Reduced Reverse Replay in Anxious Individuals Impairs Reward Prediction

Qianqian Yu[1,2,3#], Yuejia Luo[1,2#], Ray Dolan[2,4,5,6], Jianxin Ou[2,3], Chuwen Huang[1], Haiteng Wang[2], Zhibing Xiao[2], Yunzhe Liu[2,3*]

1. School of Psychology, Center for Brain Disorders and Cognitive Science, Shenzhen University, Shenzhen, China

2. State Key Laboratory of Cognitive Neuroscience and Learning, IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China

3. Chinese Institute for Brain Research, Beijing, China

4. Max Planck University College London Centre for Computational Psychiatry and Ageing Research, London, UK

5. Wellcome Centre for Human Neuroimaging, University College London, London, UK

6. Department of Psychiatry, Universitätsmedizin Berlin, Berlin, Germany

# These authors contributed equally

* Correspondence: yunzhe.liu@bnu.edu.cn (Y.L.)

Abstract

Anxiety traits can manifest as a tendency to overlook positive information, impacting memory and decision-making. Using EEG-based neural decoding, we examined future reward prediction in participants with varying levels of trait anxiety. In a resting period, following value learning, we found an increase in reverse replay for a task sequence that led to reward. This reward-related replay was significantly reduced within individuals exhibiting higher trait anxiety. Furthermore, this reduction correlated with a distorted neural representation of the reward-predictive sequence, an effect mirrored behaviorally in an attenuated preference for stimuli within the same sequence. These findings indicate that higher trait anxiety may impede formation of reward-predictive representations by reducing offline replay during rest. More broadly, our demonstration of sequential replay using EEG enables examination of offline learning and memory consolidation in wider populations, particularly those with psychopathology.

Key words: neural replay, anxiety, reward

Introduction

It is rare that an outcome is directly preceded by its cause. When playing chess, the quality of a discrete move may not be apparent until the game has concluded. This ability to infer causal relationships that go beyond immediate experience is vital for daily life. For instance, failing to anticipate the delights of a restaurant meal may result in deciding against going. Anxious individuals, in particular, are known to overlook potential rewards[1–3], engendering a reluctance to engage in beneficial activity[4–7].

To infer future reward beyond immediate experience, requires an internal model of the environment or cognitive map. It is now known that neural replay, a rapid time-compressed sequential reactivation of past experience, plays a significant role in forming such map[8–12]. When learning to predict reward, hippocampal replay events are temporally coordinated with the firing of reward-responsive cells in the dopaminergic midbrain[13]. When reward magnitude increases, reverse (opposite order to actual experience), but not forward (same order with experience) replay, is boosted accordingly[14,15]. Recent evidence points to similar replay mechanisms in humans[16,17]. In humans, instead of a mere recapitulation of past experience, replay represents the internal model of events and organizes experiences for novel inferences[16,18]. Furthermore, reverse replay increases selectively upon reward receipt, facilitating model-based reinforcement learning that extends beyond direct experience[17].

Aberrant replay of experiences is posited as a key mechanism in psychiatric disorder[19,20] for disrupting cognitive map formation[13,21–23]. It is also linked to irrational decisions for approaching risky environment, an effect more pronounced in people with higher trait anxiety[24]. This may be due to an impairment of value propagation necessary for learning in anxiety, leading to an inability to infer future reward and avoidance of potentially rewarding experiences[25–27].

To examine how replay in anxious individuals impacts on reward prediction, we adopted a sequence learning task previously shown to elicit reverse replay after value learning[16]. Here, subjects are required to learn which stimuli predict future reward. The stimuli form two sequences, one of which ends with a reward and the other with a neutral outcome. Preference for all stimuli was rated twice, before and after value learning, allowing us to measure stimulus preference changes as an index of reward effect on subjective valuation[23]. Trait anxiety, a stable personality trait reflecting a predisposition to neglect positive stimuli (and a disposition to focus on negative aspects), was used as an individual difference variable. We hypothesized that individuals with high trait anxiety would exhibit reduced reverse replay of reward sequence following value learning, accompanied by a reduced reward-induced change in stimuli preference, compared to those with low trait anxiety.

Recent methodological advancements have opened a new avenue for examining neural replay in humans non-invasively. This has been most commonly applied in magnetoencephalography (MEG) [16–18,28–30], but also fMRI[31,32]. However, the limited accessibility and portability of MEG has constrained its use, especially for psychiatry research[33–35]. In this respect, electroencephalography (EEG), offers distinct advantages over MEG or fMRI[33–36], and has been shown to be amenable to index neural replay in temporally delayed linear modeling (TDLM)[37]. In the current study, we examine replay deficits in individuals with varying levels of trait anxiety using EEG.

Results

Task. Participants were recruited carefully based on their trait anxiety score on STAI so that they were evenly spanned the spectrum (20 subjects in four score domains, 20-30, 30-40, 40-50, and above 50. The overall mean trait anxiety score is 42.57 ± 1.34). All subjects who then passed a structure learning phase, performed a model-based value learning task on novel stimuli with concurrent EEG recording. The task spanned two

days (Fig. 1a). On Day 1, subjects were shown eight stimuli embedded in two visual sequences [B'/A/D'/B] and [A'/C/C'/D] and were explicitly instructed on how to map these stimuli onto two underlying true sequences: A->B->C->D and A'->B'->C'->D' (Fig. 1a, left panel). Participants who achieved above 80% accuracy were allowed to proceed to the formal EEG experiment on Day 2.

On Day 2, participants were presented with a new set of eight visual stimuli (Fig. 1a, right panel). They were first asked to rate how much they liked each stimulus on a 1-9 scale. Then, subjects performed the task (applied learning) that they were instructed on in Day 1 (the same rule applied to the new stimuli) with concurrent EEG recording (64-channels). The task ended with another preference rating of the eight stimuli. The change of stimulus preference quantifies an effect of reward on subjective valuation[23].

The main task consisted of four phases. In the functional localizer phase, stimuli were presented in random order for training of neural decoders. To ensure subjects attended to the stimuli, on rare occasions (20% probability), stimuli were presented upside down and subjects were required to respond by pressing a button. The mean response accuracy was 97.54 ± 0.21%, and this was independent of subjects' trait anxiety score ($r = 0.02$, $p = 0.895$). Then, in the applied learning phase, subjects applied the Day-1 mapping rule to the Day-2 new stimuli. The mean accuracy is 84.58 ± 1.47%, it increases as a function of applied learning blocks (F (2, 134) = 58.59, $p < 0.0001$, $\eta^2 = 0.47$, run1: 71.13 ± 2.8%, run2: 89.06 ± 1.69%, run3: 93.54 ± 0.76%; Fig. 1b). Following this phase, the last object of each sequence, i.e., D or D' was paired with either a reward or neutral icon (value learning). This was followed by a position probe, in which subjects were probed regarding the precise position of all stimuli on a linear sequence. The overall mean accuracy is 96.98 ± 0.86%, indicating correct learning of the internal model of the task.

In the end, subjects were asked again to provide preference ratings on the stimuli. We found a negative correlation between trait anxiety score and preference increase of the

stimulus paired with reward, so that the more anxious individuals have less increase of preference for stimulus paired with reward (Fig. 1c, $r = 0.26$, $p = 0.036$). Interestingly, beyond paired stimulus, we found the preference for all reward-predictive stimuli increased as a function of its actual closeness to reward, an effect evident only for low anxious individuals (median split, $\beta = 0.31$, $t\ (142) = 2.68$, $p = 0.008$, Fig. 1d), but not for high anxious individuals ($\beta = 0.20$, $t\ (126) = 1.63$, $p = 0.106$), suggesting trait anxiety not only impact reward associative learning, but also reward prediction.

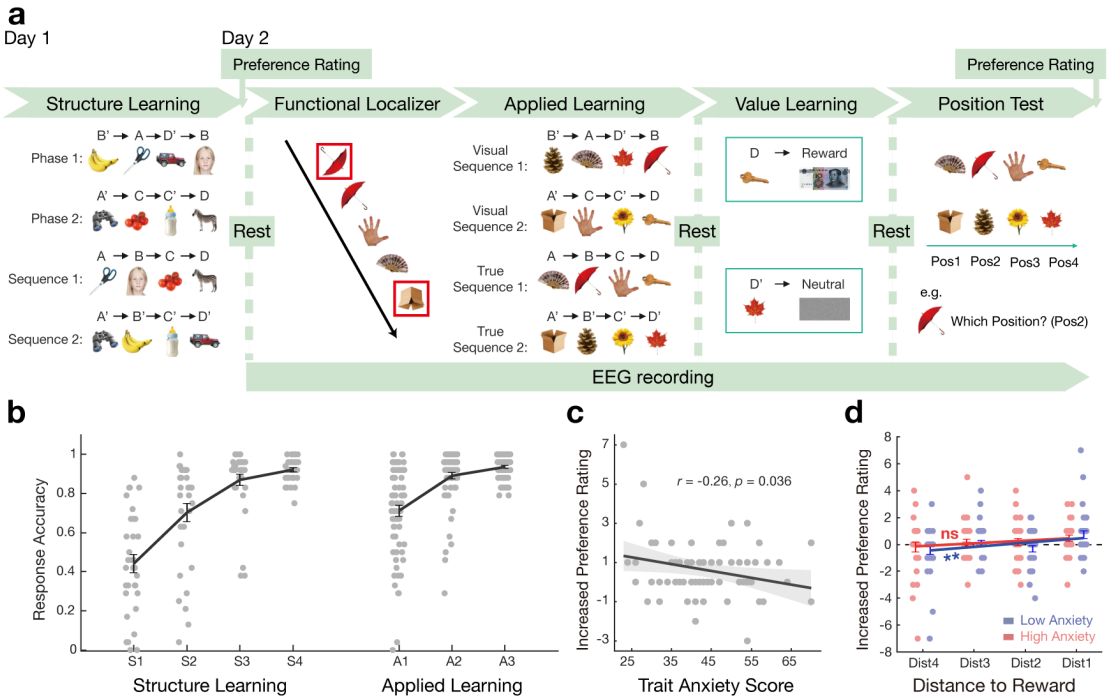


Fig. 1| Task Design and Behavioral Performance of Learning. a, On Day 1, structure learning - participants were trained on the mapping rule (i.e., structure learning) between the sequence of stimuli presentation ("visual sequence", e.g., "B'->A->D'->B") and the task-relevant order ("true sequence", e.g., "A->B->C->D"). On Day 2, participants were exposed to a novel set of stimuli but with the same mapping rule. Stimuli preferences were rated twice, once at the beginning and once at the end of the task. Subjects undertook the sequence learning task[16] while undergoing whole-brain EEG recording. In the functional localizer phase, stimuli were presented in random order for training decoders. In the applied learning phase, stimuli were presented in the same way as Day 1, and subjects were required to apply the learnt structure to reorder the new stimuli. Following applied learning, the last stimulus of each true sequence, *D* or *D'*, was paired with either reward or neutral icon. In the end, subjects were probed on all stimuli as to their position in the true sequence. There are 3 resting states across the whole task used for capturing spontaneous replay of task sequences. During rest, subjects were asked to keep their eyes open for 4 mins, without any task demand. **b,** Mean behavioral performance as a function of learning sessions,

separately for Day 1 structure learning (left), and Day 2 applied learning (right). **c,** The increased preference rating on the object paired with reward is negatively linked to trait anxiety score. **d,** Increases of stimuli preference as a function of distance to reward. There was a significant positive linear relationship between its closeness to reward and increase of preference for low anxious subjects, but not for subjects with high anxiety. Error bars show 95% SEM; each dot indicates results from one subject. The solid line reflects the best robust linear fit, $*p < 0.05$, $**p < 0.01$. Red dots represent data from individuals of high trait anxiety, and blue dots represent data from low anxiety.

Neural decoding. To measure sequential reactivation (or replay) of stimuli representations during rest, it is first necessary to distinguish neural representations associated with each of the eight stimuli. As for our previous MEG work[16–18], eight binary classifiers were trained separately using logistic regression, one for each stimulus. During functional localizer phase, participants viewed the images that were later used in the task in a random order, ensuring no task-related information was included in the classifier training.

Decoding accuracy was evaluated at each 10 ms following stimulus onset up to 750 ms post-stimulus, using a 10-fold cross-validation scheme[38]. The highest decoding accuracy was observed at 180 ms after stimulus onset (23.21 ± 0.90%, t (67) = 11.934, $p < 0.0001$, two-tailed one-sample t-test; chance: 12.50%, Fig. 2a). The predicted probability of trained models exceeded the significance threshold only when the testing state was the same as what the model was trained to detect, indicating high specificity in detecting stimulus-specific representation (Extended Data Fig. 1). We used stimulus classifiers trained at 180 ms post-stimulus onset for all subsequent analyses.

Selective increase of reverse replay after value learning. The paradigm we used has previously been shown to induce spontaneous neural replay for structurally inferred but not experienced visual sequences[16,18]. Employing a similar task design in the current EEG setting, we also included a value learning phase to test reward-specific changes on replay and behavior (preference rating), as well as the influence of trait anxiety. We quantified the evidence of EEG replay under the TDLM framework[37], as in previous

MEG studies[16–18,28–30,37,39,40]. Specifically, we first applied the trained classifiers to resting state and obtained a reactivation probability for each stimulus at each time bin. Next, we measured the degree to which these reactivation probabilities systematically followed a transition of interest, using a measure called "sequenceness". This measure defined the extent to which sequential reactivation of objects adhere to a hypothesized trajectory (e.g., inferred, or visual sequence). To formally test for learning-induced spontaneous replay, we examined off-line activity. Note there were three resting states in the main task, the first at the beginning of the task, the other two before and after value learning.

We found significant reverse replay of the reward sequence (inferred) following, but not before, value learning, evident across 50-80 ms time lags, peaking at 60 ms lag (β = 0.02 ± 0.01, one-sample t-test, t (67) = 3.20, p = 0.001, significance threshold derived from a non-parametric permutation test, family-wise error (FWE) corrected across all lags, Fig. 2b, also Extended Data Fig. 2). Across all subjects, an increase of reverse replay (after vs. before value learning) was significantly higher for the reward sequence compared to neutral sequence (β_{Reward}= 0.02 ± 0.01, $\beta_{Neutral}$= -0.00 ± 0.01, t (66) = 2.53, p = 0.007, Fig. 3a). We found no evidence for replay of the visual sequence, either before or after value learning (Extended Data Fig. 3). These results are consistent with previous MEG findings[16,18] (see Extended Data Fig. 4-5 for further validation).

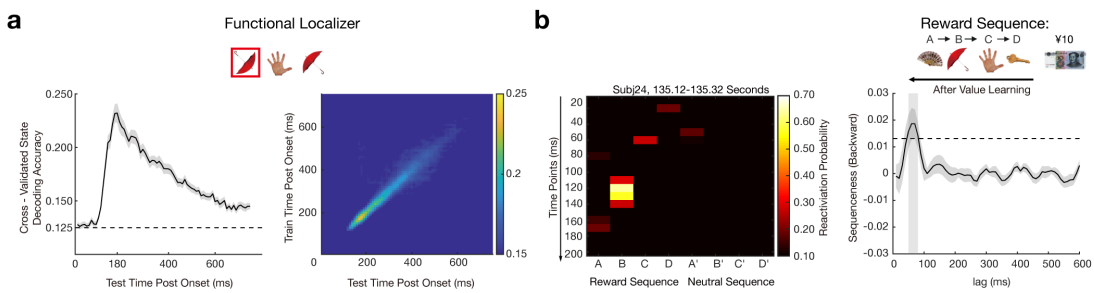


Fig. 2| Neural Decoding and Reverse Replay After Value Learning. a, Cross-validated decoding accuracy of 8 stimuli based on whole-brain EEG signal in the functional localizer phase. As with Liu et al., 2019, a binary classifier was trained independently for each stimulus using logistic regression. On the left, the decoding time course of classifiers trained and tested at the same time. The horizontal dashed

line represents the peak-level P_{FWE} = 0.05 threshold for chance-level accuracy (test-data label permutation). On the right, it is the decoding accuracy of classifiers trained on one time point (row) and tested on all time points (column) from stimuli onset to post 750 ms. Decoding performance for each state was shown in Extended Data Fig. 1. **b,** An example of replay events from one subject is shown for visualization purpose (left panel). Each row depicts reactivation probabilities of all stimuli at a given time point. For statistical purpose, data were analyzed using a ''sequenceness'' measure (see Method Details for details). Reverse replay for a reward sequence, after value learning (right panel). The dotted line is the FWE-corrected significance threshold. The grey shaded area indicates significant time lags of replay, with the peak being at 60 ms lag. This result is similar with previous studies measuring replay in MEG[16,18]. Also, as with previous MEG studies[17,39], replay strength was quantified using TDLM[37]. Other replays findings (true or visual sequences, before or after applied learning) are shown in Extended Data Fig. 2-3.

High trait anxiety is associated with decreased reverse replay.

A median split analysis of trait anxiety revealed that individuals with low anxiety exhibited a greater increase of reverse replay than those with high anxiety (β_{low}= 0.03 ± 0.01, β_{high}= 0.00 ± 0.01, t (66) = -2.74, p = 0.008, Fig. 3b). In addition, a negative correlation was evident across subjects between change of reverse replay and trait anxiety (r = -0.27, p = 0.029, Fig. 3c). Such change of reverse replay was also positively related to change of stimuli preference from pre- to post- task, specifically for stimuli within reward sequence (r = 0.25, p = 0.042, Fig. 3c).

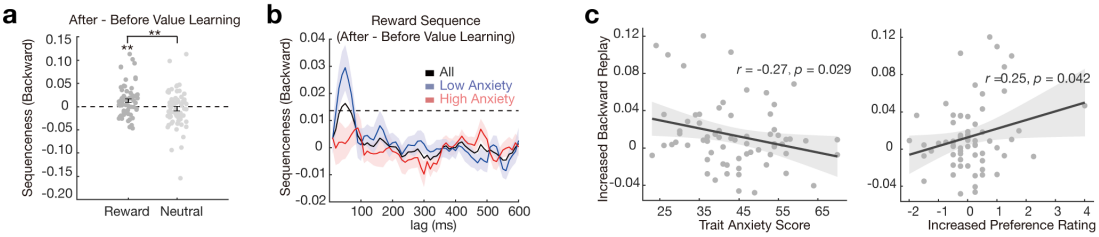


Fig. 3| Diminished Reward Effect on Reverse Replay is Associated with Trait Anxiety. a, The increased reverse replay (after – before value learning) was significantly higher for reward vs neutral sequence. **b,** A selective increase of reverse replay after value learning is more pronounced in low vs. high anxiety subjects at the 50 ms lag peak. The dashed line is the permutation threshold after controlling for multiple comparisons. **c,** The reward effect on replay is linked to behavioural change - there is a positive correlation between the increase of reverse replay and an increase in preference rating, specifically for reward sequence (left panel). Such reward effect on replay was negatively correlated with

subjects' trait anxiety (right panel). Each dot is a subject, the solid line represents the best linear fit with shaded area indicating 95% SEM.

Reward-based representation is positively related to reverse replay, negatively linked to trait anxiety. Finally, we asked whether diminished reward learning-induced replay in subjects with high trait anxiety related to a distorted neural representation of task structure. To address this, we employed a representational similarity analysis (RSA), as in previous studies[18,41], quantifying the emergence of reward representation following learning. We used localizer tasks before and after value learning (Functional Localizer and Position Test), which generated a [state x state] 'similarity change' matrix for each time point in the evoked neural response following stimuli onset. We then regressed these matrices onto a task design matrix that comprised predictors encoding a correct abstracted 'position' representation and two predictors capturing 'sequence' representations (reward and neutral sequences, Fig. 4a, further details were provided in the Method section). This approach allowed us to quantify learning-induced changes in task-specific neural representation.

We found evidence for an emergence of reward representation following learning, but only for subjects with low trait anxiety (not in high anxiety group, median split), from 480-630 ms after stimulus onset (cluster-level $P_{FWE} = 0.02$, non-parametric permutation test; peak 540 ms: $\beta = 0.05 \pm 0.02$, $t(35) = 3.37$, $p = 0.0009$, Fig. 4b, see also Extended Data Fig. 6). Such reward-based clustering positively correlated with the strength of reward-induced reverse replay ($r = 0.26$, $p = 0.035$, Fig. 4c left panel), as well as an increase in preference rating for stimuli belonging to the reward sequence ($r = 0.26$, $p = 0.034$, Fig. 4c right panel).

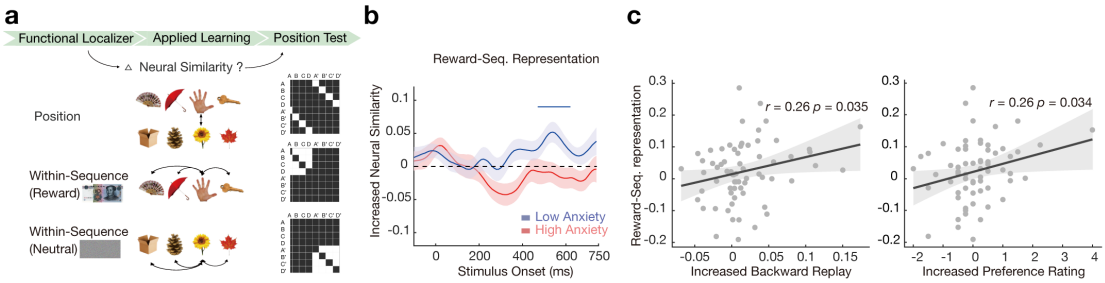


Fig. 4| Neural Representation of Reward Sequence and Trait Anxiety. a, Hypothesized patterns of learning-induced ''similarity change'' and corresponding state x state design matrices. **b,** Learning-induced change in neural representation of a reward sequence is seen only in low, but not high, anxiety subjects (median split). The horizontal blue line indicates significant cluster $P_{FWE} < 0.05$ time-window. **c,** Neural representation of reward sequence is positively linked to an increase in reverse replay, as well as preference rating, specifically for reward sequence. Each dot is a subject, the solid line represents the best linear fit with shaded area indicating 95% SEM.

Discussion

Utilizing EEG-based neural decoding, we addressed the neural mechanisms underlying reward prediction and how this is impacted by anxiety. Consistent with previous MEG-based replay findings[16,17], we show a selective increase of reverse replay for reward sequence following value learning, as captured by EEG. Intriguingly, individuals with higher trait anxiety exhibited reduced reward replay during rest compared to low anxious ones. This reduction in replay was associated with a distorted neural representation of the reward sequence during task performance, mirrored behaviorally in a diminished preference changes for reward-predictive stimuli. These findings are consistent with the idea that anxious individuals have difficulty learning from positive experiences[26] and suggest that this may be related to an aberrant replay mechanism.

Anxious individuals are known to manifest a bias towards negative information and away from positive information[26,42]. This has been extensively studied during task performance[43–46], and is often associated with memory impairments[47–50]. In the context of positive experiences[51], such memory deficits were often linked to encoding or retrieval process[52–54]. The current study suggests there may also be a contribution from offline consolidation of experience during post-task rest, when neural replay is thought

to strengthen learnt associations to reward following task structure. Across all subjects, we detected a significant reverse replay for reward sequence during rest following value learning, while no replay was found for visual sequence either before or after learning. Interestingly, there was significant reduced reward replay in high anxious individuals, accompanied by an absence of an expected preference change for stimuli predictive of reward (see also Extended Data Fig. 7). This aligns with previous resting-state studies showing disrupted functional connectivity within reward-related brain networks (including amygdala, hippocampus, and prefrontal cortex) in anxiety disorders[55–58]. Such disruptions may contribute to difficulties in consolidating reward-related cognitive map during rest, leading to deficits in subjective valuation, and that may ultimately compromise effective decision-making in those with high anxiety[1,44].

While our study has provided insights into the role of anxiety in reward prediction and the offline consolidation of positive experiences, it's imperative to note the potential limitations due to the inherent comorbidity between anxiety and depression[59,60]. Anxiety and depression often co-occur, and this can confound the understanding of neural mechanisms specific to each condition[49,61]. Additionally, our study focused primarily on the reward learning domain, while the aversive domain remained unexplored. Anxiety disorders are frequently associated with heightened anticipation of aversive outcomes[62–64]. Therefore, the mechanisms of learning from negative or aversive experiences, which might be equally, if not more, affected in anxious individuals, warrant investigation. It will be of interest in future studies to investigate deficits in anticipating dangerous or aversive outcome, which may also be due to aberrant replay mechanism[24,25].

Recent works on decoding human cognition during rest show great potential in elucidating novel mechanisms underlying learning, memory and decision-making[11], and shed new light on the psychiatric disorder[19]. This has driven methodological advances, in particular, replay analysis in non-invasive neuroimaging[16,31,37]. While previous MEG studies capitalize on its fine spatial-temporal resolution, capturing fast

sequential reactivations or replay during rest[17,18,24,25,29]. It is rather expensive and immobile (but see recent progress on optically pumped magnetometers[33–35], which offers a mobile and cheap MEG solution), limiting its use for studies in clinical settings. EEG, on the other hand, offers the same fine temporal resolution with MEG, but is much more widely available and easily accessible within clinical setting, making it an ideal candidate for neural decoding or replay studies in psychiatry research. The current study show that whole-brain 64-channel EEG is capable of decoding up to eight stimuli (Extended Data Fig. 1), and certain sensitivity to detect their sequential neural replay during rest (Extended Data Fig. 5). Consistent with previous MEG studies[16,65], we show a selective increase in reverse replay for the reward sequence following value learning, with no replay for visual sequences. We acknowledge that forward replay prior to value learning was not detected in the current EEG study, an absence that may be attributed to a lower decoding accuracy observed in the 64-channel EEG signal utilized in this study (mean accuracy 23.21 ± 0.90%), in comparison to MEG signal based upon 275 independent sensors (mean accuracy 39.60 ± 2.20%)[16]. Further research is required to develop novel methods that can improve decoding accuracy and replay sensitivity using EEG, particularly when employing high-density EEG configurations[66–69].

Understanding the neural mechanisms underlying anxiety-related alterations in inferring future reward is not only important from a basic science perspective but also has clinical implications. Cognitive-behavioral therapies have shown promise in reducing anxiety symptoms and improving cognitive functioning[70,71]. Attention bias modification training and memory reconsolidation interventions, in particular, demonstrate efficacy in addressing anxiety-related cognitive biases and memory deficits[72]. Future research could explore interventions targeting replay mechanisms to enhance positive memory consolidation and enable a more veridical cognitive map in anxious individuals. By actively updating and correcting the cognitive maps[73], individuals can develop more accurate and adaptive perceptions of the world, allowing for adaptive behavior in changing environment.

In conclusion, we show that anxious individuals exhibit difficulties in inferring future reward, evidenced by reduced reverse replay and a distorted neural representation of a reward sequence. This anxiety-related neural alteration is also evident in a blunting of preference increase for stimuli predictive of reward. These findings provide new insights into understanding of anxiety and highlighting an altered offline memory consolidation mechanism. Further research can target memory consolidation mechanisms of those with psychopathology and help guide interventions aimed at modulating memory and emotion during rest or sleep.

Methods

Participants. Eighty healthy volunteers (46 males; mean age = 20.90 ± 1.95) participated in the experiment. All participants had a normal or corrected-to-normal vision and no history of neurological or psychiatric illness. The study was approved by the local Ethical Committee of Shenzhen University (PN-202300013). The participants provided written informed consent in compliance with the Declaration of Helsinki and were compensated approximately $30 for their participation. The participants received comprehensive training of task rule (structure learning) one day prior to the EEG task and only subjects achieving at least 80% accuracy were permitted to the EEG experiment the next day. Seven participants were excluded from the analysis because of failure of passing the Day-1 structure learning, and another five subjects were excluded due to excessive movements and/or recording artifacts during EEG recording. Therefore, 68 subjects (mean trait anxiety 43.93 ± 1.33) in total were included for subsequent analysis.

Design and Materials. The task spanned two days in a manner that followed closely our previous MEG study design[16], with the exception of eliciting stimuli preference ratings before and after the main task, in order to capture a putative reward effect on behavior.

On Day-1, participants were shown eight sample stimuli embedded in two visual sequences [B'/A/D'/B] and [A'/C/C'/D] and were explicitly instructed on how to map the stimuli onto two true sequences: A->B->C->D and A'->B'->C'->D'. This mapping between visual and underlying true order was randomized across subjects. Subjects completed four runs of training where each run consisted of one learning and one probe session. During learning, two visual sequences ([B'/A/D'/B] and [A'/C/C'/D]) were presented three times, with each stimulus presented serially in the center of the screen for 900 ms, followed by a 500 ms inter-stimulus interval. Each sequence presentation ends with a 2000 ms interval of blank screen. After that, participants were quizzed about the true order of the stimuli without feedback. On each trial, the target image appeared for 5000 ms, during which subjects were required to think about which image should appear following (but not necessarily next to) the target image. Then, two candidate images appear on the screen and subjects were asked to select the correct one within 1000 ms, a limit set to minimize potential associative learning when the stimuli were presented together[16]. Participants were required to reach at least 80% accuracy in the last structure learning run to be permitted for the Day-2 experiment.

On Day 2, participants were presented with a new set of eight stimuli, and first rated their preferences for each of the eight images on a 1-9 scale (from "strongly dislike" to "strongly like"). The procedure was performed twice, one before and one after the main task, where the aim was to assess preference changes due to value learning[22,23], an implicit reward effect on behavior. Subjects then performed the main task during concurrent whole-brain EEG recording. The main task comprises four phases: functional localizer, applied learning, value learning, and a position test, and three resting states, one immediately before the functional localizer task, one before and one after value learning in order to capture reward induced change in spontaneous neural replay.

In the functional localizer phase, stimuli were presented in a random order. As this preceded applied learning, participants had no structure information about stimuli, such

that decoding models trained during functional localizer solely captured sensory level stimulus information. On each trial, this phase starts with a text presented for 1500 - 3000 ms during which subjects were required to vividly imagine its associated imagery content, followed then with the actual image for 750 ms, and a jitter ITI for 700 - 1700 ms. To ensure subjects were attending to the stimuli, on rare occasions (20% probability), random stimuli were presented upside down and subjects were required to respond to these by pressing a button. There were 400 trials in total, split in two runs, with 50 instances for each visual stimulus (20% were upside down). Only correct -side-up images (40 repetitions per image) were used for classifier training.

In the applied learning phase, subjects were required to apply the Day-1 learnt mapping rule to the Day-2 new stimuli. The use of an entirely new set of stimuli ensured neural replay signatures in EEG (if exist) could not be attributed to perceptual biases introduced during Day-1 training. Participants were explicitly informed that the mapping rule was the same with the one on Day-1. There were three runs of Applied Learning. As for structure learning, each run consisted of one learning and one probe session. During learning, two visual sequences ([B'/A/D'/B] and [A'/C/C'/D]) were presented three times, with each stimulus presented serially in the center of the screen for 900 ms, followed by a 500 ms inter-stimulus interval. Each sequence presentation ends with a 2000 ms interval of blank screen. After that, participants were quizzed about the true order of the stimuli without feedback. On each trial, the target image appeared for 5000 ms, during which subjects were required to think about which image should appear following (but not necessarily next to) the target image. Then, two candidate images were presented on the screen and subjects were tasked to select the correct one within 1000 ms. This time limit was designed to minimize potential associative learning when the stimuli were presented together. No feedback was provided. There was a 33% possibility that the wrong answer came from the same sequence but preceded instead of following the probe stimuli. This setup was designed to encourage participants to form sequential rather than clustering representations (i.e., which sequence does this object belong to).

Next, participants completed value learning, where the last object of each sequence (i.e., D, D') was paired with either reward or a neutral icon. On each trial, participants saw the object at each end of the sequence (i.e., D or D') for 900 ms, followed by an inter-stimulus interval of 3000 ms, and then either received a reward (image of 10 Chinese Yuan banknote) or no-reward (scrambled image of 10 Chinese Yuan banknote) outcome for 2000 ms, followed by an ITI of 3000 ms. There were 12 repetitions for each association, with trial order randomized. Before or after the value learning, participants underwent a 4-minute resting session.

Participants subsequently completed a position test, wherein an object was presented on screen for 1000 ms, and subjects required to think about its position within its sequence. Each object was followed by a single number (1, 2, 3, or 4), and participants indicated whether the presented number matched the sequence position of the preceding object, via a button press (1000 ms response window, chance accuracy 50%). This was followed by a 3000 ms ITI. There were 80 trials in total, 10 repetitions for each stimulus, with a constraint that the same stimulus does not appear consecutively. No feedback was provided. Finally, at the end of the main task, participants were required to write down the true sequences in the correct order.

EEG Data Acquisition

Data Acquisition and Preprocessing. In this study, a Brain Products 64 channel electroencephalography (EEG) system was used to record EEG during Day-2 experiment, with an additional electrode placed below the right eye as vertical electro-oculograms (EOG). Online EEG recordings were sampled at a rate of 1000 Hz with FCz as reference electrode during data acquisition. All electrode impedances were kept below 5 $K\Omega$.

Preprocessing was conducted separately for each session, following the approach used by previous MEG study[16]. The data were first high-pass filtered at 0.5 Hz and notch filter of 50 Hz to remove line noise, and then down-sampled to 100 Hz. The continuous recording was visually screened for noisy channels, which were interpolated using the weighted average of surrounding electrodes. All electrodes were re-referenced offline to the average. The data was subsequently segmented into epochs, with each epoch spanning from -100 ms before to 750 ms after the image onset. Trials with extremely high noises by visual inspection were manually excluded. Bad channels with abnormal voltage were interpolated with the weighted average of their neighboring electrodes. Independent component analysis (ICA) was performed to remove eye movement and other artifactual components and the remaining components were then back projected to the EEG channel space. Five participants were excluded for excessive head movement or other artifacts. Only trials with correct button press were used for subsequent analyses.

EEG Analysis

Neural Decoding Analysis. Same with Liu et al. [16], a binary classifier was trained independently for each stimulus using logistic regression. We trained each of the eight binary classifiers ($k \in \{1:8\}$) on the evoked whole-brain neural response at a signal time bin after the stimulus onset in the Functional Localizer task. Essentially, model (k) distinguished sensor patterns based on the particular stimulus k (positive examples) relative to all other stimuli ($\sim k$), which was further aided by an equal amount of 'null' data from the inter-trial interval that acted as negative examples. The inclusion of null data decreased the spatial correlation between the decoders and allowed simultaneous reporting of low probabilities in the rest data by all decoders[16,17]. To enhance sensitivity for sequence detection, we used L1 regularization and fixed lambda = 0.001, to prevent overfitting of the results to the regularization parameter. Every model k, unique to time and stimulus, was represented as a single vector of length $n + 1$ (n sensors [max. 62], plus intercept).

To quantify within-session decoding accuracy, we trained regression models using 10-fold cross-validation on the Functional Localizer data. During each validation loop, prediction accuracy was quantified as the proportion of test trials, where the decoder reported the highest probability corresponding to the trial label, giving a prediction accuracy. The average of all estimates obtained over the validation loops determined the overall accuracy, consistent with the approach in previous MEG studies[18,22,39].

The decoders trained and tested at 180 ms after the stimulus onset showed the highest decoding accuracy (mean accuracy over all participants), as seen in Fig. 2a and Extended Data Fig. 1. In other words, decoders from this time bin were utilized for sequenceness analysis[16,28,74]. We confirmed that decoding accuracy was significantly greater than chance (expected chance decoding accuracy 1/8 = 12.5%) using nonparametric tests. Notably, the association between the eight visual stimuli and the corresponding state indices was fixed within subjects but randomized across subjects. This randomization prevented any systematic dissimilarity among stimuli (e.g., stimulus preference or stimulus decodability), regardless of the consistency in subjects, from having an influence on state decoding differences at the group level. Participants were not provided with information regarding the association or possibility of the stimuli and their corresponding states during the functional localizer task. Furthermore, the stimuli were delivered in a random order with the same frequency. Consequently, the classifiers lacked any bias from the task structure. All classifiers were later used to examine for sequences in resting task.

Neural Sequence Analysis. To evaluate neural sequences, we measured the sequence strength of a pairwise state-to-state transition using a multiple regression model, where the representation of state i statistically predicts the subsequent representation of state j at a specific time lag (i.e., speed of replay). This measure is an average estimate of statistical predictiveness, incorporating the number and strength of replay events, and is referred to as "sequence strength." The reason for this approach is that neural representations (of different states) are efficiently decoded in a noisy and probabilistic

manner. The detailed methodology, including related simulations, is documented in Liu et al. (2019)[16]. We have implemented the Temporal Delayed Linear Modelling (TDLM) methodology in EEG data following its success in earlier MEG empirical work. The process began by utilizing eight-state classifiers (from the highest accuracy time bin) to characterize EEG data from each timepoint of all the resting sessions, thus generating a [time x state] reactivation probability matrix for each session. We then utilized TDLM to estimate the evidence for sequential reactivations, consistent with the task-defined sequential order.

TDLM is a method that employs multiple linear regression to quantify the degree to which a lagged reactivation time course of state i, ($X(\Delta t)_i$), where t indicates lag time) can predict the reactivation time-course of state j, (X_j). This approach comprises two stages. Initially, we conducted separate family multiple regressions (first-stage) with the reactivation time course of each state ($j \in \{1{:}8\}$) as the dependent variable and the historical (i.e., time-lagged) reactivation time courses of all states as predictor variables:

$$X_j = \sum_{i=1}^{8} X(\Delta t)_i * \beta(\Delta t)_{i,j} + C \tag{1}$$

The predictor (design) matrix from a single model contained a separate predictor for the reactivation time courses of all states ($i \in \{1{:}8\}$), lagged by $\Delta t \in$ {10 ms, 20 ms, …, 90 ms}, plus the reactivation time course of all states lagged by $\Delta t + \alpha$, where α = [100 ms, 200 ms, …], to capture autocorrelations in state time courses at a canonical alpha frequency, which predominated in human brain activity at rest, in addition to a constant term, C. We repeated the regression in Equation 1 for each $j \in \{1{:}8\}$ and $\Delta t \in$ {10, 20, 30, …, 90 ms}, and used ordinary least squares regression to obtain β. The regression coefficients from Equation 1 quantify the evidence for each empirical state-to-state reactivation pattern at a specific lag. Δt. For example, $\beta(\Delta t)_{i,j}$ quantifies the coefficient that captures the unique variance in X_j explained by $X(\Delta t)_i$. These coefficients were demonstrated in a lag-specific [8 x 8] empirical transition

matrix $\boldsymbol{B}$, representing evidence for all possible state-to-state transitions at a given time lag.

In a second-level regression, we quantified the evidence that the empirical transition matrix, $\boldsymbol{B}$, can be predicted by the underlying task transition structure (i.e., true sequences),

$$B = \sum_{r=1}^{4} Z_r * T_r \tag{2}$$

Here, $\boldsymbol{T}_r$ is the [state x state] predictor transition matrix (for regressor r, with 1 for transitions of interest, and 0 otherwise), and $\boldsymbol{Z}_r$ is the scalar regression coefficient quantifying the evidence that the hypothesized transitions, $\boldsymbol{T}_r$ predict the empirical transitions, $\boldsymbol{B}$. Four predictor matrices were considered, including $\boldsymbol{T}_{auto}$, $\boldsymbol{T}_{const}$, $\boldsymbol{T}_F$, $\boldsymbol{T}_B$, (1) $\boldsymbol{T}_F$: true sequence transitions in the *forward* direction (A->B->C->D and A' ->B' ->C' ->D' entries in $\boldsymbol{T}_F = 1$, all others set to 0), (2) $\boldsymbol{T}_B$, true sequence transitions in the *backward* direction (D->C->B->A and D' ->C' ->B' ->A', i.e. $\boldsymbol{T}_B$ is the transpose of $\boldsymbol{T}_F$), (3) $\boldsymbol{T}_{auto}$: self-transitions ([8 x 8] identity matrix), to control for auto-correlation, and (4) $\boldsymbol{T}_{const}$: a constant matrix, to capture the average of all transitions, ensuring that any weight on $\boldsymbol{T}_F$ and $\boldsymbol{T}_B$ was not due to general dynamics in background neural dynamics.

$\boldsymbol{Z}$ represents the weights of the second-level regression, which is a vector with dimension of r by 1. Each entry in $\boldsymbol{Z}$ reflects evidence for the hypothesized sequences in the empirical transitions, i.e., sequence strength. Note that this estimate of sequence strength is a relative quantity, and an estimate of zero for state i to state j does not mean there is no replay of i -> j, but rather suggests that there is no stronger replay of i -> j than that of other transitions.

The regression in Equation 2 was repeated for each time lag (Δt = 10, 20, 30, ..., 600 ms), resulting in time courses of both forward and backward sequence strength as a

function of time lag. Shorter lags indicate greater time compression, corresponding to faster speed. For statistical testing, we used non-parametric permutation tests at the second-level regression, shuffling the rows and columns of $\boldsymbol{T}_F$ (*forward* predictor matrix), defining $\boldsymbol{T}_B$ (*backward* predictor matrix) as its transpose. For each of 100 permutations we calculated the peak absolute mean sequence strength over participants and across lags (controlling for multiple comparisons across lags). Sequence strength in the unpermuted data was deemed significant at peak-level $P_{\mathrm{FWE}} < 0.05$ if its absolute magnitude exceeded 95% of within-permutation peak.

Representational Similarity Analysis (RSA). To examine how the representation of task structure change from the Functional Localizer (pre-learning) to Position Probe (post-learning) sessions, we applied Representational Similarity Analysis (RSA) on EEG data. We first z-scored the pre-processed EEG data across all trials for each sensor and time point (t, -100 to +750 ms) post stimulus-onset. Next, we regressed the [trial x 1] neural data for each trial, $Y(s)_t$ (at time point t and sensor s), onto a session design matrix, **X**, which included dummy coding for the trial stimulus label, using Equation 3:

$$Y(s)_t = X * \beta(s)_t \tag{3}$$

Here, $\beta(s)_t$ was the [stimulus x 1] vector of regression weights, as an estimate of the stimulus-specific activation for sensor s at time point t. The procedure was repeated over all sensors to yield a [sensor x stimulus] matrix at each time point. Pearson correlation distance was computed between the sensor patterns for each pair of pictures (columns), which resulted in a symmetrical [8 x 8] Representational Dissimilarity Matrix (RDM) at each time point. This procedure was repeated in Functional Localizer (FL) and Position Probe (PP) sessions to determine the learning-induced increase in representational similarity $\Delta\boldsymbol{S}_t$ at each time point, as expressed in Equation 4:

$$\Delta\boldsymbol{S}_t = RDM(FL)_t - RDM(PP)_t \tag{4}$$

Here, s_{ij} of $\Delta \boldsymbol{S}_t$ quantified the post-learning similarity increase between evoked signals for stimuli i and j at time t.

To assess the unique contribution of three predictors (an abstracted representation of ordinal position, reward within-sequence, and neutral within-sequence representation) in explaining the variance in $\Delta \boldsymbol{S}_t$, we employed a second multiple regressions approach. The data were preprocessed by smoothing $\Delta \boldsymbol{S}_t$ over time using a 90 ms Gaussian kernel before conducting regression analysis. To identify time windows (clusters) that showed significant positive evidence for each predictor, we utilized nonparametric tests while correcting for multiple comparisons over time. Specifically, we performed one-sample t-tests for each time point over all participants for each predictor and computed the sum of t-values within each continuous stretch of time points exhibiting a positive effect at $P < 0.05$.

To validate of our results, we repeated this procedure 1000 times, shuffling the rows and columns (stimulus labels) of $\Delta \boldsymbol{S}_t$ consistently across time to preserve temporal smoothness prior to the second regression (permutations). Finally, we extracted the maximal sum-of-t value for each predictor and identified a suprathreshold cluster in the unpermuted data as significant if its sum-of-t exceeded 95% of maximal within-permutation sum-of-t values. The results are displayed in Fig. 4b.

Acknowledgment

Conceptualization, Y.L., Q.Y., R.D. and Y.Luo.; Investigation, Q.Y., Y.L., C.H., J.O., Z.L., Z.X.; Writing – Original Draft, Q.Y., Y.L.; Writing – Review & Editing, Q.Y., Y.L., and R.D. This study is supported by the National Science and Technology Innovation 2030 Major Program (2022ZD0205500), the National Natural Science Foundation of China (32271093, 31920103009), the Major Project of National Social Science Foundation (20&ZD153) and Shenzhen-Hong Kong Institute of Brain Science – Shenzhen Fundamental Research Institutions (2022SHIBS0003).

Conflict of interest

The authors have indicated they have no potential conflicts of interest to disclose.

Data & code availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The analysis code will be publicly available on https://gitlab.com/liu_lab/EEGReplay.git upon publication.

References:

1. Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X. & Bishop, S. J. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat. Neurosci.* **18**, 590–596 (2015).
2. Indovina, I., Robbins, T. W., Núñez-Elizalde, A. O., Dunn, B. D. & Bishop, S. J. Fear-Conditioning Mechanisms Associated with Trait Vulnerability to Anxiety in Humans. *Neuron* **69**, 563–571 (2011).
3. Levy, I. & Schiller, D. Neural Computations of Threat. *Trends Cogn. Sci.* **25**, 151–171 (2021).
4. Dunsmoor, J. E. & Paz, R. Fear Generalization and Anxiety: Behavioral and Neural Mechanisms. *Biol. Psychiatry* **78**, 336–343 (2015).
5. Grupe, D. W. & Nitschke, J. B. Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective. *Nat. Rev. Neurosci.* **14**, 488–501 (2013).
6. Paulus, M. P. & Yu, A. J. Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends Cogn. Sci.* **16**, 476–483 (2012).
7. Sailer, U. *et al.* Altered reward processing in the nucleus accumbens and mesial prefrontal cortex of patients with posttraumatic stress disorder. *Neuropsychologia* **46**, 2836–2844 (2008).
8. Buzsáki, G. Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus* **25**, 1073–1188 (2015).
9. Carr, M. F., Jadhav, S. P. & Frank, L. M. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nat. Neurosci.* **14**, 147–153 (2011).
10. Dupret, D., O'Neill, J., Pleydell-Bouverie, B. & Csicsvari, J. The reorganization and reactivation of hippocampal maps predict spatial memory performance. *Nat. Neurosci.* **13**, 995–1002 (2010).
11. Liu, Y., Nour, M. M., Schuck, N. W., Behrens, T. E. J. & Dolan, R. J. Decoding cognition from spontaneous neural activity. *Nat. Rev. Neurosci.* **23**, 204–214 (2022).

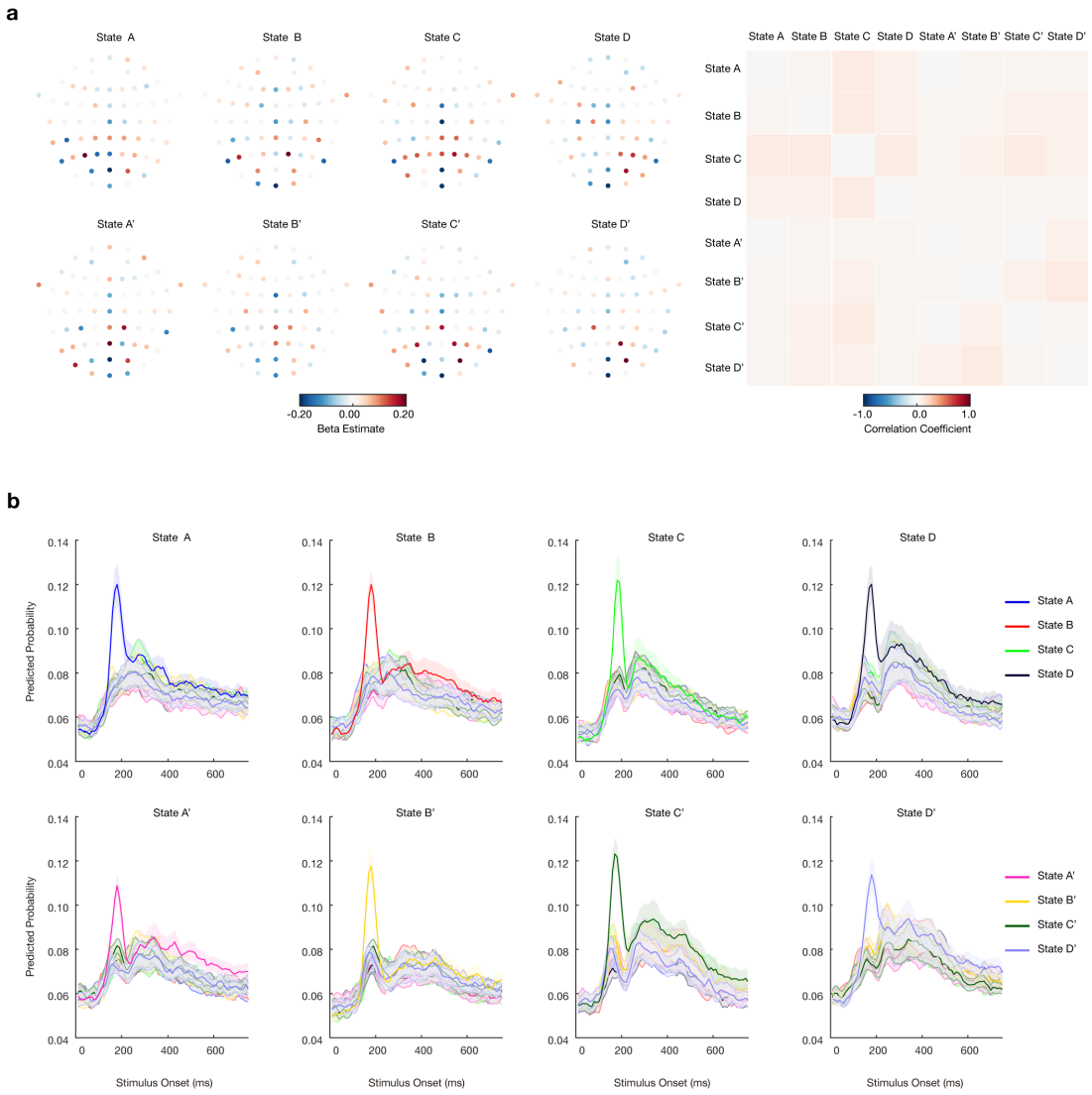
12. Pfeiffer, B. E. & Foster, D. J. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* **497**, 74–79 (2013).
13. Gomperts, S. N., Kloosterman, F. & Wilson, M. A. VTA neurons coordinate with the hippocampal reactivation of spatial experience. *eLife* **4**, e05360 (2015).
14. Ambrose, R. E., Pfeiffer, B. E. & Foster, D. J. Reverse Replay of Hippocampal Place Cells Is Uniquely Modulated by Changing Reward. *Neuron* **91**, 1124–1136 (2016).
15. Foster, D. J. & Wilson, M. A. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* **440**, 680–683 (2006).
16. Liu, Y., Dolan, R. J., Kurth-Nelson, Z. & Behrens, T. E. J. Human Replay Spontaneously Reorganizes Experience. *Cell* **178**, 640-652.e14 (2019).
17. Liu, Y., Mattar, M. G., Behrens, T. E. J., Daw, N. D. & Dolan, R. J. Experience replay is associated with efficient nonlocal learning. *Science* **372**, eabf1357 (2021).
18. Nour, M. M., Liu, Y., Arumuham, A., Kurth-Nelson, Z. & Dolan, R. J. Impaired neural replay of inferred relationships in schizophrenia. *Cell* **184**, 4315-4328.e17 (2021).
19. Heller, A. S. & Bagot, R. C. Is Hippocampal Replay a Mechanism for Anxiety and Depression? *JAMA Psychiatry* **77**, 431–432 (2020).
20. Schafer, M. & Schiller, D. The Hippocampus and Social Impairment in Psychiatric Disorders. *Cold Spring Harb. Symp. Quant. Biol.* **83**, 105–118 (2018).
21. Grégoire, L. & Anderson, B. A. Semantic generalization of value-based attentional priority. *Learn. Mem.* **26**, 460–464 (2019).
22. Kurth-Nelson, Z., Barnes, G., Sejdinovic, D., Dolan, R. & Dayan, P. Temporal structure in associative retrieval. *eLife* **4**, e04919 (2015).
23. Wimmer, G. E. & Shohamy, D. Preference by Association: How Memory Mechanisms in the Hippocampus Bias Decisions. *Science* **338**, 270–273 (2012).
24. McFadyen, J., Liu, Y. & Dolan, R. J. Differential replay of reward and punishment paths predicts approach and avoidance. *Nat. Neurosci.* **26**, 627–637 (2023).
25. Wise, T., Liu, Y., Chowdhury, F. & Dolan, R. J. Model-based aversive learning in humans is supported by preferential task state reactivation. *Sci. Adv.* **7**, eabf9616 (2021).
26. Pike, A. C. & Robinson, O. J. Reinforcement Learning in Patients With Mood and Anxiety Disorders vs Control Individuals: A Systematic Review and Meta-analysis. *JAMA Psychiatry* **79**, 313–322 (2022).
27. Vandendriessche, H. & Palminteri, S. Neurocognitive biases from the lab to real life. *Commun. Biol.* **6**, 158 (2023).
28. Wimmer, G. E., Liu, Y., Vehar, N., Behrens, T. E. J. & Dolan, R. J. Episodic memory retrieval success is associated with rapid replay of episode content. *Nat. Neurosci.* **23**, 1025–1033 (2020).
29. Wimmer, G. E., Liu, Y., McNamee, D. C. & Dolan, R. J. Distinct replay signatures for prospective decision-making and memory preservation. *Proc. Natl. Acad. Sci.* **120**, e2205211120 (2023).

30. McFadyen, J., Liu, Y. & Dolan, R. J. Differential replay of reward and punishment paths predicts approach and avoidance. *Nat. Neurosci.* **26**, 627–637 (2023).
31. Schuck, N. W. & Niv, Y. Sequential replay of nonspatial task states in the human hippocampus. *Science* **364**, eaaw5181 (2019).
32. Wittkuhn, L. & Schuck, N. W. Dynamics of fMRI patterns reflect sub-second activation sequences and reveal replay in human visual cortex. *Nat. Commun.* **12**, 1795 (2021).
33. Molins, A., Stufflebeam, S. M., Brown, E. N. & Hämäläinen, M. S. Quantification of the benefit from integrating MEG and EEG data in minimum ℓ2-norm estimation. *NeuroImage* **42**, 1069–1077 (2008).
34. Goldenholz, D. M. *et al.* Mapping the signal-to-noise-ratios of cortical sources in magnetoencephalography and electroencephalography. *Hum. Brain Mapp.* **30**, 1077–1086 (2009).
35. Cichy, R. M. & Pantazis, D. Multivariate pattern analysis of MEG and EEG: A comparison of representational structure in time and space. *NeuroImage* **158**, 441–454 (2017).
36. Stokes, M. G., Wolff, M. J. & Spaak, E. Decoding Rich Spatial Information with High Temporal Resolution. *Trends Cogn. Sci.* **19**, 636–638 (2015).
37. Liu, Y. *et al.* Temporally delayed linear modelling (TDLM) measures replay in both animals and humans. *eLife* **10**, e66917 (2021).
38. Poldrack, R. A., Huckins, G. & Varoquaux, G. Establishment of Best Practices for Evidence for Prediction: A Review. *JAMA Psychiatry* **77**, 534–540 (2020).
39. Liu, J. *et al.* Transformative neural representations support long-term episodic memory. *Sci. Adv.* **7**, eabg9715 (2021).
40. Higgins, C. *et al.* Spatiotemporally resolved multivariate pattern analysis for M/EEG. *Hum. Brain Mapp.* **43**, 3062–3085 (2022).
41. Luyckx, F., Nili, H., Spitzer, B. & Summerfield, C. Neural structure mapping in human probabilistic reward learning. *eLife* **7**, e42816 (2019).
42. Kryza-Lacombe, M. *et al.* Changes in neural reward processing following Amplification of Positivity treatment for depression and anxiety: Preliminary findings from a randomized waitlist controlled trial. *Behav. Res. Ther.* **142**, 103860 (2021).
43. De La Peña-Arteaga, V. *et al.* An fMRI study of cognitive regulation of reward processing in generalized anxiety disorder (GAD). *Psychiatry Res. Neuroimaging* **324**, 111493 (2022).
44. Etkin, A. & Fonzo, G. A. Learning in Generalized Anxiety Disorder Benefits From Neither the Carrot Nor the Stick. *Am. J. Psychiatry* **174**, 87–88 (2017).
45. McDermott, T. J. *et al.* Striatal reactivity during emotion and reward relates to approach–avoidance conflict behaviour and is altered in adults with anxiety or depression. *J. Psychiatry Neurosci.* **47**, E311–E322 (2022).
46. Penner, A. E., White, E., Stoddard, J. & Gowin, J. L. The Neural Circuitry of Reward During Sustained Threat. *Cogn. Affect. Behav. Neurosci.* **22**, 134–144 (2022).
47. Aylward, J. *et al.* Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nat. Hum. Behav.* **3**, 1116–1123 (2019).

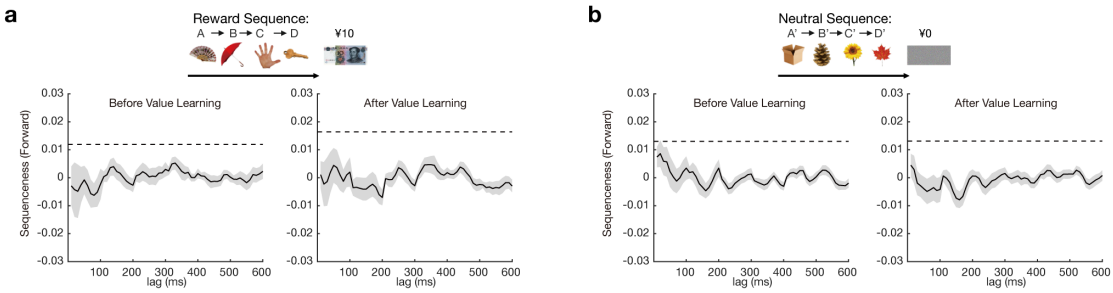
48. Chella, F. *et al.* The impact of improved MEG–MRI co-registration on MEG connectivity analysis. *NeuroImage* **197**, 354–367 (2019).
49. Gagne, C., Zika, O., Dayan, P. & Bishop, S. J. Impaired adaptation of learning to contingency volatility in internalizing psychopathology. *eLife* **9**, e61387 (2020).
50. Zinbarg, R. E., Williams, A. L. & Mineka, S. A Current Learning Theory Approach to the Etiology and Course of Anxiety and Related Disorders. *Annu. Rev. Clin. Psychol.* **18**, 233–258 (2022).
51. Sequeira, S. L., Forbes, E. E., Hanson, J. L. & Silk, J. S. Positive valence systems in youth anxiety development: A scoping review. *J. Anxiety Disord.* **89**, 102588 (2022).
52. Scult, M. A., Knodt, A. R., Radtke, S. R., Brigidi, B. D. & Hariri, A. R. Prefrontal Executive Control Rescues Risk for Anxiety Associated with High Threat and Low Reward Brain Function. *Cereb. Cortex* **29**, 70–76 (2019).
53. Sporn, S., Hein, T. & Herrojo Ruiz, M. Alterations in the amplitude and burst rate of beta oscillations impair reward-dependent motor learning in anxiety. *eLife* **9**, e50654 (2020).
54. Xia, L., Xu, P., Yang, Z., Gu, R. & Zhang, D. Impaired probabilistic reversal learning in anxiety: Evidence from behavioral and ERP findings. *NeuroImage Clin.* **31**, 102751 (2021).
55. Eckstrand, K. L. *et al.* Trauma-associated anterior cingulate connectivity during reward learning predicts affective and anxiety states in young adults. *Psychol. Med.* **49**, 1831–1840 (2019).
56. Liu, Q. *et al.* Neural function underlying reward expectancy and attainment in adolescents with diverse psychiatric symptoms. *NeuroImage Clin.* **36**, 103258 (2022).
57. Auerbach, R. P. *et al.* Reward-Related Neural Circuitry in Depressed and Anxious Adolescents: A Human Connectome Project. *J. Am. Acad. Child Adolesc. Psychiatry* **61**, 308–320 (2022).
58. Anderson, Z. *et al.* Association between reward-related functional connectivity and tri-level mood and anxiety symptoms. *NeuroImage Clin.* **37**, 103335 (2023).
59. Huys, Q. J., Pizzagalli, D. A., Bogdan, R. & Dayan, P. Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. *Biol. Mood Anxiety Disord.* **3**, 12 (2013).
60. Mason, L., Eldar, E. & Rutledge, R. B. Mood Instability and Reward Dysregulation—A Neurocomputational Model of Bipolar Disorder. *JAMA Psychiatry* **74**, 1275 (2017).
61. Gagne, C., Agai, S., Ramiro, C., Dayan, P. & Bishop, S. Biased belief priors versus biased belief updating: Differential correlates of depression and anxiety. *PLOS Comput. Biol.* **18**, e1010176 (2022).
62. LeDoux, J. E. Coming to terms with fear. *Proc. Natl. Acad. Sci.* **111**, 2871–2878 (2014).
63. LeDoux, J. E. As soon as there was life, there was danger: the deep history of survival behaviours and the shallower history of consciousness. *Philos. Trans. R. Soc. B Biol. Sci.* **377**, 20210292 (2022).

64. Michely, J., Rigoli, F., Rutledge, R. B., Hauser, T. U. & Dolan, R. J. Distinct Processing of Aversive Experience in Amygdala Subregions. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **5**, 291–300 (2020).
65. Liu, Y., Mattar, M. G., Behrens, T. E. J., Daw, N. D. & Dolan, R. J. Experience replay is associated with efficient nonlocal learning. *Science* **372**, eabf1357 (2021).
66. Amoruso, L., Pusil, S., García, A. M. & Ibañez, A. Decoding motor expertise from fine-tuned oscillatory network organization. *Hum. Brain Mapp.* **43**, 2817–2832 (2022).
67. Gurariy, G., Mruczek, R. E. B., Snow, J. C. & Caplovitz, G. P. Using High-Density Electroencephalography to Explore Spatiotemporal Representations of Object Categories in Visual Cortex. *J. Cogn. Neurosci.* **34**, 967–987 (2022).
68. Gwin, J. T. & Ferris, D. High-density EEG and independent component analysis mixture models distinguish knee contractions from ankle contractions. in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 4195–4198 (IEEE, 2011). doi:10.1109/IEMBS.2011.6091041.
69. Hsu, S.-H., Lin, Y., Onton, J., Jung, T.-P. & Makeig, S. Unsupervised learning of brain state dynamics during emotion imagination using high-density EEG. *NeuroImage* **249**, 118873 (2022).
70. Klumpp, H. *et al.* Neural responsiveness to reward and suicidal ideation in social anxiety and major depression before and after psychotherapy. *Biol. Psychol.* **178**, 108520 (2023).
71. Norbury, A., Hauser, T. U., Fleming, S. M., Dolan, R. J. & Huys, Q. J. *Different components of cognitive-behavioural therapy affect specific cognitive mechanisms*. https://osf.io/ydct5 (2023) doi:10.31234/osf.io/ydct5.
72. Shechner, T. *et al.* Attention biases, anxiety, and development: toward or away from threats or rewards?: Special Article: Attention Biases, Anxiety, and Development. *Depress. Anxiety* **29**, 282–294 (2012).
73. Raio, C. M., Hartley, C. A., Orederu, T. A., Li, J. & Phelps, E. A. Stress attenuates the flexible updating of aversive value. *Proc. Natl. Acad. Sci.* **114**, 11241–11246 (2017).
74. Kurth-Nelson, Z., Economides, M., Dolan, R. J. & Dayan, P. Fast Sequences of Non-spatial State Representations in Humans. *Neuron* **91**, 194–204 (2016).

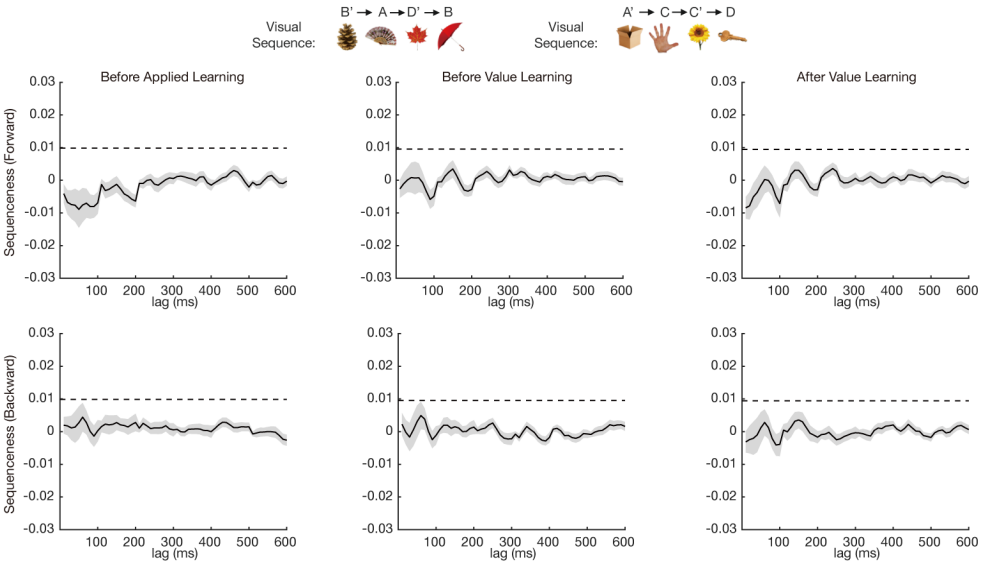
Supplementary Information



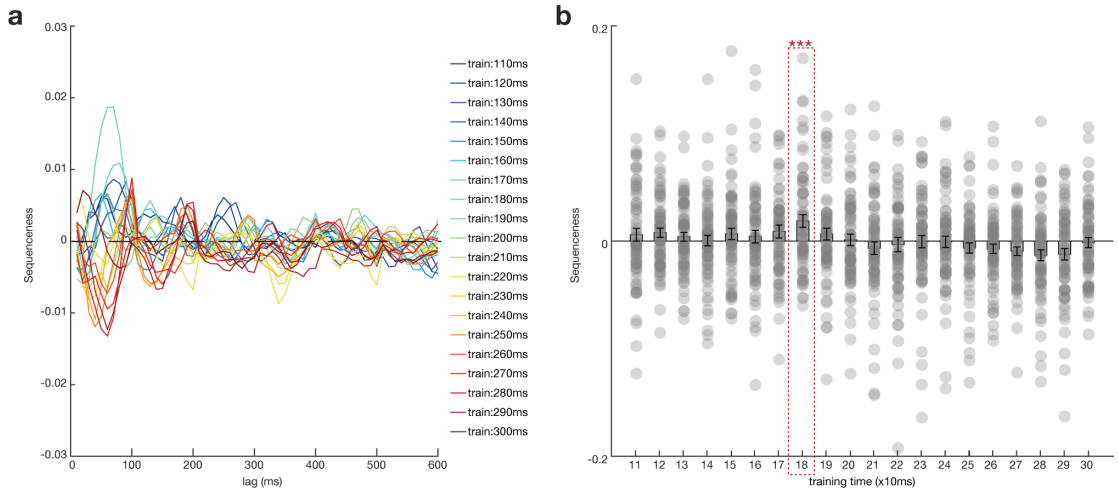
Extended Data Fig. 1| Sensor Maps, Spatial Correlation, and Classifiers Performance of Trained Lasso Logistic Regression Models, Scalp map, Related to Figure 2. a, Sensor map for each state decoding model is shown on the left, with correlation matrix between classifiers shown on the right. **b**, leave-one-out cross-validation results for each classifier in functional localizer task. These plots only use classifiers trained at 180 ms post stimulus onset. The x axis refers to the time point used for testing the classifiers. The target state of the decoder is indicated by line color, and the ground truth state is represented by a different panel for each state. The curves therefore have a different shape than plots made by varying both the training and testing time. **c**, Scalp map show response amplitude for each state at 180ms post stimulus onset.



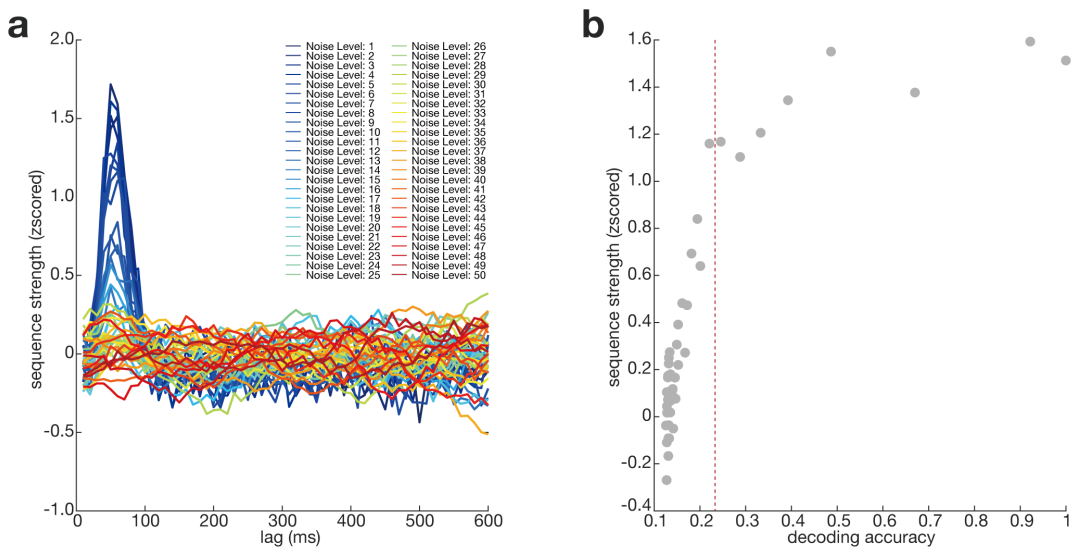
Extended Data Fig. 2| Forward Replay of True Sequences, Before and After Value Learning, related to Figure 2. No evidence of forward replay of sequences. The dotted line is the FWE-corrected significance threshold.



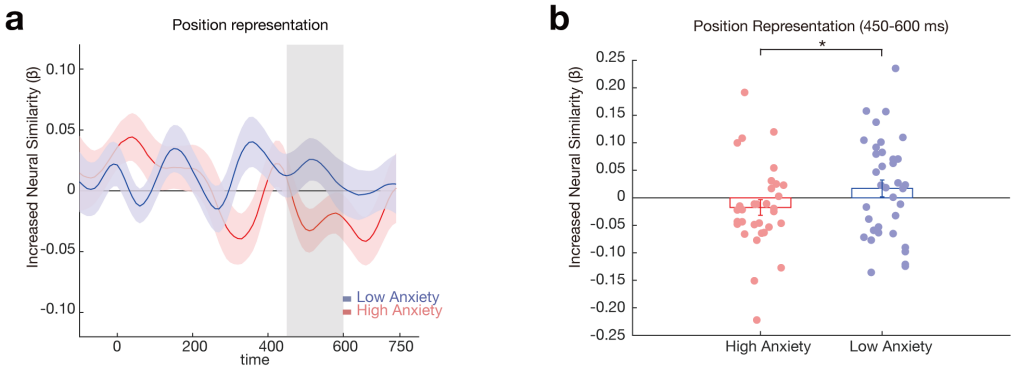
Extended Data Fig. 3| No Evidence of Replay of Visual Sequences, related to Figure 2. There was no evidence for replay of the visually experienced sequence during 3 resting states among the whole task, same with previous MEG studies. The dotted line is the FWE-corrected significance threshold.



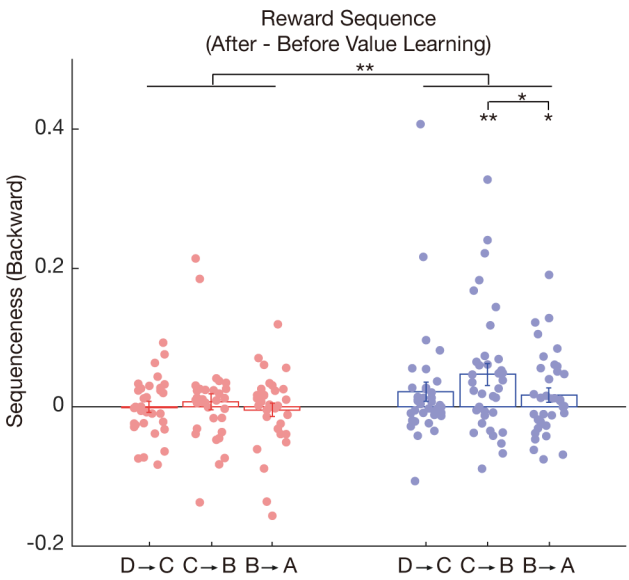
Extended Data Fig. 4| Sequenceness Analysis with Classifiers Trained at Different Times, Related to Figures 2. a, Sequenceness using classifiers trained at different times relative to stimulus onset (110 ms – 300 ms). This is from resting data after value learning. **b,** Scatterplot of the sequenceness at 60 ms time lag as a function of classifier training times during rest period after value learning. 180 ms is the training time used throughout the current study. *** is $p < 0.001$.



Extended Data Fig. 5| Simulation Results for Decoding Accuracy and Replay Strength, Related to Figures 2. In the simulation we varied decoding accuracy by injecting different levels of noise during classifier training. We then ask how does decoding accuracy influence replay strength (fixed and known in simulation) using current sequence detection methods. **a,** Sequence strengths were plotted for a classifier trained with different noise levels. **b,** The sequence strength at 60 ms lag (ground truth) is plotted against corresponding decoding accuracy. The red dotted line indicates the empirical decoding accuracy obtained in real data.



Extended Data Fig. 6| Neural Representation of Position with Anxiety, related to Figure 4. a, The temporal dynamics of the emergent (learning-induced) position representation, separately for low (blue) and high (red) anxious participants. **b,** Based on previous study[18], a 450-600ms time window after the stimuli onset (grey shaded area in panel a) was chosen as the region of interest (ROI). The ROI analysis revealed a significant increase of the position representation in low than high anxious subjects (t-test, $t(66) = 1.68$, $p = 0.049$). Error bars show 95% SEM; each dot indicates results from one subject. The solid line reflects the best robust linear fit, $*p < 0.05$.



Extended Data Fig. 7| Increase of Replay as a Function of Distance, related to Figure 3. a, Low anxious subjects (blue) showed significant increase of reverse replay for all pairwise associations, with stronger increase for transitions nearer the reward than distal to the reward (paired t test between $C \rightarrow B$ and $B \rightarrow A$, $t(35) = 1.65$, $p = 0.05$), while for subjects with high anxiety, there is no significant replay for any pairwise associations within reward sequence (red). Error bars show 95% SEM; each dot indicates results from one subject. $*p < 0.05$, $**p < 0.01$.