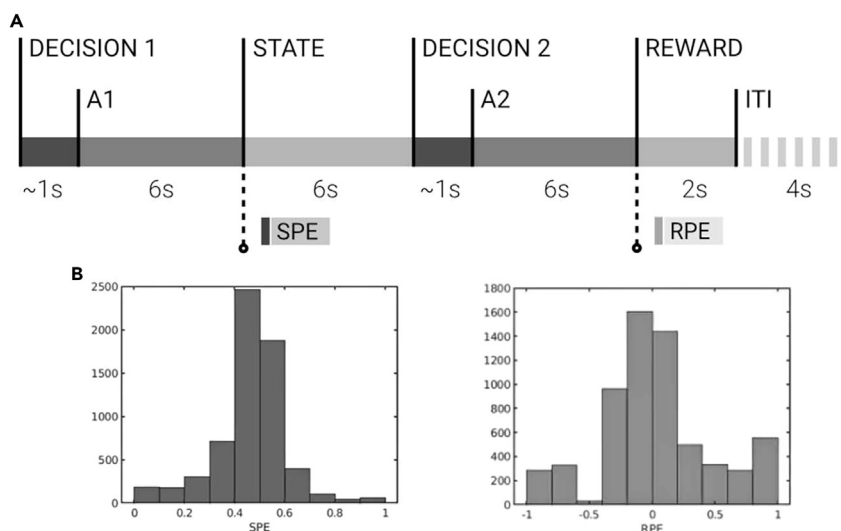


Figure 2. Trial timing and seed ROI signal distribution

(A) Timing of events in one trial of the original task design. The input of participant's decisions is not time restricted and introduces slight shifts in timing of the pattern events, therefore reducing temporal autocorrelation.

(B) Histograms of the two seed ROI BOLD responses of all participants in the original experiment. Aim for a Gaussian distribution of BOLD responses in the seed ROI to facilitate the estimation of the final GLM that predicts pattern change using these signals.

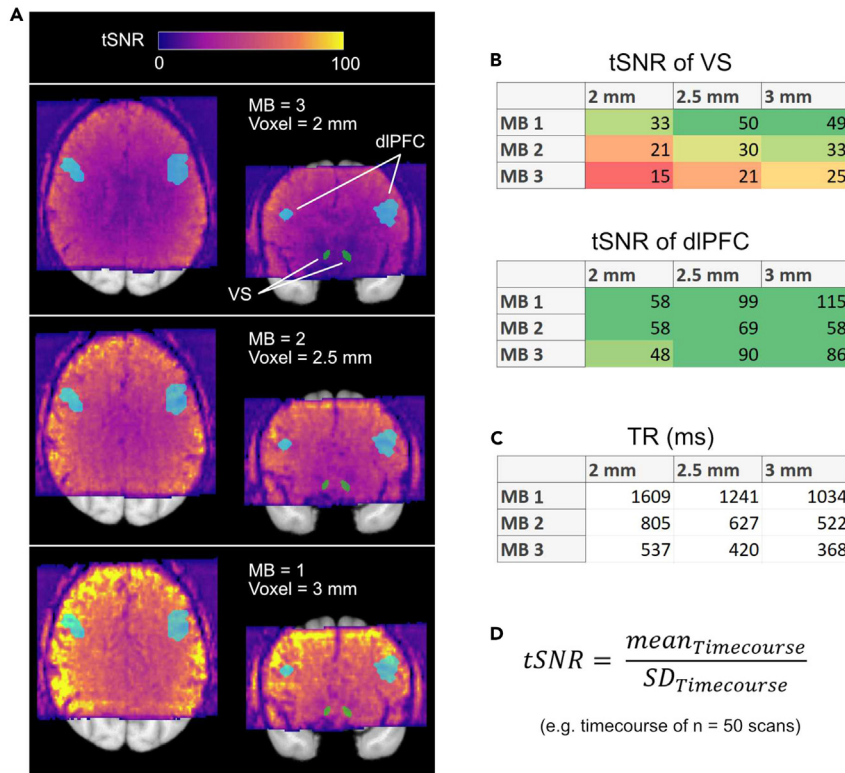


Figure 3. Different scan parameters and corresponding tSNR maps

(A) The tSNR maps for 3 different combinations of fMRI scan parameters calculated for 50 scans each. (B) Comparison of tSNR values for different combinations of voxel sizes and multiband factors (MB) in a subcortical ROI (VS) and a cortical ROI (dorsolateral prefrontal cortex, dlPFC). Subcortical regions have lower tSNR values in general and were the limiting factor in this project. We chose to use a voxel size of 2.5 mm and MB 2 for functional imaging. (C) Repetition times (TR) of the different fMRI scan parameter combinations. (D) The tSNR is defined as the mean signal strength of a series of scans divided by the standard deviation of the signal.

- Gather behavioral pilot data to ensure the task is capable of producing the desired behavioral response and neural effects.

Set fMRI scan parameters

⌚ Timing: 1 day

- Balance scan parameters for functional imaging depending on your specific needs whilst retaining a reasonable temporal signal to noise ratio (tSNR).

Note: We recommend performing test scans on your local MR scanner using multiple parameter combinations and calculate the tSNR of your target regions to guide your final decision (see [Figure 3](#)).

Optional: Especially subcortical regions (like the ventral striatum) are susceptible to signal dropout when the scanning parameters are configured in an unfavorable way. In this case, reverting the phase encoding direction might help in recovering signal in these regions.

Note: The following fMRI sequence parameters affect tSNR. We recommend selecting a combination of these parameters that maximize tSNR in the brain region with the lowest tSNR (either the seed or target region).

- a. TR.
- b. Voxel size.
- c. Acceleration (e.g., Multiband, in-plane acceleration).
- d. Scan volume (i.e., brain coverage of slice stack).

Note: The scan parameters we used in this project were chosen and optimized with the intention to achieve a very short TR. However, the analysis described in this protocol should also work on conventional TR intervals of around 2 seconds. In this specific project, MRI data collection was conducted on a Siemens Prisma 3T scanner (Siemens, Erlangen, Germany) equipped with a 64-channel head coil. Each brain volume consisted of 24 axial slices (voxel size, $2.5 \times 2.5 \times 2.5 \text{ mm}^3$, and 0.5 mm spacing between slices) acquired using a T2*-weighted echo planar imaging (EPI) protocol (repetition time, TR = 627 ms; echo time, TE = 30 ms; flip angle = 45°; field of view = 224 mm, inplane acceleration factor = 2, Multiband factor = 2, Phase encoding direction P >> A). The orientation of the slice pack was tilted approximately 30° upward to the anterior commissure–posterior commissure axis to minimize susceptibility artifacts in the orbitofrontal cortex (OFC).⁴

Note: The flip angle that maximizes SNR under repeated measurements (the “Ernst angle” is a function of the repetition time (TR) and the T1 relaxation time and is calculated as $\alpha_{Ernst} = \arccos\left(e^{-\frac{TR}{T1}}\right)$. In our Siemens Prisma scanner the T1 relaxation time of gray matter is measured as 1500 ms, which is the range reported in the literature.

8. Collect additional data for later use during preprocessing of the functional images.
 - a. Field map: acquired with a gradient echo based sequence before functional EPI scanning to measure the magnetic field inhomogeneity (TE1 = 5.29 ms, TE2 = 7.75 ms).
 - b. Anatomical scan: a high-resolution anatomical image (voxel size, $1 \times 1 \times 1 \text{ mm}^3$) acquired after the experiment using a T1-weighted MPRAGE protocol (FoV 256 mm, TR = 2300 ms, TE = 2,98 ms, phase encoding direction R >> L).

Optional: Collect physiological data (e.g. ECG, breathing belt) to remove physiological noise from the BOLD data (However, this did not improve sensitivity of the multivariate analyses in our project).

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Behavioral data	This paper	https://doi.org/10.5281/zenodo.8015563
Code	This paper	https://doi.org/10.5281/zenodo.8015561
Imaging data	This paper	Will be shared by the lead contact upon request
Software and algorithms		
MATLAB	MathWorks	https://de.mathworks.com/products/MATLAB.html , RRID:SCR_001622
SPM12	The Wellcome Centre for Human Neuroimaging	https://www.fil.ion.ucl.ac.uk/spm/ , RRID:SCR_007037
The Decoding Toolbox	Martin Hebart	https://doi.org/10.3389/fninf.2014.00088
MRlcroGL	Chris Rorden	https://www.nitrc.org/projects/microgl , RRID:SCR_002403
dcm2niix	Chris Rorden	https://github.com/rordenlab/dcm2niix
DARTEL Toolbox	The Wellcome Centre for Human Neuroimaging	https://github.com/neurodebian/spm12/blob/master/toolbox/DARTEL/tbx_cfg_dartel.m

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<i>Other</i>		
Siemens Prisma 3T scanner	Siemens, Erlangen, Germany	https://www.siemens-healthineers.com/de/magnetic-resonance-imaging/3t-mri-scanner/magnetom-prisma
Head/Neck 64 (64-channel head coil)	Siemens, Erlangen, Germany	https://www.siemens-healthineers.com/magnetic-resonance-imaging/options-and-upgrades/coils/64-channel-head-neck-coil

STEP-BY-STEP METHOD DETAILS

Perform experiment

⌚ Timing: 2 h

Note: In this major part the participant performs the experimental task inside the scanner and all data needed for the analysis are gathered. The participant should receive all necessary instructions to perform the task before entering the scanner.

1. Position participant inside the scanner and install a 64-channel head-coil.
2. Perform localizer scans.

Optional: This is a good time to check any input devices the participants will use throughout the task and run a test trial to make participants familiar with the setup.

3. Acquire a field map to measure inhomogeneities of the magnetic field before collecting the functional EPI data.

Note: The positioning of the slice pack is automatically copied from the positioning of the field map measurement.

4. Start task using a presentation software tool, perform task and acquire functional imaging data (EPI).
5. After the experiment acquire anatomical images of the participant (MPRAGE).

Preprocess data

⌚ Timing: 30 min

Note: There are many protocols for preprocessing fMRI data in various software packages (SPM, FSL, AFNI, Brain Voyager) or using a combination of them (e.g. fMRIPrep) that aims for standardization and replicability. The protocol below represents a fairly standard way of preprocessing fMRI data with SPM12, you can find the preprocessing code used in this protocol in the code depository (<https://doi.org/10.5281/zenodo.8015561>).

6. Convert all imaging data from DICOM format into NIFTI (.nii) format (e.g., using the dcm2niix software package).
7. Realign and unwarp functional imaging data (to minimize noise introduced by head movement and rotation of participants in the functional imaging data).
 - a. Create a voxel displacement map (VDM) using the fieldmap data.
 - b. Perform the SPM12 realign and unwarp⁵ on the functional imaging data using the first functional EPI scan as reference.

8. Normalization of all images to the MNI space (this step involves aligning the individual data from all participants into a common normalized space, enabling second-level group analyses).
 - a. First co-register the anatomical image to the first functional EPI scan.
 - b. Skull-strip and segment the anatomical image into gray matter, white matter, and cerebrospinal fluid (CSF) using the “Segment” tool of SPM12.
 - c. Use the results of the segmentation to create a flow-field in the DARTEL toolbox.⁶
 - d. Normalize the anatomical and functional image data by applying the flow-field to all images ([troubleshooting 1](#)).
9. Smoothing of functional images (to enhance the signal-to-noise ratio in the functional data. We applied less smoothing specifically for use in multivariate analyses to preserve spatially distributed information).
 - a. For univariate analyses smooth with a Gaussian kernel of FWHM = 6 mm.
 - b. For multivariate analyses smooth with a Gaussian kernel of FWHM = 2 mm.

Define ROI of seed region

⌚ Timing: 1 h

Note: The first major step is to define the Region of Interest (ROI) of the seed region which will be used to predict subsequent pattern changes in other areas of the brain. There are several ways to define a ROI for the seed region:

10. Anatomical ROI definition:
 - a. Anatomical mask image defined in a brain template (e.g., MNI154, Colin27).
 - b. Anatomical mask image defined individually for each participant (e.g., by tracing and anatomical region on a T1 image using the drawing tools of established software packages, e.g., MRICroN, MRICroGL).
11. Functional ROI definition:
 - a. Localizer experiment: short fMRI run presenting stimuli that activate the intended seed region.
 - b. Localizer contrast: similar to a., except that the ROI is identified with a functional contrast in the main experiment.
 - c. Model-informed analysis: using variables (e.g., a reward prediction error) derived from a computational model (in the case of this protocol we used a reinforcement learning algorithm to model prediction error signals) as parametric modulators in the definition of the design matrix in a GLM.

Note: This additional regressor will specify model-derived modulations of the height of the BOLD response (i.e., a large prediction error will result in a larger predicted BOLD response in the design matrix). This approach is similar to b., but it utilizes variables from a computational model to identify the seed ROI. In the SPM software package, which was used in the original study, the values of these model-derived variables can be included in the onset definition during model specification.

Note: In the main paper that this protocol describes, we used the option ‘Model-informed analysis’ to identify 2 ROIs that correlated with a model-derived reward and state prediction error. However, this protocol does not require 2 separate ROIs, a single explanatory seed ROI is sufficient. Below are additional details for running this analysis.

12. Specify a 1st level GLM for each participant using all experimental events in a single condition (regressor) and the model-derived computational signal as parametric modulators of these onset events.

13. Estimate the model and include the beta images for the parametric modulators in a second-level group analysis.

Note: If the modulator(s) are independent of each other, a one-sample t-test as the 2nd level design is appropriate, if the ROI is only identified via a differential contrast (e. g. modulator1 > modulator2), then a paired t-test is the more appropriate design choice.

Optional: If the statistical results do not survive family-wise error correction, but the statistical map shows wide-spread activations just below the statistical threshold, refer to [troubleshooting 2](#).

14. Save the results for each ROI as a mask image and use it subsequently to extract single-trial BOLD activations of the seed ROI.

Extracting BOLD activations from the seed ROI

⌚ Timing: 2 h

Note: The goal of this part is the creation of a vector of BOLD activations in the seed ROI that can be used as a predictor to detect correlated pattern change. Thus, the result of this step is no longer a time-series in the traditional sense (i.e. a value for every EPI image in the original BOLD time series recorded in the scanner), but rather a vector of trial-by-trial estimates of the BOLD activation for each event in the experiment. It involves 2 analysis steps:

15. Obtain the BOLD activation at the time of the explanatory event (see event B in [Figure 1](#)) by using a first-level analysis that estimates the BOLD activation for each trial separately. There are two approaches to setup this first-level design detailed in Mumford et al.⁷ We found that the Least-squares Separate method was able to slightly improve the sensitivity of the pattern change analysis described in this protocol (this is especially the case for estimating first-level BOLD activations of the pattern events). However, this method is computationally intensive, see [troubleshooting 3](#).

- a. Least-squares All: A single first-level design with as many onset regressors as the number of trials in the experiment.

Note: Each regressor contains only one trial of each event type in the experiment (i.e. it's onset). This results in as many beta images as there are trials each representing the BOLD effect for its trial.

- b. Least-squares Separate: one first level GLM for each trial containing the trial in question in a single onset regressor and all other trials in a second regressor.

Note: This requires estimating as many first level GLM as there are trials in the experiment. Only the single trial beta image is of interest for further analyses. Although being a heavier computational burden, this approach supposedly provides more precise estimates of the single trial BOLD effect (see Mumford et al.⁷ for a more thorough and nuanced discussion). [Figure 4](#) contrasts the first level designs for *Least Square – Separate* and *Least Square – All* as discussed in the Mumford paper.

16. Extract the mean BOLD effect for each trial in the seed ROI from the results of the trial-wise first-level GLM using the ROI mask image created in the previous part.

Note: This script shows how to use SPM functions to open .nii beta images, remove non numeric values (NaN, not a number) and then extract the mean BOLD signal of the ROI mask in every trial:


```

> ROI_mask_header = spm_vol(char(ROI_mask_filename));
> ROI_mask_vol = spm_read_vols(ROI_mask_header);
>
> for i_trial = 1:n_trials
> beta_image_header = spm_vol(char(beta_image_filename));
> vol = spm_read_vols(beta_image_header);
> vol(isnan(vol)) = 0;
> seed_ROI_BOLD(i_trial) = sum(vol .* ROI_mask_vol, 'all') / nnz(ROI_mask_vol);
> end

```

Calculate pattern change maps

⌚ Timing: 2 h

Note: The result of this part is a trial-by-trial image series (across the entire brain) that contains the multivariate change in BOLD activation patterns between an event in a trial (e.g. a screen with decision cues that elicit the calculation of expected values for each option) and *the same event in the next trial*. This is done using a searchlight algorithm, which averages the pattern change within a searchlight and writes the resulting pattern change index into the center voxel before moving to the next voxel.

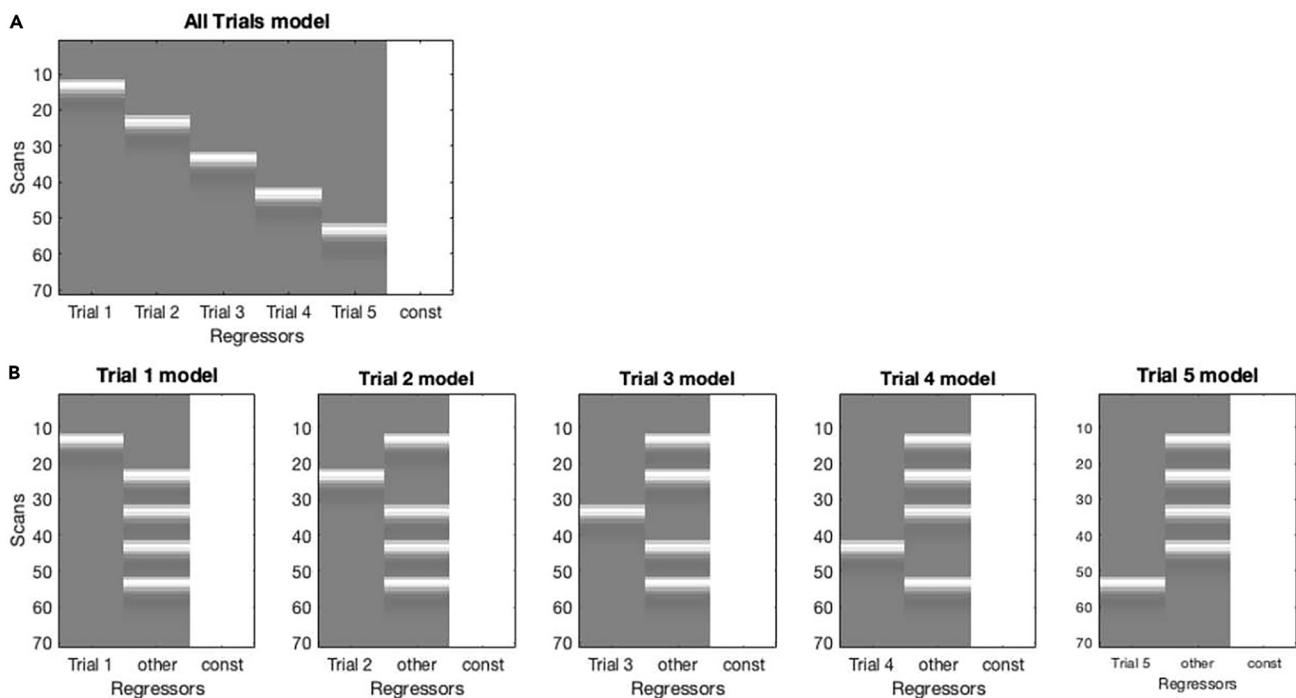


Figure 4. Examples of first-level design matrices for two different analysis approaches to obtain single-trial estimates of BOLD activation

(A) In *Least Square – All*, each trial event is included in the design matrix as an individual regressor, but all events are included in the same first level analysis.

(B) In *Least Square – Separate*, the event of interest is included in a separate regressor, whereas all other events are included in another regressor. This design is repeated for every trial. From each of these first level analyses the beta image of the regressor with the single event is retained for further analyses.

17. Obtain the BOLD activation at the time of the pattern event (see event A in [Figure 1](#)) by using a first-level analysis that estimates the BOLD activation for each trial separately.

Note: The Least-squares Separate approach (see step 15) is again favorable for the sensitivity of the analysis.

18. Employ a searchlight algorithm to calculate the pattern change of the whole brain in every trial. We used spherical searchlights with a radius of 10 mm (4 voxels).

Note: The searchlight radius should be large enough to include a sufficient number of voxels to facilitate meaningful pattern analyses, but small enough to retain spatial specificity. We found that a radius of 4 voxels is a good balance for the pattern change analysis presented in this protocol.

- a. The Decoding Toolbox (TDT)⁸ provides an adaptable framework for calculating multivariate pattern change in the context of a searchlight algorithm.

Note: Most searchlight algorithms are intended to be used for pattern classification. They only provide you with the patterns of all trials of a specified condition/class and discard information about the temporal relationship of patterns. However, for this analysis we need the patterns of all trials in chronological order.

- b. To achieve a chronological order of all trial, expand the TDT with information about the trial number for each pattern.
- c. Label all trial patterns with the same class/condition to pass them to the searchlight function in one batch.
- d. Within the searchlight function, iterate through all trial patterns in backwards order. This implies that the resulting pattern change of two adjacent trials t and $t+1$ will be stored at trial t .

Note: This order is beneficial for setting up the linear model in later analyses, because the pattern change $[t, t+1]$ is already aligned with its predicting BOLD activation in the seed ROI at trial t .

Note: The pattern change of the last trial in each run remains undefined (i.e. there is no trial to compare the last pattern with).

△ CRITICAL: The following script describes the above steps that should happen within the searchlight function. TDT passes the data of each searchlight (i.e. patterns of voxel activations for every trial) using the `current_data` 2D-array (the first dimension if this array are all trials and second dimension are all the voxels of the current searchlight sphere). The variable `current_decoding` indicates the index of the current searchlight sphere. When dealing with several runs you have to add another for-loop for all runs surrounding this code snippet.

```
> previous_pattern = NaN(length(current_data(1, :)), 1);
>
> for i_trial_forward = n_trials : -1 : 1
>
> if(isnan(previous_pattern))
> % pattern_change of the last pattern is undefined
```

```
> result(i_trial, current_decoding) = 0;
> else
> pattern_change = [calculate pattern change here];
> result(i_trial, current_decoding) = pattern_change;
> end
>
> previous_pattern = current_data(i_trial, :);
> end
```

Note: You can find the code that uses TDT functions for this analysis in the code repository of this protocol (<https://doi.org/10.5281/zenodo.8015561>).

19. An important choice is the multivariate distance measure that is used for calculating the pattern change. We choose to utilize a Fisher z-transformed correlation distance as explained below.
 - a. Calculate the Pearson correlation coefficient between the BOLD activation patterns in the searchlight between the same event in two adjacent trials (If some of the resulting correlation coefficients are undefined (i.e., NaN), refer to [troubleshooting 4](#)).

```
> correlation_matrix = corrcoef(searchlight_voxels(t, :), searchlight_voxels(t+1, :));
```

- b. Apply a Fisher z-transformation to improve the distribution of correlation coefficients toward a Gaussian distribution.

Note: [Figure 5](#) exemplifies the effect of this transformation: data points (voxel activation at both ends of the distribution are exaggerated, whereas those in the middle are mostly left unchanged). The Fisher-z transformation can be performed by calculating the `atanh()` of a correlation coefficient.

- c. Flip the sign of the transformed correlation distance.

Note: That way, two highly correlated patterns will result in a low amount of pattern change and vice versa.

```
> pattern_change = - atanh(correlation_matrix(1,2));
```

- d. Write the pattern change of the searchlight sphere to the center voxel.
20. Store the resulting pattern change maps of all trials in NIfTI format for subsequent analyses. This can be achieved with `spm_write_vol()` when using SPM.

Note: To store the data with the same volume settings (e.g. rotation, voxel size) as your input data it is easiest to use the header of existing NIfTI files (we used the `ROI_mask_header` here, [troubleshooting 5](#)).

```
> output_header = ROI_mask_header;
>
> for i_trial = 1:n_trials
```

```
> filename = sprintf('betas_%03d.nii', i_trial);
>
> output_header.fname = fullfile(output_dir, filename);
> output_header.private.dat.fname = output_header.fname;
> spm_write_vol(output_header, result(i_trial, :))
> end
```

Optional: There are different options when choosing the multivariate distance measure for pattern change calculation. We achieved best results when using the described Fisher z-transformed correlation distance, but here are examples of other possible distance measures to try: Mahalanobis distance: similar to the correlation distance, but it also considers the covariance between the whole distribution of voxel activations in the searchlight sphere over all trials. First, you have to calculate the covariance matrix of all patterns (all trials in a given searchlight) to then determine the Mahalanobis distance of each individual pattern to the whole distribution of patterns in that searchlight. Sum of squared differences: the simplest (and unscaled) distance measure corresponding to the cost function in ordinary least squares designs (e.g. multiple regression analysis)

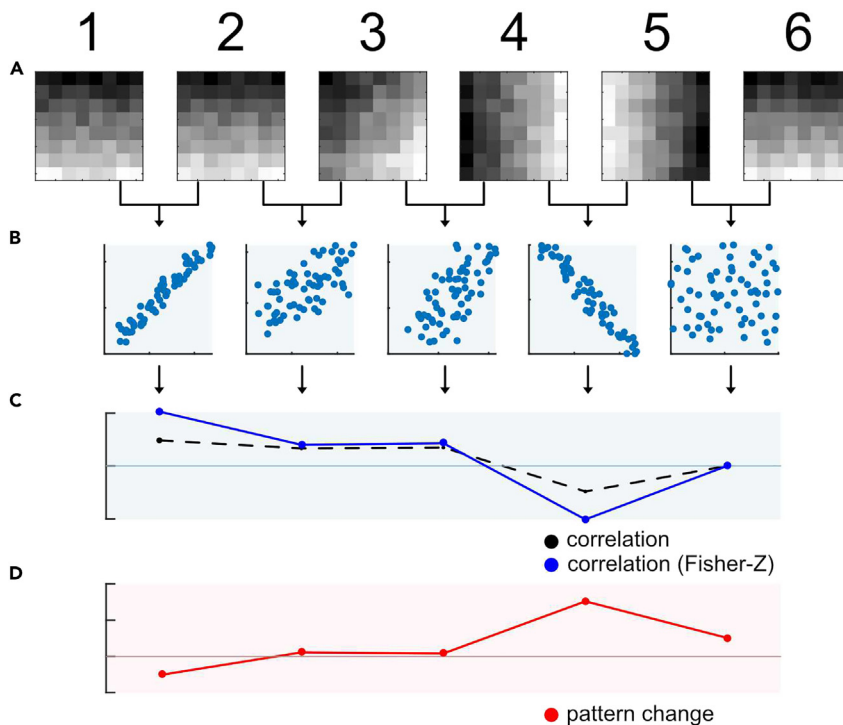


Figure 5. A correlation based pattern change metric

This figure illustrates how to calculate the magnitude of pattern change between two patterns.

(A) Simulated patterns of a searchlight in trials 1 to 6 with linear gradients and different gradient directions. Patterns 1 and 2 have the same direction. Patterns 2 and 3 as well as 3 and 4 have slightly different gradient directions (45°). Patterns 4 and 5 show completely opposite gradients (180°). Patterns 5 and 6 feature a medium change in gradients (90°).

(B) Scatter plots of patterns in adjacent trials to visualize the pattern correlation.

(C) Pearson's correlation coefficients (black dashed line) and Fisher-Z transformed correlation values (blue line).

(D) The pattern change metric (red line) is a measure of the magnitude of change. Similar patterns produce low pattern change values (patterns A and B) while inverse patterns result in high pattern change (patterns D and E).

Predict pattern change using BOLD of seed ROI

⌚ Timing: 30 min

Note: In this major step of the analysis we combine the extracted BOLD signal of the seed ROI and the whole brain pattern change maps described above in a univariate GLM to localize areas of the brain in which changes in neural patterns can be predicted by the seed ROI activity.

21. Specify a first-level GLM for each participant by regressing the trial-by-trial pattern change image series (dependent variable) onto the trial-by-trial BOLD activations from the seed ROI (independent variable, predictor):

```
> pattern_change ~ b0 + b1 * seed_ROI_BOLD
```

22. Estimate the model and include the resulting beta images (b1) in a second-level group analysis.
23. Identify brain areas in which changes in neural patterns can be significantly predicted by the BOLD activity of the seed ROI using a one-sample t-test in the second-level group analysis.

Note: If the statistical results do not survive family-wise error correction, but the statistical map shows wide-spread activations just below the statistical threshold, refer to [troubleshooting 2](#).

24. Visualize significant clusters by overlaying statistical results onto an anatomical image using established software packages (e.g., MRICron, MRICroGL).

Further analyses

Optional: Several analyses can be explored to further investigate the observed pattern changes, validate related hypotheses, or affirm the results of the pattern change predictions outlined in this protocol. All of the methods listed below are well established and thus we do not provide a detailed explanation for them in this protocol. Instead we want to illustrate ways to expand on the results gained in the primary analysis.

25. Investigate the temporal dynamics of pattern changes.
 - a. The trial-by-trial pattern change maps created in this protocol can be used to plot the trial-by-trial evolution of pattern changes in different brain regions at specific events in the task. [Figure 6A](#) plots the mean magnitude of pattern change over all subjects in the orbitofrontal cortex (OFC) before and after a reward reversal in the task.
26. Pattern classification: the pattern change analysis described in this protocol is capable of quantifying differences in neural information content over time. Other analysis strategies have to be utilized to gain insights about the information encoded in those patterns. This can be achieved through a decoding analysis (multivariate classification) or representational similarity analysis.⁹
27. Permutation analysis: confirm the results of the primary pattern change analysis by employing a permutation test in the final GLM. The concept is to simulate an analysis with many randomly shuffled combinations of the input data and then compare the test-statistic of the analysis using the real input data against the distribution of permutations (see [Figure 6B](#)).
 - a. During the first-level analysis (single subject): within each subject, permute the trial-by-trial BOLD in seed ROI (predictor) while keeping the true order of trial-by-trial pattern changes.
 - b. Calculate the second-level group analysis using all subjects and store the resulting second-level t-statistics of all permutations.
 - c. Compare the second-level t-statistic of the real analysis against the distribution of permuted t-statistics.

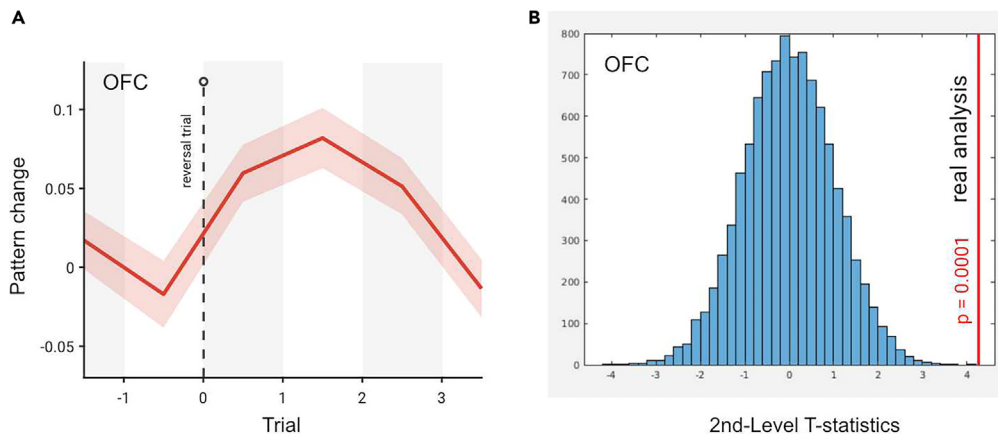


Figure 6. Further analyses to consider

(A) Investigate the trial-by-trial temporal dynamics of pattern changes in different brain regions at important events in the task. This figure shows the pattern change in the orbitofrontal cortex (OFC) after a reward reversal in the task. The red line depicts the mean pattern change over all subjects and the light red area illustrates the standard error. (B) Confirm the results of the primary pattern change analysis using a permutation test in which the trial-wise predictor variable (BOLD in seed ROI) is randomly shuffled during the first-level analysis within each subject. Then plot a histogram of the second-level T-statistic and compare the T-statistic of the real data against the distribution of permutations (here we simulated the analysis with $n = 10000$ random permutations).

28. Conjunction analysis: when testing multiple hypotheses you can test the statistical significance of a combination of several analyses throughout the brain by employing a true conjunction analysis.¹⁰

Note: Conjunction analyses must not be performed on threshold-free cluster enhancement (TFCE) enhanced results.

EXPECTED OUTCOMES

After performing all analyses described above, the expected outcome is a statistical map of the whole brain displaying the significance that pattern change in a brain region is predicted by the preceding BOLD signal strength of the seed ROI. Figure 7 depicts an overview of all major analysis steps and the final GLM, which produces the statistical map.

LIMITATIONS

The multivariate analysis outlined in this protocol can draw conclusions about the magnitude of pattern changes in different areas of the brain, the temporal dynamics of these changes (e.g., for how many trials the pattern change persists) and the relationship to potential seed regions inside the brain. It does not make assumptions about the information contained in the observed patterns. To explore the information content of patterns other existing analysis types can be employed (see further analyses).

The results of using this analysis strategy can also be impaired by different factors. These include weak BOLD signal in seed and target regions, problems during seed ROI definition and insufficient statistical power.

TROUBLESHOOTING

Problem 1

During the first data analysis you cannot obtain robust statistical group-analysis results for simple contrasts (e.g., motor contrasts) (step 8).

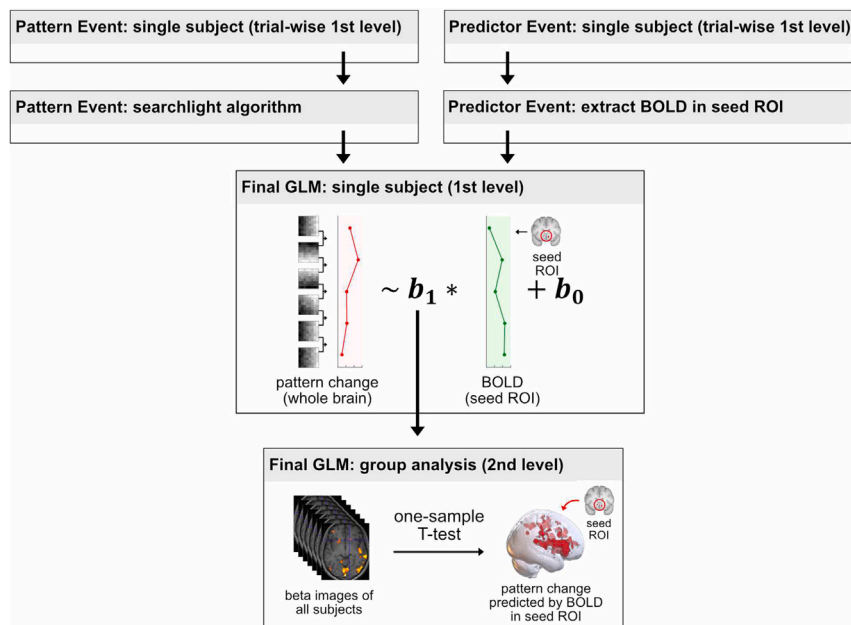


Figure 7. Overview of all major analyses

Schematic overview of all major steps and the final GLM to predict changes in neural patterns using BOLD in a seed ROI. The final GLM regresses the trial-by-trial pattern change in every voxel of the brain (dependent variable) onto the trial-by-trial BOLD responses of the seed ROI (independent variable). The resulting beta images of all subjects are then tested for statistical significance in a group analysis using a one-sample T-test.

Potential solution

- It might be worthwhile to manually check if the normalization process worked for all participants. You can check this by overlaying all normalized images in a viewer like MRICroGL.
- Also check if the order of your NIfTI files is correct and potential dummy scans at the start of the experiment have been removed properly.

Problem 2

The statistical results of the final second-level pattern change GLM (or other GLMs, e.g., for functional seed ROI definition) do not survive family-wise error correction (steps 13 and 23).

Potential solution

Depending on the shape of the activation clusters (e.g., wide-spread and rather flat vs. spiky with smaller peaks in a larger cluster), TFCE (threshold-free cluster enhancement), an iterative permutation-based correction for multiple comparisons, can be used.¹¹

Problem 3

The Least-squares Separate method requires the estimation of one whole GLM per trial in the experiment. It may be extremely time and memory consuming (step 15).

Potential solution

Use parallel computing to estimate GLMs of all trials in the experiment. When a GLM is finished, directly copy the beta image of interest (i.e., the beta image of the current trial in its own regressor) into a new folder and discard the GLM to save memory.

Problem 4

The pattern change searchlight algorithm produces NaN values in the trial-by-trial pattern change maps of some subjects (step 19).

Potential solution

- `Corrcoef()` gives NaN values as an output, when one of the input variables contains a NaN. You can use the option `corrcoef(X, Y, 'rows', 'complete')` or `corrcoef(X, Y, 'rows', 'pairwise')` to ignore NaN in the input. You should also check the cause for the existence of NaN values in the input and fix potential problems.
- `Corrcoef()` also produces NaN values, when at least one of the input patterns contains only identical values (i.e., the standard deviation of one pattern is zero). This results in a divide by zero computation and the correlation of these two patterns is undefined. If this impairs the second level GLM results and happens rarely you can try to replace these NaN values with zeros. If there are many undefined correlation coefficients you should inspect the potential cause of that problem.

Problem 5

The stored NIfTI image has the wrong orientation or is stored in integer format so that information is lost (step 20).

Potential solution

Make sure to use the correct header file when manually saving the NIfTI images. When saving the results transforms your data into integer format you have to use the header of one of the realigned and normalized images instead of the mask image header.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jan Gläscher (glaescher@uke.de).

Technical contact

Questions about the technical specifics of performing the protocol should be directed to and will be answered by the technical contact, Leon Möhring (l.moehring@uke.de).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Behavioral data have been deposited at Zenodo and are publicly available as of the date of publication. DOIs are listed in the [key resources table](#). Imaging data reported in this paper will be shared by the [lead contact](#) upon request.
- All original code has been deposited at Zenodo and is publicly available as of the date of publication. DOIs are listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

ACKNOWLEDGMENTS

This work was funded by the German Research Foundation (SFB TRR 169 "Cross-modal Learning").

AUTHOR CONTRIBUTIONS

J.G. and L.M. designed the experimental task, developed analytical protocols, and wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Möhring, L., and Gläscher, J. (2023). Prediction errors drive dynamic changes in neural patterns that guide behavior. *Cell Rep.* 42, 112931. <https://doi.org/10.1016/j.celrep.2023.112931>.
2. Durnez, J., Degryse, J., Moerkerke, B., Seurinck, R., Sochat, V., Poldrack, R.A., and Nichols, T.E. (2016). Power and sample size calculations for fMRI studies based on the prevalence of active peaks. Preprint at bioRxiv. <https://doi.org/10.1101/049429>.
3. Mumford, J.A. (2012). A power calculation guide for fMRI studies. *Soc. Cogn. Affect. Neurosci.* 7, 738–742. <https://doi.org/10.1093/scan/nss059>.
4. Deichmann, R., Gottfried, J.A., Hutton, C., and Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage* 19, 430–441. [https://doi.org/10.1016/s1053-8119\(03\)00073-9](https://doi.org/10.1016/s1053-8119(03)00073-9).
5. Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., and Turner, R. (2002). Image distortion correction in fMRI: A quantitative evaluation. *Neuroimage* 16, 217–240. <https://doi.org/10.1006/nimg.2001.1054>.
6. Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage* 38, 95–113. <https://doi.org/10.1016/j.neuroimage.2007.07.007>.
7. Mumford, J.A., Turner, B.O., Ashby, F.G., and Poldrack, R.A. (2012). Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage* 59, 2636–2643. <https://doi.org/10.1016/j.neuroimage.2011.08.076>.
8. Hebart, M.N., Gørgen, K., and Haynes, J.-D. (2014). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Front. Neuroinform.* 8, 88. <https://doi.org/10.3389/fninf.2014.00088>.
9. Freund, M.C., Etzel, J.A., and Braver, T.S. (2021). Neural Coding of Cognitive Control: The Representational Similarity Analysis Approach. *Trends Cogn. Sci.* 25, 622–638. <https://doi.org/10.1016/j.tics.2021.03.011>.
10. Nichols, T., Brett, M., Andersson, J., Wager, T., and Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage* 25, 653–660. <https://doi.org/10.1016/j.neuroimage.2004.12.005>.
11. Smith, S.M., and Nichols, T.E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44, 83–98. <https://doi.org/10.1016/j.neuroimage.2008.03.061>.

2. Description of the publication with references

2.1 Introduction

Reinforcement learning (RL) encompasses a group of algorithms focused on guiding intelligent agents in selecting actions to maximize accumulated rewards (Sutton and Barto 2018). These algorithms have had a significant impact on computational neuroscience by serving as effective models for understanding the learning and decision-making mechanisms in the brain. When navigating the complexities of the world, organisms, including humans, rely on responsive behaviors driven by internal models of their environment (Dayan and Daw 2008). These internal models encode expectations, enabling the anticipation of future events and facilitating efficient responses. In dynamic environments, where the world undergoes constant changes, adaptive adjustments in behavioral strategies are essential. Learning, in this context, revolves around minimizing the disparity between internal expectations and real-world outcomes. Prediction errors (PEs) play a pivotal role in this process, serving as computational signals representing deviations from expected outcomes and acting as instructive cues in reinforcement learning paradigms.

In the realm of RL, the continual updating of prediction errors and expected values is crucial for driving the learning process. Ideally, this iterative adaptation results in a convergence where internal expectations align perfectly with reality, and prediction errors approach zero. However, the intricate neural mechanisms underlying these adjustments in a biological system, such as the human brain, remain a profound question. Beyond the changes in the representations of core computational variables, the neural implementation of this process may also involve alterations in other cognitive systems (Niv 2009). Attentional resources, working memory, and motor planning likely contribute to the reshaping of internal models following a prediction error.

Over the past two decades, considerable research has delved into the neural correlates of prediction errors (Garrison et al. 2013). However, none of these studies investigated the subsequent changes in activity patterns in brain networks that underlie reinforcement learning. This study integrates model-derived univariate and multivariate functional magnetic resonance imaging (fMRI) analyses, unveiling the short-term influence of prediction errors on patterns of neural activity in the human brain (Möhring and Gläscher 2023).

2.2 Model-free and model-based Reinforcement learning

To increase the specificity of observed changes in neural patterns induced by prediction errors, this study employs two distinct types of such errors. Using a task designed to evoke

orthogonal prediction errors, wherein one prediction error is large while the other is not, our goal is to establish a form of in-brain control group. Thus, enabling us to distinguish genuine error-related pattern changes from general effects of arousal. The two types of prediction errors used here are the reward and state prediction error.

In the origins of research on reinforcement learning Schultz and colleagues were the first to observe a prediction error signature in neural recordings (Schultz 2019). They made the remarkable discovery that the phasic activity of midbrain dopaminergic neurons encodes a reward prediction error (RPE), which serves as the central teaching signal in model-free reinforcement learning (Schultz et al. 1997).

Model-free RL is essentially rooted in a habitual principle, where actions followed by a reinforcer tend to be repeated in the future (Jocham et al. 2011, Lerner et al. 2021). However, beyond this habitual aspect, there is evidence suggesting the development and mental simulation of an internal task model to enhance effective decision-making (Daw et al. 2011, Lee et al. 2014). From a computational perspective, these learning mechanisms are viewed as two distinct forms of reinforcement learning. Model-free reinforcement learning focuses exclusively on learning the values of potential actions in each state. The driving force behind model-free learning is the reward prediction error, which arises when anticipated rewards are not met. In contrast, model-based reinforcement learning aims to construct a mental representation of the state space and potential state transitions. Initially, the acquisition of information about the world's structure is independent of reward-related information. The central element responsible for updating expectations in model-based learning is the state prediction error (SPE), which occurs when there are deviations from the expected upcoming state (Gläscher et al. 2010). Subsequently, this knowledge of state transitions can be integrated with rewards, leading to a more sophisticated decision-making process.

In an earlier study, Glaescher et al. utilized a probabilistic Markov decision task to distinguish between two distinct neural correlates of model-based and model-free reinforcement learning (Gläscher et al. 2010). They employed computational modeling with a novel hybrid-learner, which integrated value estimates from both model-based and model-free learning algorithms using a free weighting parameter. Through this hybrid approach, they identified reward prediction errors in the ventral striatum (VS), replicating previous findings. Additionally, evidence for neural correlates of the state prediction error was observed in the intraparietal sulcus (IPS) and the dorsolateral prefrontal cortex (dlPFC). Importantly, they demonstrated that these signals were not merely related to general attention or salience but were better explained by the concept of model-based state prediction errors. These findings

suggest that human decision-making is influenced by a combination of these learning systems, and error signals from each system are computed in distinct brain areas.

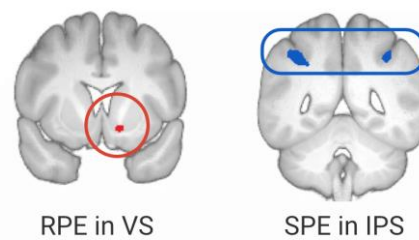


Figure 1. Neural correlates of the reward and state prediction error in the human brain. Using fMRI the RPE can be observed in BOLD signals of the ventral striatum (red) while the SPE is located in the intraparietal sulcus (blue). The figure is adapted from Möhring and Gläscher (2023).

2.3 Neural correlates of the reward prediction error

The reward prediction error signal initially observed by Schultz et al. is being distributed throughout the brain by dopaminergic neurons in the ventral tegmental area (VTA). On a rapid sub-second scale, known as phasic dopamine activity, these dopaminergic neurons encode highly time-specific reward prediction errors and impact swift behavioral responses (Schultz et al. 1997). The dopamine reactions can be directly assessed through electrophysiological methods and voltammetry. Using fMRI one can detect RPE signatures in blood oxygenation level-dependent (BOLD) responses of the ventral striatum (O'Doherty et al. 2003). The BOLD activations in this region likely reflects peri-synaptic activity caused by projections of midbrain dopaminergic neurons.

In addition to transmitting the reward prediction error signal, dopamine is involved in multiple other cognitive systems. Slower alterations in dopamine levels, spanning seconds to minutes, correlate with diverse brain functions such as stress, reward, punishment, attention, and movement (Büchel et al. 2017, Datta et al. 2002, Howe et al. 2013, Young 2004). These gradual shifts in dopamine levels likely occur independently of rapid phasic activations and some theories suggest they may arise from presynaptic interactions or slower changes in impulses (Anzalone et al. 2012, Threlfell et al., 2012). Longer-lasting fluctuations in dopamine concentrations might modulate sensitivity to phasic dopamine impulses (Grace et al. 2007). Additionally, dopamine exerts a continuous influence on postsynaptic neurons and dopamine receptors. In Parkinson's disease, the depletion of dopamine levels can be alleviated through pharmacological activation of dopamine receptors, which mimics healthy tonic dopamine levels but fails to trigger phasic responses (Frank 2005). Tonic dopamine

effects are associated with movement, cognition, and motivation (Schultz 2016). In conclusion, dopamine contributes to a variety of neuronal processes and affects behavior across phasic, intermediate, and tonic timeframes.

2.4 The BOLD response and model-informed fMRI analyses

In functional magnetic resonance imaging, the disparity in magnetic properties between oxygenated and deoxygenated hemoglobin is exploited to detect alterations in brain regions' oxygenation and perfusion over time, stemming from preceding neuronal activity. The resultant measure in fMRI is termed the blood oxygenation level-dependent signal, which responds to neuronal activations with a delay of approximately 4 to 6 seconds. To infer neuronal activity from its effects on the BOLD signal, researchers utilize a hemodynamic response function (HRF) to model the data. The HRF aims to characterize the temporal relationship between neuronal activations and the resulting increase in oxygenation due to vessel dilation and enhanced perfusion (Huettel et al. 2009).

The statistical framework used for investigating functional neuroimaging data is known as statistical parametric mapping (SPM). Within this framework, a parametric statistical model is applied to each voxel of the collected fMRI data using the general linear model (GLM). The conventional approach for analyzing fMRI data is known as univariate analysis, where each voxel is examined separately (Friston et al. 1994). The objective is to characterize the variability in the data in terms of both experimental effects of interest and confounding effects, while accounting for some level of residual variability. First, multiple preprocessing steps are applied to the collected fMRI data to clean images from various potential noise sources. This typically involves correcting artifacts caused by head movement, transforming the data of all participants to a shared normalized space and enhancing the signal-to-noise ratio by applying slight smoothing.

A design matrix is constructed, which includes regressors at defined onsets during the task for all experimental events as well as for potential confounding events in the data. Additionally, a regressor can be modulated in magnitude over time by applying a parametric modulator to it, often used to represent varying strengths of neuronal processes. These regressors are then convolved with the hemodynamic response function, and the resulting GLM is estimated based on the acquired BOLD data (Friston et al. 1994). Given the large number of statistical tests conducted for all voxels, it is crucial to address the issue of multiple comparisons to control false positive rates.

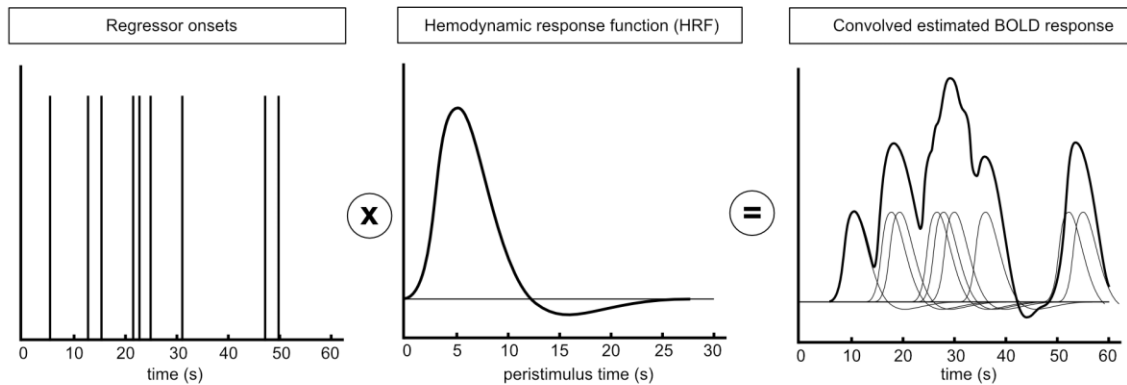


Figure 2. The hemodynamic response function and estimated BOLD signals.

The regressors of the design matrix are convolved with the hemodynamic response function to emulate a hypothetical combined BOLD response which would result from the hypothesized underlying neural activations throughout the task. Subsequently, a general linear model is used to test the hypothesis of the convolved BOLD response being present for every voxel of the brain while accounting for some level of residual variability.

Model-informed fMRI represents a powerful method for testing hypotheses regarding the underlying neuronal computations involved in cognitive systems of the human brain. Computational models, such as Reinforcement learning, can be calibrated to the behavioral data of participants gathered during the experiment and subsequently assessed based on the degree to which the model mirrors observed behaviors. The trial-by-trial time course of a variable (i.e. prediction errors) derived from this computational model is then employed as a parametric modulator assigned to a regressor in the GLM (Gläscher and O’Doherty 2010). The model-informed fMRI analysis results in a statistical map of the entire brain, highlighting regions where the observed BOLD signal corresponds with the expected convolved regressors of the computational model.

2.5 Predicting multivariate change of brain patterns using PEs

The advent of multivariate methods has provided a novel perspective on stimulus representations in the human brain, unveiling the information embedded in spatially distributed patterns of neural activity, as measured by functional magnetic resonance imaging (Haynes 2015, Kahnt 2018, Mahmoudi et al. 2012). While some studies have utilized machine learning algorithms to decode information about stimuli (Liu et al. 2022, Wilbertz et al. 2017, Wittkuhn and Schuck 2021), others have explored the (dis)similarity of

activation patterns in relation to perceptual or categorical similarity (Freund et al. 2021, Schuck et al. 2016, Yan et al. 2016). This approach has extended to the investigation of learning-induced changes in the similarity of activation patterns (Howard and Kahnt 2018).

Prediction errors, as quantitative signals reflecting the deviation of expectations from environmental stimuli, necessitate proportional changes in the neural activity patterns encoding these internal representations. This study adopts a unique approach, simultaneously employing two distinct types of prediction errors: reward prediction errors and state prediction errors. This dual-error paradigm enables the identification of brain regions exclusively modulated by a specific type of prediction error, shedding light on error-specific changes in neural patterns.

The analytical approach developed in this study integrates model-informed univariate and multivariate fMRI analyses, unveiling the influence of prediction errors on subsequent changes in patterns of neural activity. This method dynamically quantifies the magnitude of multivariate changes in neural patterns over time and predicts these pattern changes using trial-by-trial BOLD responses in error-coding regions (RPEs in VS and SPEs in IPS). The study introduced a metric 'pattern change' to quantify multivariate changes in neural patterns over time. Searchlight-based analyses were used to calculate whole-brain maps of pattern change for each trial (Möhring and Gläscher 2023).

A searchlight analysis in fMRI involves systematically moving a small spherical "searchlight" across each voxel in the brain, performing a local analysis within each searchlight. The size of the searchlight, typically defined by a radius, determines the number of neighboring voxels included in the analysis. At each voxel location, the surrounding voxels within the chosen radius are considered, and a local analysis is conducted to characterize the patterns of activity within the searchlight. In this study, we employed a correlation-based metric to assess patterns of neighboring trials within the same searchlight, allowing us to quantify the degree of pattern change over time. The results of the local analysis are then mapped back onto the original brain space, creating a spatial map showing the outcome of the analysis at each voxel location (Etzel et al. 2013).

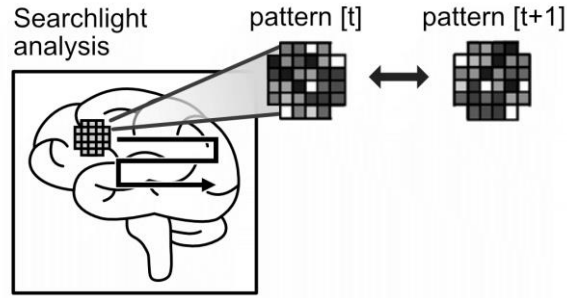


Figure 3. Schematic of a searchlight analysis.

This figure illustrates a searchlight analysis on fMRI data. The analysis utilizes a small spherical "searchlight" which is systematically moved across each voxel in the brain. Within each searchlight, we performed a local analysis comparing neural patterns from adjacent trials, utilizing a correlation-based distance metric. The outcomes of this local analysis were then mapped back onto the center voxels of each searchlight sphere.

There are numerous options of multivariate distance measures that can be used for assessing pattern change. In our methodology, we adopt a Fisher z-transformed correlation distance for this purpose. Firstly, we calculate the Pearson correlation coefficient between the BOLD activation patterns in the searchlight for corresponding events in adjacent trials, with provisions to handle undefined correlation coefficients. Subsequently, we apply the Fisher z-transformation to these correlation coefficients to normalize their distribution towards a Gaussian distribution and allow for statistical evaluation of the pattern change metric. This transformation involves taking the $\text{arctanh}()$ of each correlation coefficient. Afterward, we flip the sign of the transformed correlation distance, ensuring that highly correlated patterns correspond to low pattern change and vice versa (Möhring and Gläscher 2024).

$$pattern\ change[t, t + 1] = -\text{arctanh}(\text{corr}(\text{pattern}[t], \text{pattern}[t + 1]))$$

Finally, we write the pattern change of the searchlight sphere to the center voxel and store the resulting pattern change maps of all trials.

2.6 Results of the publication

In this study, 44 participants completed a two-step Markov decision task over five runs. Initially, participants made decisions leading to transitions into either a blue or red state. Subsequently, in the second step, participants selected one of two boxes, each offering

either 0 €, 4 €, or 6 €. The task incorporated reversal and catch trials to induce unexpected changes, forcing participants to update their internal task model throughout the experiment. Hereby, prediction errors in model-based (SPE) and model-free (RPE) learning emerge. Importantly, the task design produced orthogonal SPEs and RPEs, allowing independent investigation of their effects.

To model participant behavior, a hybrid-learning algorithm blending model-based and model-free reinforcement learning was employed, emphasizing the interplay between both learning strategies during task execution. This hybrid approach utilized a constant weighting parameter to specify the impact of each learning strategy for an individual participant. During model evaluation, the hybrid-learning algorithm outperformed pure model-free and model-based approaches.

Neural correlates of SPEs were localized in the intraparietal sulcus and precuneus (PCu), while RPEs were identified in the ventral striatum. Notably, these results replicate the findings of an earlier study by Gläscher et al. (Gläscher et al. 2010). Regions of interest (ROIs), which will serve as seed regions for later predictive multivariate analyses, were defined based on a conjunction of those findings and the present study.

The study introduced a metric 'pattern change' to quantify multivariate changes in neural patterns over time. Our results revealed that the mean pattern change over all trials shows a distinct distribution different from the distribution of mean change in blood oxygenation level-dependent signal strength throughout the brain. This suggests that the introduced pattern change metric does capture information varying from pure univariate BOLD signals.

Trial-wise BOLD responses induced by SPEs and RPEs were extracted in error coding ROIs and utilized to predict subsequent pattern change throughout the brain. This pattern change analyses unveils that RPE-evoked BOLD responses in the ventral striatum predicted subsequent pattern changes in orbitofrontal cortex, anterior cingulate cortex (ACC), dorsolateral prefrontal cortex, and bilateral insula. Contrarily, SPE-evoked BOLD responses in the intraparietal sulcus predicted pattern changes in superior parietal lobule (SPL), precuneus, supplementary motor area (SMA), and the intraparietal sulcus itself. Our results reveal short-term (20 - 30 s) effects of prediction errors on neural activation patterns in the brain. Importantly, these findings demonstrate a dissociation between RPE- and SPE-driven pattern changes, supporting the idea that distinct PEs influence different neural networks.

Additionally, we explored the information content of evolving neural patterns. Clusters in the anterior cingulate cortex and ventromedial prefrontal cortex (vmPFC) coded for the value of the chosen action at reward-decision. The precuneus contained information about expectations of the next state at the time of state-decision.

Importantly, the introduced pattern change metric significantly predicted subsequent adaptations in behavioral policy of participants. Clusters in the primary motor cortex (M1), ventromedial prefrontal cortex, bilateral insula, and supplementary motor area were found to correlate with alterations in action probabilities for state-decisions. Furthermore, clusters in the orbitofrontal cortex, and dorsolateral prefrontal cortex were linked to changes in action probabilities for reward-decisions.

Conjunction analyses identified brain regions where pattern change, modulated by PE-evoked BOLD responses, predicted ensuing behavioral adaptations. Specifically, OFC exhibited such conjunction for reward-decision and RPEs, while vmPFC showed it for state-decision and SPEs. These findings underscore the intricate relationship between prediction errors, neural activation patterns, and subsequent behavioral adjustments.

2.7 Discussion

In our study, we delved into the intricate workings of reinforcement learning in the human brain, focusing on the pivotal role of prediction errors in orchestrating dynamic changes in neural patterns that ultimately guide behavior. Despite its critical importance, the mechanisms leading to the transformation of these neural representations have been largely overlooked in previous research.

Our findings show that prediction errors, specifically two distinct types – model-based SPEs and model-free RPEs – play a central role in driving short-term changes in neural activation patterns. These prediction errors exert specific effects on two learning systems, inducing alterations in neural patterns of particular brain regions. Notably, we established a direct link between the magnitude of pattern change and subsequent modifications in participants' behavioral policies, underscoring the profound impact of prediction errors on both neural representations and behavior.

Our research sheds light on the segregation of these two prediction error types in different brain regions. RPEs were found to modulate pattern changes in regions associated with reward-related processes, including the orbitofrontal cortex, dorsolateral prefrontal cortex, and anterior cingulate cortex. These regions are known for encoding expected rewards, attributing values to stimuli, and participating in decision-making based on reward predictions (Hare et al. 2008, Klein-Flügge et al. 2013, Li et al. 2016, Yan et al. 2016). The involvement of ACC in encoding the value of chosen actions further emphasizes its crucial role in decision-making processes (Becker et al. 2016, Silvetti et al. 2013, Vassena et al. 2014).

Conversely, SPEs were associated with pattern changes in the Precuneus and superior parietal lobule. The Precuneus, in particular, emerged as a key player in conscious processing of external information based on internal beliefs and updates to the internal model of the environment (Bzdok et al. 2015, Lyu et al. 2021).

We also explored the distribution of RPE signals throughout the brain, highlighting the role of dopamine in mediating these effects. This suggests a potential biochemical substrate for the observed effects on neural activity patterns, implicating dopamine's ability to modulate short-term synaptic plasticity (Brzosko et al. 2017, Condon et al. 2019, Yagishita et al. 2014). Our study supports the idea that prediction errors are involved in priming brain regions for subsequent functional reconfigurations, providing insights into the neurobiological underpinnings of learning.

A major aspect of our findings is the connection between pattern changes in the orbitofrontal cortex and ventromedial prefrontal cortex and subsequent adaptations in behavioral policy of reward-related decisions. This aligns with existing research indicating that OFC value representations inform the general value signal constructed in vmPFC, guiding goal-directed behavior (Algermissen et al. 2022, Howard and Kahnt 2017, Lee et al. 2021, Schuck et al. 2016, Vaidya and Badre 2020). Additionally, we observed correlations between pattern changes in motor-related regions (M1 and SMA) and changes in behavioral policy of state-related decisions, suggesting a link between neural activity patterns, action planning, and execution.

Our study contributes to the understanding of evolving multivoxel patterns of BOLD activity during learning, emphasizing the temporal changes in neural activation patterns. Leveraging multivariate approaches, trial-by-trial searchlight algorithms, and whole-brain general linear models, we provide a comprehensive account of the interplay between prediction error signals, neural activity patterns, and behavioral adaptations. The proposed pattern change metric captures differences in neural information content over time, offering valuable insights into the nature of neural activation patterns.

In conclusion, our work offers insights into the mechanisms through which prediction errors drive short-term reconfigurations of neural representations, influencing subsequent shifts in behavioral policy. The dual perspective on neural patterns, considering both their temporal changes and the information they encode, contributes to a more nuanced understanding of learning within a theoretical framework that views learning as the process of updating internal models for more accurate predictions in the future.

2.8 References

- Algermissen J, Swart JC, Scheeringa R, Cools R, den Ouden HEM (2022) Striatal BOLD and Midfrontal Theta Power Express Motivation for Action. *Cereb Cortex*. 32(29):24–42.
- Anzalone A, Lizardi-Ortiz JE, Ramos M, De Mei C, Hopf FW, Iaccarino C, et al. (2012) Dual control of dopamine synthesis and release by presynaptic and postsynaptic dopamine D2 receptors. *J Neurosci*. 32(90):23–34.
- Becker MPI, Nitsch AM, Hewig J, Miltner WHR, Straube T (2016) Parametric modulation of reward sequences during a reversal task in ACC and VMPFC but not amygdala and striatum. *Neuroimage*. 143(50):7.
- Brzosko Z, Zannone S, Schultz W, Clopath C, Paulsen O (2017) Sequential neuromodulation of Hebbian plasticity offers mechanism for effective reward-based navigation. *eLife*. 6:e27756.
- Büchel C, Peters J, Banaschewski T, Bokde ALW, Bromberg U, Conrod PJ, et al. (2017) Blunted ventral striatal responses to anticipated rewards foreshadow problematic drug use in novelty-seeking adolescents. *Nat Commun*. 8:14140.
- Bzdok D, Heeger A, Langner R, Laird AR, Fox PT, Palomero-Gallagher N, et al. (2015) Subspecialization in the human posterior medial cortex. *Neuroimage*. 106:55–71.
- Condon MD, Platt NJ, Zhang Y-F, Roberts BM, Clements MA, Vietti-Michelina S, et al. (2019) Plasticity in striatal dopamine release is governed by release-independent depression and the dopamine transporter. *Nat Commun*. 10:4263.
- Datla KP, Ahier RG, Young AMJ, Gray JA, Joseph MH (2002) Conditioned appetitive stimulus increases extracellular dopamine in the nucleus accumbens of the rat. *Eur J Neurosci*. 16(19):87–93.
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 69(12):04–15.
- Dayan P, Daw ND (2008) Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*. 8(4):29–53.
- Etzel JA, Zacks JM, Braver TS (2013) Searchlight analysis: promise, pitfalls, and potential. *Neuroimage*. 78(26):1–9.
- Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. *J Cogn Neurosci*. 17:51–72.
- Freund MC, Etzel JA, Braver TS (2021) Neural Coding of Cognitive Control: The Representational Similarity Analysis Approach. *Trends Cogn Sci*. 25(6):22–38.
- Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ (1994) Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*. 2:189–210.
- Garrison J, Erdeniz B, Done J (2013) Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev*. 37(1):297–310.
- Gläscher J, Daw N, Dayan P, O'Doherty JP (2010) States versus Rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*. 66(5):85–95.
- Gläscher JP, O'Doherty JP. Model-based approaches to neuroimaging: combining reinforcement learning theory with fMRI data. *WIREs Cognitive Science* 2010;1:501–10. <https://doi.org/10.1002/wcs.57>.
- Grace AA, Floresco SB, Goto Y, Lodge DJ. Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends Neurosci* 2007;30:220–7. <https://doi.org/10.1016/j.tins.2007.03.003>.

- Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A. Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. *J Neurosci* 2008;28:5623–30. <https://doi.org/10.1523/JNEUROSCI.1309-08.2008>.
- Haynes J-D. A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. *Neuron* 2015;87:257–70. <https://doi.org/10.1016/j.neuron.2015.05.025>.
- Howard JD, Kahnt T. Identity prediction errors in the human midbrain update reward-identity expectations in the orbitofrontal cortex. *Nat Commun* 2018;9:1611. <https://doi.org/10.1038/s41467-018-04055-5>.
- Howard JD, Kahnt T. Identity-Specific Reward Representations in Orbitofrontal Cortex Are Modulated by Selective Devaluation. *J Neurosci* 2017;37:2627–38. <https://doi.org/10.1523/JNEUROSCI.3473-16.2017>.
- Howe MW, Tierney PL, Sandberg SG, Phillips PEM, Graybiel AM. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* 2013;500:575–9. <https://doi.org/10.1038/nature12475>.
- Huettel SA, Song AW, McCarthy G. *Functional Magnetic Resonance Imaging*. Oxford University Press, Incorporated; 2009.
- Jocham G, Klein TA, Ullsperger M. Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *J Neurosci* 2011;31:1606–13. <https://doi.org/10.1523/JNEUROSCI.3904-10.2011>.
- Kahnt T. A decade of decoding reward-related fMRI signals and where we go from here. *Neuroimage* 2018;180:324–33. <https://doi.org/10.1016/j.neuroimage.2017.03.067>.
- Klein-Flügge MC, Barron HC, Brodersen KH, Dolan RJ, Behrens TEJ. Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. *J Neurosci* 2013;33:3202–11. <https://doi.org/10.1523/JNEUROSCI.2532-12.2013>.
- Lee S, Yu LQ, Lerman C, Kable JW. Subjective value, not a gridlike code, describes neural activity in ventromedial prefrontal cortex during value-based decision-making. *Neuroimage* 2021;237:118159. <https://doi.org/10.1016/j.neuroimage.2021.118159>.
- Lee SW, Shimojo S, O'Doherty JP. Neural computations underlying arbitration between model-based and model-free learning. *Neuron* 2014;81:687–99. <https://doi.org/10.1016/j.neuron.2013.11.028>.
- Lerner TN, Holloway AL, Seiler JL. Dopamine, Updated: Reward Prediction Error and Beyond. *Curr Opin Neurobiol* 2021;67:123–30. <https://doi.org/10.1016/j.conb.2020.10.012>.
- Li Y, Vanni-Mercier G, Isnard J, Mauguière F, Dreher J-C. The neural dynamics of reward value and risk coding in the human orbitofrontal cortex. *Brain* 2016;139:1295–309. <https://doi.org/10.1093/brain/awv409>.
- Liu Y, Nour MM, Schuck NW, Behrens TEJ, Dolan RJ (2022) Decoding cognition from spontaneous neural activity. *Nat Rev Neurosci*. 23(20):4–14.
- Lyu D, Pappas I, Menon DK, Stamatakis EA (2021) A Precuneal Causal Loop Mediates External and Internal Information Integration in the Human Brain. *J Neurosci*. 41(99):44–56.
- Mahmoudi A, Takerkart S, Regragui F, Boussaoud D, Brovelli A (2012) Multivoxel pattern analysis for fMRI data: a review. *Comput Math Methods Med*. 2012:961257.
- Möhring L, Gläscher J (2024) Protocol for predicting multivariate change of brain patterns using model-informed fMRI activations. *STAR Protoc*. 5:102978.
- Möhring L, Gläscher J (2023) Prediction errors drive dynamic changes in neural patterns that guide behavior. *Cell Rep* 42:112931.

- Niv Y (2009) Reinforcement learning in the brain. *Journal of Mathematical Psychology*. 53(1):39–54.
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. *Neuron*. 38(3):29–37.
- Schuck NW, Cai MB, Wilson RC, Niv Y (2016) Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron*. 91(140):2–12.
- Schultz W (2019) Recent advances in understanding the role of phasic dopamine activity. *F1000Res*. 8:F1000 Faculty Rev-1680.
- Schultz W (2016) Dopamine reward prediction-error signalling: a two-component response. *Nat Rev Neurosci*. 17(1):83–95.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science*. 275(159):3–9.
- Silvetti M, Seurinck R, Verguts T (2013) Value and prediction error estimation account for volatility effects in ACC: a model-based fMRI study. *Cortex*. 49(16):27–35.
- Sutton RS, Barto AG (2018) *Reinforcement Learning: An Introduction*. Second Edition. Cambridge, MA: MIT Press.
- Threlfell S, Lalic T, Platt NJ, Jennings KA, Deisseroth K, Cragg SJ (2012) Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron*. 75:58–64.
- Vaidya AR, Badre D (2020) Neural Systems for Memory-based Value Judgment and Decision-making. *J Cogn Neurosci*. 32(1):896–923.
- Vassena E, Krebs RM, Silvetti M, Fias W, Verguts T (2014) Dissociating contributions of ACC and vmPFC in reward prediction, outcome, and choice. *Neuropsychologia*. 59(1):12–23.
- Wilbertz G, van Kemenade BM, Schmack K, Sterzer P (2017) fMRI-based decoding of reward effects in binocular rivalry. *Neurosci Conscious*. 2017:nix013.
- Wittkuhn L, Schuck NW (2021) Dynamics of fMRI patterns reflect sub-second activation sequences and reveal replay in human visual cortex. *Nat Commun*. 12:1795.
- Yagishita S, Hayashi-Takagi A, Ellis-Davies GCR, Urakubo H, Ishii S, Kasai H (2014) A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science*. 345(16):16–20.
- Yan C, Su L, Wang Y, Xu T, Yin D-Z, Fan M-X, et al. (2016) Multivariate Neural Representations of Value during Reward Anticipation and Consummation in the Human Orbitofrontal Cortex. *Sci Rep*. 6:29079.
- Young AMJ (2004) Increased extracellular dopamine in nucleus accumbens in response to unconditioned and conditioned aversive stimuli: studies using 1 min microdialysis in rats. *J Neurosci Methods*. 138:57–63.

3. Summary

3.1 English version

Surviving in constantly changing environments necessitates the brain's ability to update internal models and predict future events. A fundamental aspect of learning involves minimizing the gap between internal expectations and real-world outcomes, wherein prediction errors (PEs) serve as instructive signals. This study focuses on elucidating how PEs, specifically reward prediction errors (RPEs) and state prediction errors (SPEs), drive dynamic changes in neural patterns and consequently shape adaptive behavior. Multivariate analyses of functional magnetic resonance imaging (fMRI) data offer a unique perspective into neural representations, allowing the investigation of the effects of distinct PEs on the brain's spatially distributed activity patterns. Searchlight-based algorithms were employed to calculate pattern change across the entire brain, which was then related to trial-by-trial BOLD responses of PEs. Pattern change analyses revealed that RPE-evoked BOLD responses in the ventral striatum predicted subsequent pattern changes in orbitofrontal cortex (OFC) and anterior cingulate cortex (ACC). Conversely, SPE-evoked BOLD responses in the intraparietal sulcus predicted pattern changes in superior parietal lobule (SPL) and precuneus (PCu). Notably, changes of neural activation patterns, modulated by RPEs and SPEs, were found to predict subsequent adaptations in participants' behavioral policy. This study provides comprehensive insights into how PEs, specifically RPEs and SPEs, drive short-term changes in neural activation patterns, influencing subsequent behavioral adaptations.

3.2 German version

Das Leben in ständig wechselnden Umgebungen erfordert die Fähigkeit des Gehirns, interne Modelle zu aktualisieren, um zukünftige Ereignisse vorherzusagen. Ein zentraler Aspekt des Lernens besteht darin, die Diskrepanz zwischen internen Erwartungen und realen Ereignissen zu minimieren, wobei Vorhersagefehler (PEs) als Lernsignale dienen. Diese Studie erforscht die Mechanismen, durch welche PEs, insbesondere Belohnungsvorhersagefehler (RPEs) und Zustandsvorhersagefehler (SPEs), dynamische Veränderungen in neuronalen Mustern beeinflussen und folglich adaptives Verhalten ermöglichen. Multivariate Analysen funktioneller Magnetresonanztomographie (fMRT)-Daten bieten eine einzigartige Perspektive auf neuronale Repräsentationen und ermöglichen die Untersuchung der Auswirkungen unterschiedlicher PEs auf räumlich verteilte Aktivitätsmuster des Gehirns. Searchlight-basierte Algorithmen wurden verwendet, um Musteränderungen im gesamten Gehirn zu berechnen, die dann mit BOLD-Antworten von PEs in Beziehung gesetzt werden. Die Analysen der Musteränderungen zeigen, dass RPE-ausgelöste BOLD-Antworten im ventralen Striatum nachfolgende Musteränderungen im orbitofrontalen Cortex (OFC) und im anterioren cingulären Cortex (ACC) vorhersagen. Im Gegensatz dazu können SPE-ausgelöste BOLD-Antworten im intraparietalen Sulcus nachfolgende Musteränderungen im superioren parietalen Lappen (SPL) und im Precuneus (PCu) prognostizieren. Hierbei konnte gezeigt werden, dass Veränderungen der neuronalen Aktivierungsmuster, die durch RPEs und SPEs moduliert werden, nachfolgende Anpassungen in der Verhaltensstrategie der Teilnehmer vorhersagen. Diese Studie liefert umfassende Einblicke wie PEs, insbesondere RPEs und SPEs, kurzfristige Veränderungen in neuronalen Aktivierungsmustern beeinflussen und nachfolgende Verhaltensanpassungen bewirken.

4. Declaration of self-contribution

Leon Möhring and Jan Gläscher designed the experiment and generated initial hypotheses about the experimental paradigm. Leon Möhring and Jan Gläscher also conceptualized the task that participants performed in this project. Leon Möhring created and programmed the task. Leon Möhring instructed participants and acquired the functional magnetic resonance imaging data. Leon Möhring and Jan Gläscher developed analytical protocols including the novel searchlight algorithm used to calculate pattern change maps of the whole brain. Leon Möhring conducted the statistical analyses involving preprocessing of fMRI data, hierarchical behavioral modeling of participant's choices using Reinforcement Learning algorithms, univariate model-informed analyses locating prediction error signals in the brain, multivariate analyses calculating trial-by-trial pattern change maps of the brain, whole brain GLMs to combine univariate prediction error signals and multivariate pattern changes, multivariate classification of neural patterns using support vector machines to decode the information content of observed brain patterns. Leon Möhring and Jan Gläscher wrote both publications. Leon Möhring wrote the dissertation. Leon Möhring created all illustrations and figures of the final publications and the dissertation.

5. Acknowledgements

First and foremost, I would like to express my appreciation to Dr. Jan Gläscher for his exceptional supervision, inspiring support, and invaluable guidance throughout the process of writing these publications and the dissertation. His expertise, insightful feedback, and encouragement have been instrumental in shaping the research and refining its outcomes.

I also like to thank Professor Christian Büchel for his support and guidance throughout the publication process and the supervision of my dissertation.

Furthermore, I am grateful to my colleagues and peers for their collaboration and stimulating discussions and to my family and friends for their support and encouragement.

This dissertation and the accompanying publications would not have been possible without the support and contributions mentioned above.

6. Curriculum vitae / Lebenslauf

Karriere und Studium

Seit 02/2024	Assistenzarzt in der Klinik für Diagnostische und Interventionelle Radiologie und Nuklearmedizin, Universitätsklinik Hamburg-Eppendorf
12/2023	Ärztliche Approbation
04/2020 – 04/2021	Doktorandenstipendium am Institut für Systemische Neurowissenschaften, UKE (DFG SFB - TransRegio 169, Crossmodal Learning: Adaptivity, Prediction and Interaction)
10/2016 – 12/2023	Studium der Humanmedizin am Universitätsklinik Hamburg-Eppendorf
05/2016	Abitur am Gymnasium Vegesack, Bremen

Publikationen und Konferenzen

03/2024	Protocol for predicting multivariate change of brain patterns using model-informed fMRI activations. (STAR Protocols, 2024, Möhring and Gläscher)
08/2023	Prediction errors drive dynamic changes in neural patterns that guide behavior. (Cell Reports, 2023, Möhring and Gläscher)
03/2023	DGKN 2023, Hamburg: (Deutsche Gesellschaft für Klinische Neurophysiologie) Changes in neural patterns modulated by two distinct types of prediction errors. (DOI: https://doi.org/10.1016/j.clinph.2023.02.150)
08/2022	CCN 2022, San Francisco: (Conference on Cognitive Computational Neuroscience) Common and distinct changes in brain activation patterns modulated by two different types of prediction errors. (DOI: https://doi.org/10.32470/CCN.2022.1087-0)

7. Statutory declaration / Eidesstattliche Versicherung

Ich versichere ausdrücklich, dass ich die Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die von mir angegebenen Quellen und Hilfsmittel nicht benutzt und die aus den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen einzeln nach Ausgabe (Auflage und Jahr des Erscheinens), Band und Seite des benutzten Werkes kenntlich gemacht habe.

Ferner versichere ich, dass ich die Dissertation bisher nicht einem Fachvertreter an einer anderen Hochschule zur Überprüfung vorgelegt oder mich anderweitig um Zulassung zur Promotion beworben habe.

Ich erkläre mich einverstanden, dass meine Dissertation vom Dekanat der Medizinischen Fakultät mit einer gängigen Software zur Erkennung von Plagiaten überprüft werden kann.

Unterschrift: