



Published in final edited form as:

Neuroimage. 2022 December 01; 264: 119709. doi:10.1016/j.neuroimage.2022.119709.

Large-scale neural network computations and multivariate representations during approach-avoidance conflict decision-making

Nicole Moughrabi^a, Chloe Botsford^c, Tijana Sagorac Gruichich^c, Ameera Azar^a, Mickela Heilicher^c, Jaryd Hiser^c, Kevin M. Crombie^a, Joseph E. Dunsmoor^{a,b}, Zach Stowe^c, Josh M. Cisler^{a,b,*}

^aDepartment of Psychiatry and Behavioral Sciences, University of Texas at Austin

^bInstitute for Early Life Adversity Research, University of Texas at Austin

^cDepartment of Psychiatry, University of Wisconsin-Madison

Abstract

Many real-world situations require navigating decisions for both reward and threat. While there has been significant progress in understanding mechanisms of decision-making and mediating neurocircuitry separately for reward and threat, there is limited understanding of situations where reward and threat contingencies compete to create approach-avoidance conflict (AAC). Here, we leverage computational learning models, independent component analysis (ICA), and multivariate pattern analysis (MVPA) approaches to understand decision-making during a novel task that embeds concurrent reward and threat learning and manipulates congruency between reward and threat probabilities. Computational modeling supported a modified reinforcement learning model where participants integrated reward and threat value into a combined total value according to an individually varying policy parameter, which was highly predictive of decisions to approach reward vs avoid threat during trials where the highest reward option was also the highest threat option (i.e., approach-avoidance conflict). ICA analyses demonstrated unique roles for salience, frontoparietal, medial prefrontal, and inferior frontal networks in differential encoding of reward vs threat prediction error and value signals. The left frontoparietal network uniquely encoded degree of conflict between reward and threat value at the time of choice. MVPA demonstrated that delivery of reward and threat could accurately be decoded within salience and inferior frontal networks, respectively, and that decisions to approach reward vs avoid threat were predicted by the relative degree to which these reward vs threat representations were active at the time of choice. This latter result suggests that navigating AAC decisions involves generating mental

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

*Corresponding author: Health Discovery Building, 1601 Trinity Blvd, Austin, TX 78701. josh.cisler@austin.utexas.edu (J.M. Cisler).

Credit author statement

All authors contributed meaningful to the manuscript. Contributed to design / hypotheses: JMC, NM, JED, ZS. Contributed to data collection: MH, CB, JH, TSG, AA, KMC, NM. Contributed to analyses and writing: MH, CB, JH, TSG, AA, KMC, NM, ZS, JED, JMC.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi: [10.1016/j.neuroimage.2022.119709](https://doi.org/10.1016/j.neuroimage.2022.119709).

representations for possible decision outcomes, and relative activation of these representations may bias subsequent decision-making towards approaching reward or avoiding threat accordingly.

Keywords

Reward learning; Threat learning; Approach-avoidance conflict; Computational modeling; Multivariate pattern analysis

1. Introduction

There is increasing interest in the complex neurobehavioral mechanisms that support effective decision-making to optimize reward and avoid threat. The complex nature of decision-making is revealed in part by work on the high dimensionality of the stimulus feature space (Leong et al., 2017; Niv et al., 2015), multi-step / sequential decision-making (Daw et al., 2011; Feher da Silva and Hare, 2020; Miller and Venditto, 2021), and forming and inferring latent states / cognitive maps of the learning environment (Behrens et al., 2018; Schuck et al., 2016; Wikenheiser and Schoenbaum, 2016). An important domain of real-world decision-making involves navigating situations that involve the probability of receiving some amount of reward at the risk of an aversive outcome. That is, the same behavior or set of behaviors involved in obtaining potential reward (e.g., attending a social gathering in hopes of having a wonderful time) often contain some potential for an aversive outcome (e.g., attending the gathering at the risk of feeling awkward or socially rejected). This dual-outcome aspect of decision-making can create conflict between motivation / goals to approach vs avoid, termed approach-avoidance conflict, and has received increased recent research attention due to its high relevance for stress and anxiety-related disorders (Aupperle et al., 2015; Bach et al., 2014; Kirlic et al., 2017; Weaver et al., 2020), substance use disorders (Fleming and Bartholow, 2014; Martin Braunstein et al., 2016; Piercy et al., 2021; Stormark et al., 1997), and bipolar spectrum disorders (Alloy et al., 2006, 2009; Hirshfeld-Becker et al., 2003). For example, in the case of posttraumatic stress disorder (PTSD), the problem is not simply that an individual seeks to avoid trauma reminders; rather, through avoidance of trauma reminders the individual sacrifices rewarding and meaningful aspects of their life, thereby leading to functional impairment. Accordingly, increased mechanistic understanding of approach-avoidance conflicted decision-making may aid in the development of cognitive models of decision-making relevant to the development, maintenance, and treatment of psychopathology.

Prior neuroimaging of approach-avoidance conflict demonstrates lateral prefrontal cortex (PFC), anterior insula and inferior frontal gyrus, and dorsal anterior cingulate cortex (dACC) engagement when participants make decisions during conflict between reward and threat contingencies (Aupperle et al., 2015; Ironside et al., 2020; Schlund et al., 2016; Shenhav et al., 2014). Studies using an adapted version of foraging tasks (Bach et al., 2014; Khemka et al., 2017) or a task where individuals first learn reward / threat value of face-scene pairs and then have to recall these associations (O'Neil et al., 2015) also demonstrate an important contribution of the hippocampus to approach-avoidance conflict. One recent study using a sophisticated hierarchical Bayesian modeling approach to account for heterogeneity in

task activation due to deliberation time and individual variability in approach-avoidance tradeoff preferences suggested approach-avoidance conflicted decision-making may more specifically engage lateral and dorsomedial PFC (Zorowitz et al., 2019). In the current study, we build upon and address gaps in this literature by investigating neurocomputational mechanisms that manage decisions under dual-outcomes of risk and threat in the context of an approach-avoidance paradigm. Specifically, we used an adapted version of bandit task (Weaver et al., 2020)(Behrens et al., 2007, 2008; Daw et al., 2006) that includes varying probabilities of receiving both reward and threat outcomes, which allowed us to 1) apply a computational model of learning to directly compare neurocircuitry encoding of decision-making for reward vs threat and during approach-avoidance conflict, and 2) characterize episodic memory and prospective mental representations for reward vs threat as mechanisms resolving approach-avoidance conflict.

One key element that has been missing in the existing literature of neurocircuitry mediating approach-avoidance conflict decision-making is leveraging of the vast existing computational neuroscience literature on learning and decision-making using variations of bandit tasks (Sutton and Barto, 2018). Indeed, some common versions of approach-avoidance conflict tasks lack a learning component and instead present all necessary information to participants, which then resets on each trial. This precludes the ability to probe dynamic processes by which individuals interact with and learn from the environment. Additionally, adapting bandit tasks to probe approach-avoidance conflict allows using versions of the well-studied Rescorla-Wagner (RW) model of learning. Moreover, adopting the RW model to understand approach-avoidance learning affords the ability to model reward and threat processes, and their interactions, simultaneously. For example, applying the RW model to a bandit task with concurrent threat and reward allows identifying brain regions that uniquely encode reward vs threat computations (e.g., value expectation, prediction errors), and would formalize approach-avoidance conflict as jointly elevated reward and threat expectations for a given action. While there have been many separate investigations of neurocircuitry encoding reward or threat computations (Daw et al., 2006, 2011; Delgado et al., 2008; Garrison et al., 2013; Homan et al., 2019; Li et al., 2011; Lindström et al., 2018; Niv et al., 2012), there has been a paucity of research directly comparing the computational neurocircuitry of reward and threat (Metereau and Dreher, 2013). Some work investigating neural correlates of prediction error encoding defines valence as the sign of the prediction error signal (Fouragnan et al., 2018). However, this definition conflates the magnitude of the outcome (more or less than expected) with the appetitive vs aversive affective properties of the outcome (Carroll et al., 1999). That is, more reward than expected (i.e., positive prediction error) would be appetitive, but more threat than expected (i.e., also a positive prediction error) would be aversive. Similarly, less reward than expected would be aversive (Tom et al., 2007), but less threat than expected would be appetitive (Kalisch et al., 2019). One way to address this conceptual and empirical gap is through tasks that concurrently manipulate reward and threat probabilities, allowing a direct comparison of neurocircuitry encoding of reward vs threat learning. Given that approach-avoidance conflict necessarily involves concurrent reward and threat expectations, differentiating reward and threat learning is essential for understanding approach-avoidance conflict.

An additional gap in the literature on the neurocircuitry of approach-avoidance conflict is lack of integration of the role of episodic memory and mental simulation in decision-making (Biderman et al., 2020; Dasgupta and Gershman, 2021; Gershman and Daw, 2017; Schacter et al., 2017). Some work in this literature suggests that at the time of choice, reactivation of memory representations, via reactivation of neural patterns at the time of encoding (i.e., neural replay), functions to maintain a cognitive map of the decision-making environment (Gillespie et al., 2021; Mızrak et al., 2021). In this way, memory retrieval at the time of choice functions to retrieve evidence to inform the subsequent choice. At the same time, other work in this literature highlights the role of memory reactivation at the time of choice as a form of mental simulation of future events (i.e., neural ‘preplay’) (Biderman et al., 2020; Doll et al., 2015; Gluth et al., 2015; Schacter et al., 2017; Shadlen and Shohamy, 2016; Yu and Frank, 2015; Zielinski et al., 2020). In this way, memory retrieval functions to prospectively predict possible outcomes of a chosen action(s), thereby potentially biasing behavior based on the perceived prospective value of the predicted outcome. This mechanism for memory in decision-making fits well into the established literature on model-based learning systems (Daw, 2018; Daw et al., 2011; Feher da Silva and Hare, 2020; Gläscher et al., 2010), where individuals form an abstract cognitive map of the learning environment that is used to inform a prospective, rather than trial-and-error, strategy for decision-making (Doll et al., 2015).

Both of these non-mutually exclusive roles for memory in decision-making could inform the neurocircuitry mechanisms of approach-avoidance conflict. When confronted with approach-avoidance conflict, it seems more likely that the source of the conflict is less about the affective properties of the presented stimuli that signal the possibility for reward or threat and more about the imagined future reward or threat outcome states. As such, there is likely a key role for reactivation of memory representations during these decisions, either as a means of navigating the cognitive map of the task space or as a means of prospectively predicting the consequences of an action. While memory reactivation during choice has been investigated during both reward tasks and aversive learning tasks (Castegnetti et al., 2020; Doll et al., 2015; Wise et al., 2021), it has never been investigated during approach-avoidance conflict. Here, we seek to leverage a multivariate pattern analysis (MVPA) approach to characterize the role of reactivation of reward and threat memory representations at the time of choice during approach-avoidance conflict decision-making. In line with prior studies (Doll et al., 2015; Wise et al., 2021), a role for episodic memories in approach-avoidance decision-making would be supported by observing that representations of reward vs threat at the time of choice predict subsequent approach vs avoidance decisions during conflict.

Finally, we sought to define neurocircuitry encoding of threat, reward, and approach-avoidance conflict, and prospective memory representations for reward and threat as mechanisms for resolving approach-avoidance conflict, using large-scale network analyses. In contrast to univariate voxelwise approaches, dominant models of human brain organization emphasizes distributed information processing across spatially distributed networks (Avena-Koenigsberger et al., 2018; Bressler and Menon, 2010; Bullmore and Sporns, 2009, 2012; Menon, 2011; Meunier et al., 2010; Pessoa, 2017, 2018). For example, three canonical networks that are highly reproducible across both rest and numerous types of

cognitive tasks are frontoparietal, salience, and default mode networks (Bressler and Menon, 2010; Menon, 2011; Smith et al., 2009). In line with these models, we use independent component analysis (ICA) (Calhoun et al., 2001, 2003; Calhoun and de Lacy, 2017) to define large-scale functional networks. ICA provides a spatial map of each network (e.g., spatial map indicating each voxel within the salience network) as well as a timecourse of activity for each network. Notably, identifying large-scale networks with ICA allows one to characterize both 1) the activity of a network in response to the computational mechanisms (e.g., to what degree does the salience network's timecourse of activity scale with the magnitude of prediction errors?) and 2) multivariate patterns of activity of the voxels within a network that represent mental representations for threat and reward. In line with prior RW modeling studies using ICA networks (Cisler et al., 2019; Letkiewicz et al., 2020, 2022; Ross et al., 2018), we hypothesized that reward and threat prediction errors would be encoded in salience and striatum networks, and reward and threat value expectation would be encoded in medial prefrontal networks. In line with the previous study approach-avoidance conflict (Zorowitz et al., 2019), we hypothesized approach-avoidance conflict would be encoded in a frontoparietal network. With respect to specific networks encoding prospective mental representations, there is not clear prior data upon which to draw; however, given their roles in reward and threat learning (Cisler et al., 2019, 2020; Delgado et al., 2008; Fullana et al., 2016; Li et al., 2011; Shackman et al., 2011), we hypothesized memory reactivation would be encoded in striatum and salience networks.

2. Methods

2.1. Participants

Participants consisted of 28 healthy adults ($n = 9$ male) recruited from the general community. Mean age was 31.2 (SD = 9.27). Participants underwent structured clinical assessments (Tolin et al., 2018) to confirm absence of current mental health disorders and the Wechsler Abbreviated Scale of Intelligence to rule out neurocognitive disabilities. Exclusion criteria consisted of MRI contraindications, estimated full scale IQ < 75 , current mental health disorders, major medical disorders, loss of consciousness greater than 30 min, and age less than 21 or greater than 50. An additional 4 participants completed the study but were removed from analyses due to either missing $> 15\%$ of trials on the task ($n = 3$) or excessive head motion ($n = 1$).

This study was conducted with approval from the Institutional Review Board of the University of Wisconsin Madison and all participants provided informed consent.

2.2. Approach avoidance conflict task (AACT)

In this fMRI adaptation of our previously used task (Weaver et al., 2020), participants completed 150 trials of a three-arm bandit task (see Fig. 1), in which each arm was associated with separate probabilities for receiving both reward (points, ranging from -10 to $+10$) and threat (electric shock) outcomes. The first 30 trials consisted of a baseline phase, where only reward outcomes were presented. Trials 31–150 presented both reward and threat outcomes, and probabilities for reward and threat associated with each arm switched every 30 trials in order to alternate between two distinct phases. In congruent phases,

the arm most likely to lead to higher point outcomes was least likely to lead to threat outcomes. In conflict phases, the arm most likely to lead to higher point outcomes was most likely to lead to threat outcomes, thus creating approach-avoidance conflict. To motivate performance, participants were told they could earn additional bonus compensation based on their performance on the task (i.e., how many points they earned).

Each trial consisted of three phases: choice, anticipation, and outcome. In the choice phase, participants were presented with each of the three response arms. Upon making a choice, the selected arm was high-lighted. Failing to make a response within 3.5 s resulted in a notification that the participant took too long to respond and missed the trial. The anticipation phase for threat then followed, which consisted of a fixation cross lasting 2–5 s. Following the anticipation phase for threat was the outcome phase for threat, during which an image appeared in the location of the selected arm that indicated either a shock, concurrent with the actual delivery of shock, or an image indicating the absence of a shock. These shock outcome images appeared for 3 s. Next was the anticipation phase for reward, which consisted of a fixation cross lasting 2–5 s. Following the anticipation phase for reward was the outcome phase for reward, during which an image appeared in the location of the selected arm indicating the number of points won in green font or number of points lost in red font for 3 s, followed by an additional fixation inter-trial-interval (ITI) of 2–4 s.

2.3. Computational modeling

Towards the goal of defining computational mechanisms of reward, threat, and approach-avoidance decision-making, we compared fit of three models, all of which are variations of a standard Rescorla-Wagner (RW) model (Rescorla and Wagner, 1972; Sutton and Barto, 2018): 1) a reward only model (R model), positing that participants ignore the irrelevant threat outcomes and only use information about reward outcomes to guide decisions, 2) a reward and threat model using a single learning rate (RT single model), positing that participants integrate reward and threat expectations into a combined total value and using one learning rate for both reward and threat outcomes, and 3) the same reward and threat model but using separate learning rates for reward and threat outcomes (RT double model). As the primary objective of the task given to participants was to maximize points earned, participants were told monetary compensation was in proportion to points earned on the task, and the threat outcomes are irrelevant to this primary task, we did not test a threat only model.

The R model updates expected reward value of a chosen option (i.e. reward expectation), VR , based on the magnitude of the prediction error, δ_R (observed outcome – V), scaled by a learning rate α (ranging from 0 to 1): $VR_{t+1} = VR_t + \delta_R * \alpha$. The learning rate, α_R , controls the speed of updating VR , with higher learning rates resulting in faster changes in expected value.

The RT single model used an identical model for updating reward expectation and also included a parallel model for threat value expectations based on threat prediction errors, δ_T , VT : $VT_{t+1} = VT_t + \delta_T * \alpha$. Here the same learning rate was used for both the VR and VT updating. The separate expectations for VR and VT were then integrated into a combined

value, VC , using an individually varying and scalar policy parameter, π , that represents an individual's preference for threat vs reward: $VC_t = VR_t (1-\pi) + (1-VT_t) \pi$.

The RT double model was identical to the RT single model, except a separate learning rate was used for reward, α_R , and threat, α_T , outcomes.

For each model, a softmax function transformed expected value into action probabilities by using an exploration / exploitation β parameter. A higher softmax β represents a tendency to exploit high value responses / respond more consistently; a lower softmax β represents a tendency to explore lower value options / respond more inconsistently.

Accordingly, the R model had 2 free parameters (α , β), the RT single model had 3 free parameters (α , β , π), and the RT double model had 4 free parameters (α_R , α_T , β , π).

Model fitting was conducted using a Hierarchical Bayesian Inference (HBI) approach (Piray et al., 2019). Models were compared using both log likelihood comparisons, as the R and RT single models are both nested within the RT double model, and Akaike Information Criterion comparison. VR , VT , δ_R , and δ_T , from the best fitting model (RT single model, see below) were carried forward to the fMRI analyses using mean sample parameters (Cisler et al., 2019; Daw et al., 2006). We additionally created a threat vs reward conflict variable, calculated as the max difference across the three arms in expected reward and threat on a given trial (i.e., $VR - (1-VT)$); that is, the greatest amount of concurrently heightened expectations for reward and threat in any arm on the current trial.

2.4. MRI procedures and first-level analyses

Data acquisition and preprocessing are described in the supplement. Following preprocessing, design matrices were created using AFNI's 3dDeconvolve. These design matrices included columns for the following task events: missed trials, choice phase, anticipation phase for threat, threat outcome, anticipation phase for reward, and reward outcome. Choice phase was modulated by reward expectation, threat expectation, and reward-threat conflict; anticipation phase for threat was modulated by threat expectation, anticipation phase for reward was modulated by reward expectation, threat outcome was modulated by threat prediction errors, and reward outcome was modulated by reward prediction errors. Parametric modulation was implemented with 3dDeconvolve's "stim_times_AM2" option.

2.5. Electric shock delivery

BIOPAC MP160 Data Acquisition System and BIOPAC STM100C module was used to administer shocks using pre-gelled electrodes placed on the skin of the fleshy portion of the mediolateral, left lower leg, directly over the tibialis anterior. Amperage on the stimulation device was set to the maximum (50 mA) to allow the greatest range of intensity selections. Participants were told to select an intensity of a 5/10.

2.6. Independent component analysis

Towards the goal of identifying network-level neurocircuitry processes engaged during approach-avoidance decision-making, we used Independent Component Analysis (ICA)

(Calhoun et al., 2001), which provides both a spatial map indicating distributed voxels that comprise the network as well as a timecourse of activity for the network. We used ICA with a model order of 40 components, which delivered a good balance between component reliability estimated across 50 ICASSO iterations and interpretability of canonical networks. Eight of the 40 components were deemed functional networks of interest after visual inspection (see Fig. 2). Components arising from artifacts of head motion or CSF and components of non-interest (i.e. motor, sensorimotor, and visual networks), which are not as relevant for the cognitive processes of interest here, were excluded.

Network timecourses were regressed onto design matrices described above using custom scripts in Matlab in order to estimate β coefficients that were carried forward into second-level analyses. Second-level analyses consisted of linear mixed effects models (LMEMs) testing for overall vs domain (reward vs threat) specific activation for each network, conducted separately for the β coefficients for PE, value expectation during choice, value expectation during anticipation, and reward vs threat expectation conflict during the choice phase. These models additionally included covariates for head motion and the threat weighting policy parameter, π (Zorowitz et al., 2019). Bonferroni correction controlled for multiple comparisons across networks (i.e., $0.05 / 8 = \text{corrected } p < .0063$).

2.7. Voxelwise mass univariate analyses

Complimentary to the primary ICA analyses, we conducted standard voxelwise GLMs, regressing each voxel's timecourse onto the above-described design matrix using AFNIs 3dREMLfit. Second-level analyses were identical to that described for the ICA analyses, resulting in separate whole-brain t value maps for PEs, value expectation during choice, value expectation during anticipation, and reward vs threat expectation conflict during choice. We corrected for whole-brain comparisons using cluster-level thresholding based on contemporary recommendations (Cox et al., 2017; Eklund et al., 2016), in which a corrected $p < .05$ was achieved through 21 contiguous voxels surviving an uncorrected $p < .001$.

2.8. Multivariate pattern analyses of mental representations during choice

We sought to test the hypothesis that representations of threat vs reward delivery that were active at the time of choice were related to subsequent decisions to approach vs avoid on high conflict trials. Fig. 3 provides an overview of the analytical approach, which was implemented separately for each of the eight ICA networks of interest. The first step was to demonstrate that network activity patterns at the time of threat and reward delivery could accurately be decoded. Each participants' trial-by-trial activation patterns at the time of point and shock delivery were characterized using 3dLSS. The timepoint \times voxel matrices were centered within each timepoint to ensure no differences in overall activation levels across trials. Support vector machines (SVM), using a radial basis function kernel implemented in Matlab through libsvm (Chang and Lin, 2011), were used to decode shock (binary classification) and points (epsilon regression for the continuous points variable scaled to a range of -1 to 1). We established the accuracy of the decoders using 5-fold cross-validation (matlab's `cvpartition.m` function) across subjects (i.e., decoders were trained on one set of subjects and then tested on an independent set of subjects' data). To account for any bias in the random partitioning of subjects into the 5 folds, we repeated this process 10 times and

took the average across all iterations (Zhou et al., 2021). The shock decoder accuracy was defined as the mean of sensitivity and specificity, the point decoder accuracy was defined as the Pearson correlation between predicted points and observed points. This process was repeated separately for each of the 8 functional networks of interest.

After testing accuracy of the shock and point decoders, the next step was to apply the shock and point decoders to participant's data at the time of choice. 3dLSS was used to define trial-by-trial activation at the time of choice. A leave-one-out approach was used, such that the shock and reward decoders were trained on all participants' shock and points activity data except for one, and the resulting decoders were applied to the left-out subject's choice data. This resulted in hyperplane distances representing the degree to which the trained multivariate patterns (shock outcomes or scalar reward values) were active at the time of choice. This process was repeated separately for each ICA network of interest, resulting in unique predictions (i.e., hyperplane distances) about reward and threat for each separate network. For example, reward and threat decoders were defined the salience network and then applied to the left-out subject's choice data, resulting in predictions about reward and threat specifically from the salience network. Next, this process was repeated for the frontoparietal network, resulting in unique predictions about reward and threat specifically from the frontoparietal network. This allowed us to compare the degree to which patterns unique to each network were reactivated at the time of choice.

Our primary interest was comparing activation of the reward vs threat representations for a given network on high conflict trials where the individual chose to approach reward vs avoid threat. We defined high conflict trials as those trials in which: 1) there was a unique high reward arm (i.e., there was not more than one high reward arm with equal reward expectations), and 2) the high reward arm was also the arm with the highest threat expectations. On these high conflict trials, we defined approach decisions as those where the individual chose the high reward arm. We defined avoid decisions on these trials as those where the individual chose the lowest threat arm. These criteria created a very stringent definition of approach and avoidance decisions: approach decisions only referred to choosing the highest reward arm despite it being the highest threat arm; avoidance decisions only referred to choosing the *lowest* threat arm when the high reward arm was also the high threat arm. The mean percentage of approach vs avoidance decisions on these conflict trials for this sample was 0.71 (SD = 0.15). Supplemental Figure 1 provides a histogram of the extent of approach vs avoidance decisions among the sample.

We again used LMEMs to compare hyperplane distances on high conflict trials where the individual subsequently chose to approach vs avoid (i.e., a within-subjects factor), and including covariates for head motion and the threat weight policy parameter to account for between-subject sources of variance. Separate LMEMs were run for the point and shock hyperplane distances and for each network, using Bonferroni correction for alpha inflation across the separate models for each network.

3. Results

Behavioral Performance on RL tasks.

The adaptation of the RW model to integrate reward and threat values according to an individually varying threat weight policy parameter, with a single learning rate, fit the data significantly better than the reward only model. There were no significant differences in model fit between the RT single and RT double models (Figs. 4A–B). Figs. 4C and 4D depict individual differences in the three free parameters of the model. Demonstrating validity of the threat weight policy parameter, we observed a strong relationship between individual differences in this parameter and the fraction of high conflict trials where individuals chose to approach vs avoid (Fig. 4E).

Network encoding of prediction errors.

As indicated in Table 1 and Fig. 5A, the LMEMs demonstrated that several networks significantly encoded prediction errors. The right frontoparietal network significantly encoded overall negative prediction errors, with no differences between reward and threat. There was a significant difference between reward and threat PE encoding in the salience, lateral PFC, right inferior frontal, and default mode networks. The salience and inferior frontal networks both encoded positive threat prediction errors and negative reward prediction errors. The lateral PFC uniquely encoded negative threat prediction errors and did not encode reward prediction errors. The DMN encoded positive reward prediction errors and negative threat prediction errors.

Network encoding of value during choice.

As indicated in Table 1 and Fig. 5B, the LMEMs demonstrated that several networks significantly encoded value during choice. The DMN significantly encoded overall value during choice with no differences between reward and threat. The medial PFC, lateral PFC, and right frontoparietal networks each positively encoded threat value and negatively encoded reward value. The right inferior frontal network positively encoded reward value but negatively encoded threat value.

Network encoding of value during anticipation.

As indicated in Table 1 and Fig. 5C, the LMEMs demonstrated that several networks significantly encoded value during anticipation. The striatum significantly encoded overall value during anticipation with no differences between reward and threat. The salience and right inferior frontal networks positively encoded threat value but not reward value. The DMN negatively encoded threat value but not reward value.

Network encoding of conflict between reward and threat value during choice.

As indicated in Table 1 and Fig. 5D, the LMEMs demonstrated that the only network that was significantly related to conflict during choice was the left frontoparietal network.

Complementary voxelwise analyses.

As depicted in Figs. 6 and Supplementary Tables 1, the voxelwise analyses demonstrated univariate activation patterns that largely mirrored the ICA networks. Of note, however, the voxelwise analyses demonstrated significant encoding of reward prediction errors in the anterior striatum (caudate, putamen, and nucleus accumbens; Fig. 6B), whereas the ICA striatum network, which included the entire striatum, was not detected as related to reward prediction errors.

3.1. Multivariate representations during choice

The cross-validation tests using five-fold cross validation across subjects demonstrated that reward and threat could each be decoded accurately in all networks tested (Fig. 7A–B). This analysis validated that reward and threat decoding from one group of subjects could accurately predict reward and threat in an independent group of subjects. It was next tested whether the reward and threat decoding could be applied at the time of choice to define trial-by-trial representations that predicted approach vs avoidance decisions during high conflict trials of the task. The approach used a leave-one-out procedure, such that the reward and threat decoders were trained on an N-1 set of subjects and then applied to the choice data of the left-out subject.

LMEMs demonstrated that predictions from the reward decoder (i.e., hyperplane distances representing predictions about reward) in the salience network significantly differed between approach and avoidance trials, $t(52) = 4.33$, $p < .001$, such that reward representations were higher during choice phases where the participant subsequently chose to approach compared to choice phases where the participant chose to avoid. No other networks were significant for the reward decoder. By contrast, the LMEM demonstrated that predictions from the threat decoder in the right inferior frontal network were significantly greater during choice phases where the participant subsequently chose to avoid compared to choice phases where the individual subsequently chose to approach, $t(52) = 3.02$, $p = .004$. No other networks were significant for the threat decoder.

To understand the relative degree of activation of these reward and threat representations on conflict trials, we z-scored each subjects hyperplane predictions across all trials and compared predictions from the decoders for high conflict trials, separated into approach and avoid trials, vs low conflict trials. As can be seen in Fig. 7C and Fig. 7D, the degree of reward and threat representations in these networks during approach and avoid trials reflect significant or marginally significant deviations from non-conflict trials for both networks (salience network: approach trials vs low conflict trial $t(52) = 2.76$, $p = .008$; avoid trials vs low conflict trials $t(52) = -2.86$, $p = .006$; right inferior frontal network: approach trials vs low conflict $t(52) = -2.71$, $p = .009$; avoid trials vs low conflict trials $t(52) = 1.83$, $p = .07$).

4. Discussion

This study investigated the neurocomputational mechanisms supporting approach-avoidance conflict decision-making. Results demonstrated large-scale functional network encoding of computational mechanisms related to reward and threat, approach-avoidance conflict,

and prospective mental representations during decision-making. Given the relatively small sample size ($N = 28$), results should be considered preliminary and in need of additional replication. We found that the single learning rate model fit the data as well as the model including separate learning rates for reward and threat. This could suggest that individuals learn at similar rates for reward and threat. An important caveat in this interpretation, however, is that the primary goal of the task was to earn reward and the threat outcome was unrelated to this goal. As such, the observation that participants nonetheless learned at the same rate for reward and threat outcome could actually suggest a primacy for threat learning. Future research comparing threat and reward learning in different contexts is needed to further differentiate these domains of learning.

With respect to encoding of computational parameters, results suggested both unique and common network encoding of reward and threat. The salience network and right inferior frontotemporal network both encoded positive threat prediction errors and negative reward prediction errors. While the sign of encoding differed between reward and threat prediction errors, the affective valence of encoding was in agreement. The DMN encoded positive reward prediction errors and negative threat prediction errors, again consistent with affective valence-specific prediction error encoding. By contrast, the right frontoparietal network encoded both reward and threat negative prediction errors, suggesting a sign-specific role for this network in negative prediction error encoding. Value expectation during choice was encoded for both reward and threat in the DMN. The medial PFC, lateral PFC, and right frontoparietal networks all demonstrate affective valence-specific encoding of expected value, such that these networks encoded greater threat expectancies and lower reward expectancies (i.e., reward losses). The right inferior frontotemporal network was again reversed in encoding, with greater reward expectancies and lower threat expectancies encoded in this network. The anticipation of expected value demonstrated similar valence-specific encodings in the mPFC, with greater expectancies for reward and less expectancies for threat encoded in this network. By contrast, the salience and right inferior frontotemporal networks uniquely encoded expectation for threat and not at all for reward, whereas the striatum encoded expected value for both reward and threat consistent with domain general role for expected value encoding the striatum.

Overall, these data more broadly suggest affective valence-specific encoding of reward and threat computations, such that the delivery or anticipation of a threatening stimulus is encoded similarly to delivery or expectation of reward loss, and lack of delivery or expectation of lack of a threatening stimulus is encoded similarly to delivery of positive reward (Kalisch et al., 2019; Tom et al., 2007), rather than suggesting that reward and threat computations are encoded entirely distinctly. Note that this definition of affective valence refers to positive valence as appetitive and negative valence as aversive (Carroll et al., 1999; Kuppens et al., 2013), and not to the sign of the prediction error (Fouragnan et al., 2018), which appears to better account for shared encoding between appetitive / aversive nature of reward vs threat obtainment vs loss. However, the data did suggest a sign-specific, rather than affective valence specific, role for the frontoparietal network in encoding negative prediction errors, suggesting that the tracking of negative prediction errors, regardless of valence, is unique to the frontoparietal network. Similarly, the striatum and DMN encoded positive expectations for both reward and threat, suggesting a sign-specific, rather than

valence-specific, function for expectation of outcomes in these networks (Delgado et al., 2008; Metereau and Dreher, 2013). While we did not find significant encoding of PEs, either for reward or threat, in the striatal network, the voxelwise analyses demonstrated significant overall encoding of PEs in the ventral striatum (Supplementary Table 1), consistent with prior research (Delgado et al., 2008; Li et al., 2011), and that ventral striatum encoding of PEs was significantly greater for reward compared to threat.

With respect to approach-avoidance encoding, the current results are consistent with a prior modeling study demonstrating that approach-avoidance conflict was encoded in dmPFC and lateral PFC (Zorowitz et al., 2019) and demonstrate significant conflict encoding in a left frontoparietal network. The results are not consistent with all prior neuroimaging investigations of approach-avoidance conflict, namely in not identifying salience network activity, which would have corresponded with prior reports of dACC and anterior insula activation (Aupperle et al., 2015; Ironside et al., 2020). Here, we modeled conflict as the degree of concurrent expectations for reward and threat while also separately controlling for degree of threat expectation and reward expectation. This approach, while limited in the possibility for shared variance between the three regressors, ensures that any variance in brain activity due to conflict is not confounded by threat or reward expectation. Additionally, the modeling approach used here allowed for specific calculations about reward and threat expectations unique to each individual and trial. As such, the conflict estimate could approximate the individual's internal model of the task, rather than assuming the individual's internal model corresponded with the task probabilities. This is not a trivial difference: because the reward and threat outcomes are probabilistic, it often occurs that at various times throughout the task the outcomes the individual experiences widely diverge from the true probabilities. The modeling approach allowed us to account for this variance and likely create reward and threat expectations that more accurately represented the individual's internal model. Using this approach, and consistent with the prior modeling study (Zorowitz et al., 2019), these data suggest that conflict between approach and avoidance, may more specifically engage the frontoparietal network and likely the higher-order prospective decision-making functions associated with it.

With respect to prospective mental representations at the time of approach-avoidance decision-making, the current results are consistent with prior reports (Doll et al., 2015; Wise et al., 2021) that mental representations of the possible outcomes of an action are active at the time of choice and predict subsequent decisions. Here, mental representations of reward and threat were differentially encoded in the salience network and inferior frontotemporal network. Greater representations for reward in the salience network at the time of choice on high conflict trials predicted decisions to approach; greater representations for threat in the inferior frontal network at the time choice predicted decisions to avoid. The salience network identified here had relatively greater involvement of the dorsal ACC / dmPFC with additional loadings in the anterior insula, whereas the inferior frontotemporal network had relatively greater involvement of the inferior frontal gyrus and anterior insula. These dissociations are noteworthy for the hypothesized role of the dACC / dmPFC in conflict representation, outcome prediction, and adaptive control (Alexander and Brown, 2011; Brown and Braver, 2005; Shackman et al., 2011), and it seems intuitive that reward representations within an adaptive control mechanism become activated when overriding

avoidance tendencies to approach reward. Similarly, threat representations in the anterior insula and inferior frontal gyrus could be consistent with roles of aversive and anticipatory processing in the anterior insula and inferior frontal gyrus (Cornwell et al., 2017; Drabant et al., 2011; Geng et al., 2018; Radoman et al., 2021).

One account of these findings is that threat and reward representations at the time of choice act as prospective predictions and simulation of future events (i.e., memory ‘preplay’) (Gilbert and Wilson, 2007; Schacter et al., 2017). In this way, differential simulation of future reward vs threat would bias behavior towards approach vs avoidance, respectively. A more recent and alternative view is memory activation at the time of choice does not bias behavior directly; rather, memory activation strengthens representations of specific experiences to facilitate long-term memory (Gillespie et al., 2021). In this way, memory activation at the time of choice reinforces specific elements of the architecture of the cognitive map that later informs behavior, rather than memory activation being directly linked to a behavior. An interesting observation in the current data is that representations for reward vs threat were not just greater on high conflict trials where individuals approached vs avoided; these representations were also suppressed on trials where the opposing behavior was selected relative to non-conflict trials. These data suggest that conflict between reward and threat contingencies, and resulting conflict in approach and avoidance response tendencies, may elicit greater memory reactivation, with the specific reward or threat memory reactivated dependent upon trial-specific characteristics (e.g., acute motivation for reward vs threat avoidance, habituation to reward vs threat, fatigue, certainty regarding contingencies, exploration vs exploitation, etc.). Indeed, this interpretation is consistent with a prior study using magnetoencephalography (MEG) during an aversive learning task that also did not find trialwise evidence for reactivation of outcome states during the choice phase; rather, they found that only on trials where participants were using a model-based (i.e., prospective) learning strategy were memory patterns associated with outcome states activated at the time of choice (Wise et al., 2021). Consistent with this MEG study, in the current data the approach-avoidance conflict likely elicits greater use of model-based prospective strategies and thereby the formation of a prospective mental representation. A causal role of the memory reactivation in driving behavior vs merely reflecting the imagined outcome of an already chosen behavior or the maintenance of a specific part of the cognitive cannot be determined in these data. Nonetheless, these data demonstrate the engagement of specific mental representations for reward vs threat that are predictive of behavior during approach avoidance conflict decision-making and implicate these representations as one possible route of resolving approach-avoidance conflict.

The current data shed light on shared vs unique neurocircuitry encoding of reward vs threat, frontoparietal network activity during approach-avoidance conflict, and mental representations for reward vs threat during approach-avoidance conflict decision-making. While this study did not include a clinical sample, results from this study may be informative for future studies among clinical samples for which approach-avoidance conflict may be relevant (e.g., anxiety disorders, substance use, bipolar spectrum disorders, etc.). For example, results from the current study might lead to the hypothesis that a population characterized by heightened avoidance at the expense of reward (e.g., PTSD) may demonstrate heightened threat memory reactivation in the inferior frontotemporal

network at the time of choice relative to controls, and that this bias in memory reactivation patterns might mediate heightened decisions to avoid. Further investigations along these lines are warranted. Despite the possible implications of the current results, this initial study is not without limitation. First, the relative salience and motivational value of the reward vs threat stimuli are likely not matched in the task, and possibly individually varying. This could result in any comparison of reward vs threat encoding being confounded with salience and/or motivational value. This could also affect the resulting conflict created by competing reward and threat expectation. Second, and relatedly, the reward and threat stimuli used (points as a proxy for subsequent monetary compensation and electro tactile stimulation) may not necessarily generalize to other domains of reward (e.g., social reward) or threat (e.g., trauma-related stimuli). Third, we did not have an independent localizer task for the reward and threat decoders, and instead they were trained from the outcome phases of the task. While we attempted to prevent any ‘double dipping’ of data being used both for training and test by first training the classifiers on a unique set of participants then testing the classifier on an independent participant and stage and the task, it would nonetheless have been stronger to use a localizer task to specifically and separately characterize neurocircuitry patterns for reward and threat. It would additionally have been stronger to have a larger sample of participants and have been able to have a true training set of participants and completely independent set of test participants, which would more stringently test generalizability of the models. Finally, an additional consideration is the fixed order of the threat outcomes and then the reward outcomes in our task paradigm. While it was necessary to fix the order to allow the participant to anticipate the threat and reward outcomes, respectively, it nonetheless raises the possibility that the order here impacted results. For example, perhaps the initial processing of threat impacted the subsequent processing of reward. Without having counterbalanced the order, we cannot rule out the possibility. We jittered the duration of these events in order to facilitate statistical separation of them in our models, but this would not address a carryover effect of threat impacting the subsequent processing of reward. Additional iterations of this paradigm in future research might consider flipping the order, or perhaps alternating the order across runs after explicitly informing the participant to continue to allow anticipation. Future research should address these limitations and continue to probe the computational neurocircuitry and cognitive mechanisms of decision-making for reward, threat, and approach-avoidance conflict.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

KMC is supported by a *National Institute on Alcohol Abuse and Alcoholism* (NIAAA/NIH) training grant (T32AA007471). JMC is supported by MH119132, MH108753, and the Institute for Early Life Adversity Research at the University of Texas at Austin.

Data availability

I have shared my data and code, please see attached data statement

Abbreviations:

AAC	Approach-Avoidance Conflict
AACT	Approach Avoidance Conflict Task
ICA	Independent Component Analyses
MVPA	Multivariate Pattern Analysis
PFC	prefrontal cortex
dACC	dorsal anterior cingulate cortex
RW	Rescorla-Wagner
R model	reward only model
RT single model	reward and threat model using a single learning rate
RT double model	reward and threat model using separate learning rates for reward and threat

References

- Alexander WH, Brown JW, 2011. Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci* 14 (10), 1338. [PubMed: 21926982]
- Alloy LB, Abramson LY, Walshaw PD, Cogswell A, Smith JM, Neeren AM, Hughes ME, Iacoviello BM, Gerstein RK, Keyser J, Urosevic S, Nusslock R, 2006. Behavioral Approach System (BAS) sensitivity and bipolar spectrum disorders: a retrospective and concurrent behavioral high-risk design. *Motiv. Emot* 30 (2), 143–155. doi: 10.1007/s11031-006-9003-3.
- Alloy LB, Abramson LY, Walshaw PD, Gerstein RK, Keyser JD, Whitehouse WG, Urosevic S, Nusslock R, Hogan ME, Harmon-Jones E, 2009. Behavioral approach system (BAS)-relevant cognitive styles and bipolar spectrum disorders: concurrent and prospective associations. *J. Abnorm. Psychol* 118 (3), 459–471. doi: 10.1037/a0016604. [PubMed: 19685944]
- Aupperle RL, Melrose AJ, Francisco A, Paulus MP, Stein MB, 2015. Neural substrates of approach-avoidance conflict decision-making. *Hum. Brain Mapp* 36 (2), 449–462. doi: 10.1002/hbm.22639. [PubMed: 25224633]
- Avena-Koenigsberger A, Masic B, Sporns O, 2018. Communication dynamics in complex brain networks. *Nature Rev. Neurosci* 19 (1), 17–33. doi: 10.1038/nrn.2017.149.
- Bach DR, Guitart-Masip M, Packard PA, Miró J, Falip M, Fuentemilla L, Dolan RJ, 2014. Human hippocampus arbitrates approach-avoidance conflict. *Curr. Biol.: CB* 24 (5), 541–547. doi: 10.1016/j.cub.2014.01.046. [PubMed: 24560572]
- Behrens TEJ, Hunt LT, Woolrich MW, Rushworth MFS, 2008. Associative learning of social value. *Nature* 456 (7219), 245–249. doi: 10.1038/nature07538. [PubMed: 19005555]
- Behrens TEJ, Muller TH, Whittington JCR, Mark S, Baram AB, Stachenfeld KL, Kurth-Nelson Z, 2018. What is a cognitive map? Organizing knowledge for flexible behavior. *Neuron* 100 (2), 490–509. doi: 10.1016/j.neuron.2018.10.002. [PubMed: 30359611]
- Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS, 2007. Learning the value of information in an uncertain world. *Nat. Neurosci* 10 (9), 1214–1221. doi: 10.1038/nn1954. [PubMed: 17676057]
- Biderman N, Bakkour A, Shohamy D, 2020. What are memories for? The hippocampus bridges past experience with future decisions. *Trends Cogn. Sci. (Regul. Ed.)* 24 (7), 542–556. doi: 10.1016/j.tics.2020.04.004.

- Bressler SL, Menon V, 2010. Large-scale brain networks in cognition: emerging methods and principles. *Trends Cogn. Sci. (Regul. Ed.)* 14 (6), 277–290. doi: 10.1016/j.tics.2010.04.004.
- Brown JW, Braver TS, 2005. Learned predictions of error likelihood in the anterior cingulate cortex. *Science* 307 (5712), 1118–1121. [PubMed: 15718473]
- Bullmore E, Sporns O, 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Rev. Neurosci* 10 (3), 186–198. doi: 10.1038/nrn2575. [PubMed: 19190637]
- Bullmore E, Sporns O, 2012. The economy of brain network organization. *Nature Reviews Neuroscience* 13 (5), 336–349. doi: 10.1038/nrn3214. [PubMed: 22498897]
- Calhoun VD, Adali T, Hansen LK, Larsen J, & Pekar JJ (2003). ICA of functional MRI data: an overview. http://orbit.dtu.dk/fedora/objects/orbit:50430/datastreams/file_2787076/content
- Calhoun VD, Adali T, Pearlson GD, Pekar JJ, 2001. A method for making group inferences from functional MRI data using independent component analysis. *Hum Brain Mapp* 14 (3), 140–151. [PubMed: 11559959]
- Calhoun VD, de Lacy N, 2017. Ten Key Observations on the Analysis of Resting-state Functional MR Imaging Data Using Independent Component Analysis. *Neuroimaging Clin. N. Am* 27 (4), 561–579. doi: 10.1016/j.nic.2017.06.012. [PubMed: 28985929]
- Carroll JM, Yik MSM, Russell JA, Barrett LF, 1999. On the Psychometric Principles of Affect. *Review of General Psychology* 3 (1), 14–22. doi: 10.1037/1089-2680.3.1.14.
- Castegnetti G, Tzovara A, Khemka S, Melin š ak F, Barnes GR, Dolan RJ, Bach DR, 2020. Representation of probabilistic outcomes during risky decision-making. *Nat Commun* 11 (1), 2419. doi: 10.1038/s41467-020-16202-y. [PubMed: 32415145]
- Chang C–C, Lin C–J, 2011. LIBSVM: a Library for Support Vector Machines. *ACM Trans. Intell. Syst. Technol* 2 (3). doi: 10.1145/1961189.1961199, 27:1–27:27.
- Cisler JM, Esbensen K, Sellnow K, Ross M, Weaver S, Sartin-Tarm A, Herringa RJ, Kilts CD, 2019. Differential roles of the salience network during prediction error encoding and facial emotion processing among female adolescent assault victims. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 4 (4), 371–380. [PubMed: 30343131]
- Cisler JM, Privratsky AA, Sartin-Tarm A, Sellnow K, Ross M, Weaver S, Hahn E, Herringa RJ, James GA, Kilts CD, 2020. L-DOPA and consolidation of fear extinction learning among women with posttraumatic stress disorder. *Transl Psychiatry* 10 (1), 1–11. [PubMed: 32066695]
- Cornwell BR, Garrido MI, Overstreet C, Pine DS, Grillon C, 2017. The Unpredictive Brain Under Threat: a Neurocomputational Account of Anxious Hypervigilance. *Biol. Psychiatry* 82 (6), 447–454. doi: 10.1016/j.biopsych.2017.06.031. [PubMed: 28838469]
- Cox RW, Chen G, Glen DR, Reynolds RC, Taylor PA, 2017. fMRI clustering and false-positive rates. *Proceedings of the National Academy of Sciences* 114 (17), E3370–E3371. doi: 10.1073/pnas.1614961114.
- Dasgupta I, Gershman SJ, 2021. Memory as a Computational Resource. *Trends Cogn. Sci. (Regul. Ed.)* 25 (3), 240–251. doi: 10.1016/j.tics.2020.12.008.
- Daw ND, 2018. Are we of two minds? *Nat. Neurosci* 21 (11), 1497–1499. doi: 10.1038/s41593-018-0258-2. [PubMed: 30349102]
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ, 2011. Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69 (6), 1204–1215. [PubMed: 21435563]
- Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ, 2006. Cortical substrates for exploratory decisions in humans. *Nature* 441 (7095), 876–879. doi: 10.1038/nature04766. [PubMed: 16778890]
- Delgado MR, Li J, Schiller D, Phelps EA, 2008. The role of the striatum in aversive learning and aversive prediction errors. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363 (1511), 3787–3800. doi: 10.1098/rstb.2008.0161. [PubMed: 18829426]
- Doll BB, Duncan KD, Simon DA, Shohamy D, Daw ND, 2015. Model-based choices involve prospective neural activity. *Nat. Neurosci* 18 (5), 767–772. doi: 10.1038/nn.3981. [PubMed: 25799041]
- Drabant EM, Kuo JR, Ramel W, Blechert J, Edge MD, Cooper JR, Goldin PR, Hariri AR, Gross JJ, 2011. Experiential, autonomic, and neural responses during threat anticipation vary

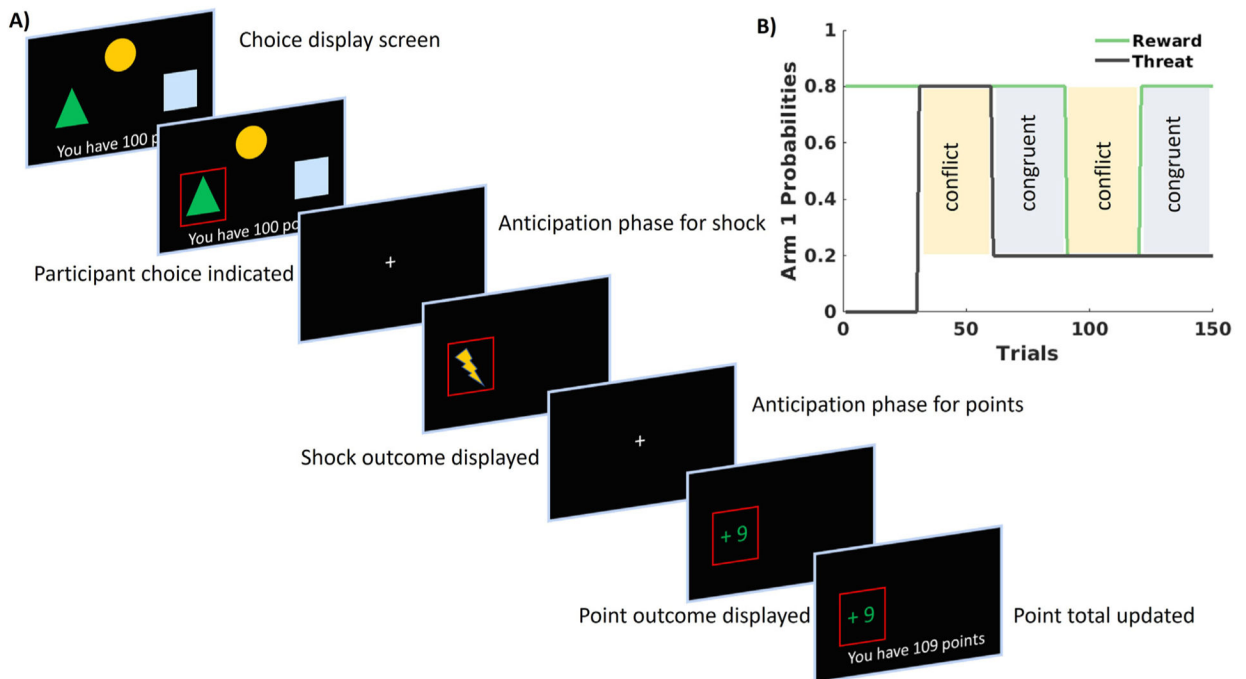
as a function of threat intensity and neuroticism. *Neuroimage* 55 (1), 401–410. doi: 10.1016/j.neuroimage.2010.11.040. [PubMed: 21093595]

- Eklund A, Nichols TE, Knutsson H, 2016. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences* 113 (28), 7900–7905. doi: 10.1073/pnas.1602413113.
- Feher da Silva C, Hare TA, 2020. Humans primarily use model-based inference in the two-stage task. *Nature Human Behaviour* 1–14. doi: 10.1038/s41562-020-0905-y.
- Fleming KA, Bartholow BD, 2014. Alcohol cues, approach bias, and inhibitory control: applying a dual process model of addiction to alcohol sensitivity. *Psychol. Addict. Behav* 28 (1), 85–96. [PubMed: 23438245]
- Fouragnan E, Retzler C, Philiastides MG, 2018. Separate neural representations of prediction error valence and surprise: evidence from an fMRI meta-analysis. *Hum Brain Mapp* 39 (7), 2887–2906. doi: 10.1002/hbm.24047. [PubMed: 29575249]
- Fullana MA, Harrison BJ, Soriano-Mas C, Vervliet B, Cardoner N, Àvila-Parcet A, Radua J, 2016. Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Mol. Psychiatry* 21 (4), 500–508. [PubMed: 26122585]
- Garrison J, Erdeniz B, Done J, 2013. Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev* 37 (7), 1297–1310. doi: 10.1016/j.neubiorev.2013.03.023. [PubMed: 23567522]
- Geng H, Wang Y, Gu R, Luo Y–J, Xu P, Huang Y, Li X, 2018. Altered brain activation and connectivity during anticipation of uncertain threat in trait anxiety. *Hum Brain Mapp* 39 (10), 3898–3914. doi: 10.1002/hbm.24219. [PubMed: 29882617]
- Gershman SJ, Daw ND, 2017. Reinforcement Learning and Episodic Memory in Humans and Animals: an Integrative Framework. *Annu Rev Psychol* 68 (1), 101–128. doi: 10.1146/annurev-psych-122414-033625. [PubMed: 27618944]
- Gilbert DT, Wilson TD, 2007. Propection: experiencing the future. *Science* 317 (5843), 1351–1354. doi: 10.1126/science.1144161. [PubMed: 17823345]
- Gillespie AK, Astudillo Maya DA, Denovellis EL, Liu DF, Kastner DB, Coulter ME, Roumis DK, Eden UT, Frank LM, 2021. Hippocampal replay reflects specific past experiences rather than a plan for subsequent choice. *Neuron* 109 (19), 3149–3163. doi: 10.1016/j.neuron.2021.07.029.e6. [PubMed: 34450026]
- 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 (4), 585–595. [PubMed: 20510862]
- Gluth S, Sommer T, Rieskamp J, Büchel C, 2015. Effective Connectivity between Hippocampus and Ventromedial Prefrontal Cortex Controls Preferential Choices from Memory. *Neuron* 86 (4), 1078–1090. doi: 10.1016/j.neuron.2015.04.023. [PubMed: 25996135]
- Hirshfeld-Becker DR, Biederman J, Calltharp S, Rosenbaum ED, Faraone SV, Rosenbaum JF, 2003. Behavioral inhibition and disinhibition as hypothesized precursors to psychopathology: implications for pediatric bipolar disorder. *Biol. Psychiatry* 53 (11), 985–999. doi: 10.1016/S0006-3223(03)00316-0. [PubMed: 12788244]
- Homan P, Levy I, Feltham E, Gordon C, Hu J, Li J, Pietrzak RH, South-wick S, Krystal JH, Harpaz-Rotem I, Schiller D, 2019. Neural computations of threat in the aftermath of combat trauma. *Nat. Neurosci* 22 (3), 470–476. doi: 10.1038/s41593-018-0315-x. [PubMed: 30664770]
- Ironside M, Amemori K–I, McGrath CL, Pedersen ML, Kang MS, Amemori S, Frank MJ, Graybiel AM, Pizzagalli DA, 2020. Approach-Avoidance Conflict in Major Depressive Disorder: congruent Neural Findings in Humans and Nonhuman Primates. *Biol. Psychiatry* 87 (5), 399–408. doi: 10.1016/j.biopsych.2019.08.022. [PubMed: 31672243]
- Kalisch R, Gerlicher AMV, Duvarci S, 2019. A Dopaminergic Basis for Fear Extinction. *Trends Cogn. Sci. (Regul. Ed.)* 23 (4), 274–277. doi: 10.1016/j.tics.2019.01.013.
- Khemka S, Barnes G, Dolan RJ, Bach DR, 2017. Dissecting the Function of Hippocampal Oscillations in a Human Anxiety Model. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 37 (29), 6869–6876. doi: 10.1523/JNEUROSCI.1834-16.2017. [PubMed: 28626018]

- Kirlic N, Young J, Aupperle RL, 2017. Animal to human translational paradigms relevant for approach avoidance conflict decision making. *Behav Res Ther* 96, 14–29. [PubMed: 28495358]
- Kuppens P, Tuerlinckx F, Russell JA, Barrett LF, 2013. The relation between valence and arousal in subjective experience. *Psychol Bull* 139 (4), 917–940. doi: 10.1037/a0030811. [PubMed: 23231533]
- Leong YC, Radulescu A, Daniel R, DeWoskin V, Niv Y, 2017. Dynamic interaction between reinforcement learning and attention in multidimensional environments. *Neuron* 93 (2), 451–463. [PubMed: 28103483]
- Letkiewicz AM, Cochran AL, Cisler JM, 2020. Frontoparietal network activity during model-based reinforcement learning updates is reduced among adolescents with severe sexual abuse. *J Psychiatr Res.*
- Letkiewicz AM, Cochran AL, Privratsky AA, James GA, Cisler JM, 2022. Value estimation and latent-state update-related neural activity during fear conditioning predict posttraumatic stress disorder symptom severity. *Cognitive, Affective, & Behavioral Neuroscience* 22 (1), 199–213. doi: 10.3758/s13415-021-00943-4.
- Li J, Schiller D, Schoenbaum G, Phelps EA, Daw ND, 2011. Differential roles of human striatum and amygdala in associative learning. *Nat. Neurosci* 14 (10), 1250–1252. [PubMed: 21909088]
- Lindström B, Haaker J, Olsson A, 2018. A common neural network differentially mediates direct and social fear learning. *Neuroimage* 167, 121–129. doi: 10.1016/j.neuroimage.2017.11.039. [PubMed: 29170069]
- Martin Braunstein L, Kuerbis A, Ochsner K, Morgenstern J, 2016. Implicit Alcohol Approach and Avoidance Tendencies Predict Future Drinking in Problem Drinkers. *Alcohol Clin. Exp. Res* 40 (9), 1945–1952. [PubMed: 27421061]
- Menon V, 2011. Large-scale brain networks and psychopathology: a unifying triple network model. *Trends Cogn. Sci. (Regul. Ed.)* 15 (10), 483–506. doi: 10.1016/j.tics.2011.08.003.
- Metereau E, Dreher J-C, 2013. Cerebral Correlates of Salient Prediction Error for Different Rewards and Punishments. *Cerebral Cortex* 23 (2), 477–487. doi: 10.1093/cercor/bhs037. [PubMed: 22368086]
- Meunier D, Lambiotte R, Bullmore ET, 2010. Modular and Hierarchically Modular Organization of Brain Networks. *Front Neurosci* 4. doi: 10.3389/fnins.2010.00200.
- Miller KJ, Venditto SJC, 2021. Multi-step planning in the brain. *Curr Opin Behav Sci* 38, 29–39. doi: 10.1016/j.cobeha.2020.07.003.
- Mızrak E, Bouffard NR, Libby LA, Boorman ED, Ranganath C, 2021. The hippocampus and orbitofrontal cortex jointly represent task structure during memory-guided decision making. *Cell Rep* 37 (9), 110065. doi: 10.1016/j.celrep.2021.110065. [PubMed: 34852232]
- Niv Y, Daniel R, Geana A, Gershman SJ, Leong YC, Radulescu A, Wilson RC, 2015. Reinforcement learning in multidimensional environments relies on attention mechanisms. *Journal of Neuroscience* 35 (21), 8145–8157. [PubMed: 26019331]
- Niv Y, Edlund JA, Dayan P, O’Doherty JP, 2012. Neural Prediction Errors Reveal a Risk-Sensitive Reinforcement-Learning Process in the Human Brain. *Journal of Neuroscience* 32 (2), 551–562. doi: 10.1523/JNEUROSCI.5498-10.2012. [PubMed: 22238090]
- O’Neil EB, Newsome RN, Li IHN, Thavabalasingam S, Ito R, Lee ACH, 2015. Examining the Role of the Human Hippocampus in Approach–Avoidance Decision Making Using a Novel Conflict Paradigm and Multivariate Functional Magnetic Resonance Imaging. *Journal of Neuroscience* 35 (45), 15039–15049. doi: 10.1523/JNEUROSCI.1915-15.2015. [PubMed: 26558775]
- Pessoa L, 2017. A Network Model of the Emotional Brain. *Trends Cogn. Sci. (Regul. Ed.)* 21 (5), 357–371. doi: 10.1016/j.tics.2017.03.002.
- Pessoa L, 2018. Understanding emotion with brain networks. *Curr Opin Behav Sci* 19, 19–25. doi: 10.1016/j.cobeha.2017.09.005. [PubMed: 29915794]
- Piercy H, Manning V, Staiger PK, 2021. Pushing or Pulling Your “Poison”: clinical Correlates of Alcohol Approach and Avoidance Bias Among Inpatients Undergoing Alcohol Withdrawal Treatment. *Front Psychol* 12, 663087. [PubMed: 34113294]

- Piray P, Dezfouli A, Heskes T, Frank MJ, Daw ND, 2019. Hierarchical Bayesian inference for concurrent model fitting and comparison for group studies. *PLoS Comput. Biol* 15 (6), e1007043. doi: 10.1371/journal.pcbi.1007043. [PubMed: 31211783]
- Radoman M, Lieberman L, Jimmy J, Gorka SM, 2021. Shared and unique neural circuitry underlying temporally unpredictable threat and reward processing. *Soc Cogn Affect Neurosci* 16 (4), 370–382. doi: 10.1093/scan/nsab006. [PubMed: 33449089]
- Rescorla RA, Wagner AR, 1972. A theory of Pavlovian conditioning. Variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF (Eds.), *Classical conditioning II: Current research and theory*. Appleton-Century-Crofts.
- Ross MC, Lenow JK, Kilts CD, Cisler JM, 2018. Altered neural encoding of prediction errors in assault-related posttraumatic stress disorder. *J Psychiatr Res* 103, 83–90. doi: 10.1016/j.jpsychires.2018.05.008. [PubMed: 29783079]
- Schacter DL, Benoit RG, Szpunar KK, 2017. Episodic Future Thinking: mechanisms and Functions. *Curr Opin Behav Sci* 17, 41–50. doi: 10.1016/j.cobeha.2017.06.002. [PubMed: 29130061]
- Schlund MW, Brewer AT, Magee SK, Richman DM, Solomon S, Ludlum M, Dymond S, 2016. The tipping point: value differences and parallel dorsal–ventral frontal circuits gating human approach–avoidance behavior. *Neuroimage* 136, 94–105. doi: 10.1016/j.neuroimage.2016.04.070. [PubMed: 27153979]
- Schuck NW, Cai MB, Wilson RC, Niv Y, 2016. Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron* 91 (6), 1402–1412. doi: 10.1016/j.neuron.2016.08.019. [PubMed: 27657452]
- Shackman AJ, Salomons TV, Slagter HA, Fox AS, Winter JJ, Davidson RJ, 2011. The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Rev. Neurosci* 12 (3), 154–167. doi: 10.1038/nrn2994. [PubMed: 21331082]
- Shadlen MN, Shohamy D, 2016. Decision making and sequential sampling from memory. *Neuron* 90 (5), 927–939. doi: 10.1016/j.neuron.2016.04.036. [PubMed: 27253447]
- Shenhav A, Straccia MA, Cohen JD, Botvinick MM, 2014. Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nat. Neurosci* 17 (9), 1249–1254. doi: 10.1038/nn.3771. [PubMed: 25064851]
- Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, Filippini N, Watkins KE, Toro R, Laird AR, Beckmann CF, 2009. Correspondence of the brain’s functional architecture during activation and rest. *Proc. Natl. Acad. Sci* 106 (31), 13040–13045. doi: 10.1073/pnas.0905267106. [PubMed: 19620724]
- Stormark KM, Field NP, Hugdahl K, Horowitz M, 1997. Selective processing of visual alcohol cues in abstinent alcoholics: an approach-avoidance conflict? *Addict. Behav* 22 (4), 509–519. [PubMed: 9290860]
- Sutton RS, Barto AG, 2018. *Reinforcement Learning: An Introduction*. MIT press.
- Tolin DF, Gilliam C, Wootton BM, Bowe W, Bragdon LB, Davis E, Hannan SE, Steinman SA, Worden B, Hallion LS, 2018. Psychometric properties of a structured diagnostic interview for DSM-5 anxiety, mood, and obsessive-compulsive and related disorders. *Assessment* 25 (1), 3–13. doi: 10.1177/1073191116638410. [PubMed: 26988404]
- Tom SM, Fox CR, Trepel C, Poldrack RA, 2007. The neural basis of loss aversion in decision-making under risk. *Science* 315 (5811), 515–518. [PubMed: 17255512]
- Weaver SS, Kroska EB, Ross MC, Sartin-Tarm A, Sellnow KA, Schaumberg K, Kiehl KA, Koenigs M, Cisler JM, 2020. Sacrificing reward to avoid threat: characterizing PTSD in the context of a trauma-related approach–avoidance conflict task. *J. Abnorm. Psychol* 129 (5), 457. [PubMed: 32437204]
- Wikenheiser AM, Schoenbaum G, 2016. Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. *Nat. Rev. Neurosci* 17 (8), 513–523. doi: 10.1038/nrn.2016.56. [PubMed: 27256552]
- Wise T, Liu Y, Chowdhury F, Dolan RJ, 2021. Model-based aversive learning in humans is supported by preferential task state reactivation. *Sci. Adv* 7 (31), eabf9616. doi: 10.1126/sciadv.abf9616. [PubMed: 34321205]

- Yu JY, Frank LM, 2015. Hippocampal–cortical interaction in decision making. *Neurobiol. Learn. Mem* 117, 34–41. doi: 10.1016/j.nlm.2014.02.002. [PubMed: 24530374]
- Zhou F, Zhao W, Qi Z, Geng Y, Yao S, Kendrick KM, Wager TD, Becker B, 2021. A distributed fMRI-based signature for the subjective experience of fear. *Nat. Commun* 12 (1), 6643. doi: 10.1038/s41467-021-26977-3. [PubMed: 34789745]
- Zielinski MC, Tang W, Jadhav SP, 2020. The role of replay and theta sequences in mediating hippocampal-prefrontal interactions for memory and cognition. *Hippocampus* 30 (1), 60–72. doi: 10.1002/hipo.22821. [PubMed: 29251801]
- Zorowitz S, Rockhill AP, Ellard KK, Link KE, Herrington T, Pizzagalli DA, Widge AS, Deckersbach T, Dougherty DD, 2019. The neural basis of approach-avoidance conflict: a model based analysis. *eNeuro* 6 (4). doi: 10.1523/ENEURO.0115-19.2019, ENEURO.0115–19.2019.

**Fig. 1.**

A) Depiction of the approach-avoidance task. Each trial began with each arm, denoted with a geometric shape, presented. After participants selected an arm, a fixation period was presented in anticipation of the delivery of the shock outcome, then another fixation period was presented in anticipation of the delivery of the point outcome, and finally the point total tracking their performance was updated. B) depiction of the reward and threat probabilities across the task for 1 arm. The first 30 trials did not contain any probability for shock presentation and served as a baseline period. Probabilities then alternated every 30 trials to create conflict (i.e., the high threat arm was also the high reward arm) and congruent (i.e., the high reward arm was the low threat arm) blocks. One arm was always 50% probability for both reward or shock (except for first 30 trials when it was set to 0) and the other arm was a reverse mirror to the arm presented in this figure, such that the probabilities of reward and shock for those two arms always summed to 1.

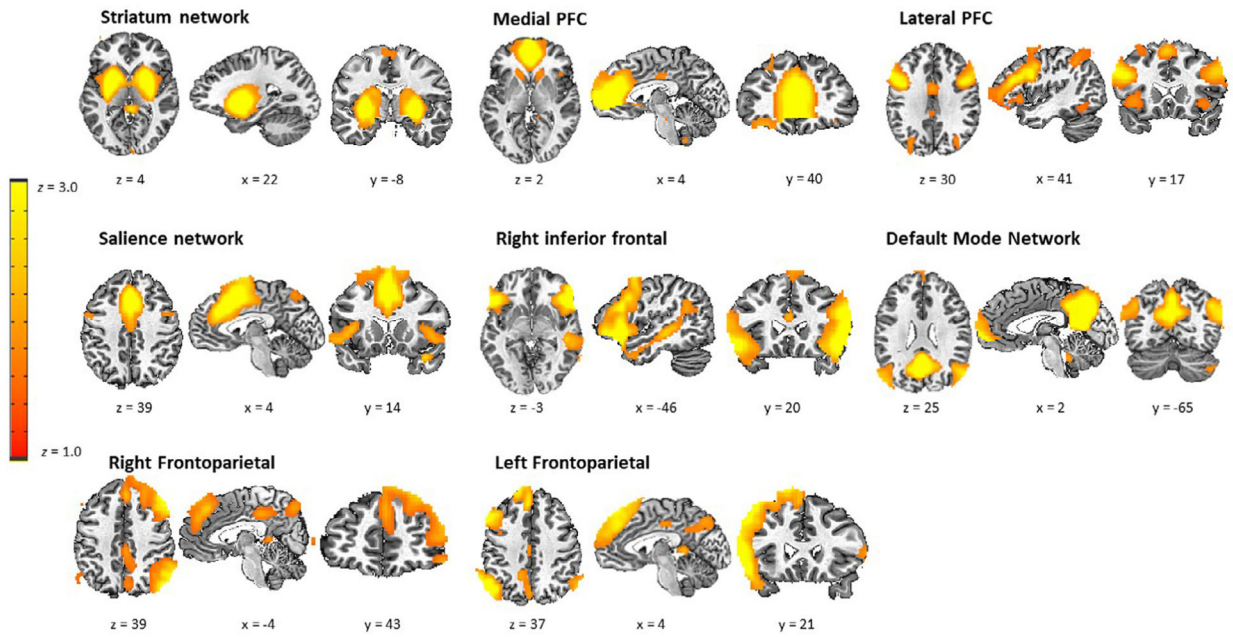


Fig. 2. Depictions of the eight ICA networks of interest used in subsequent analyses.

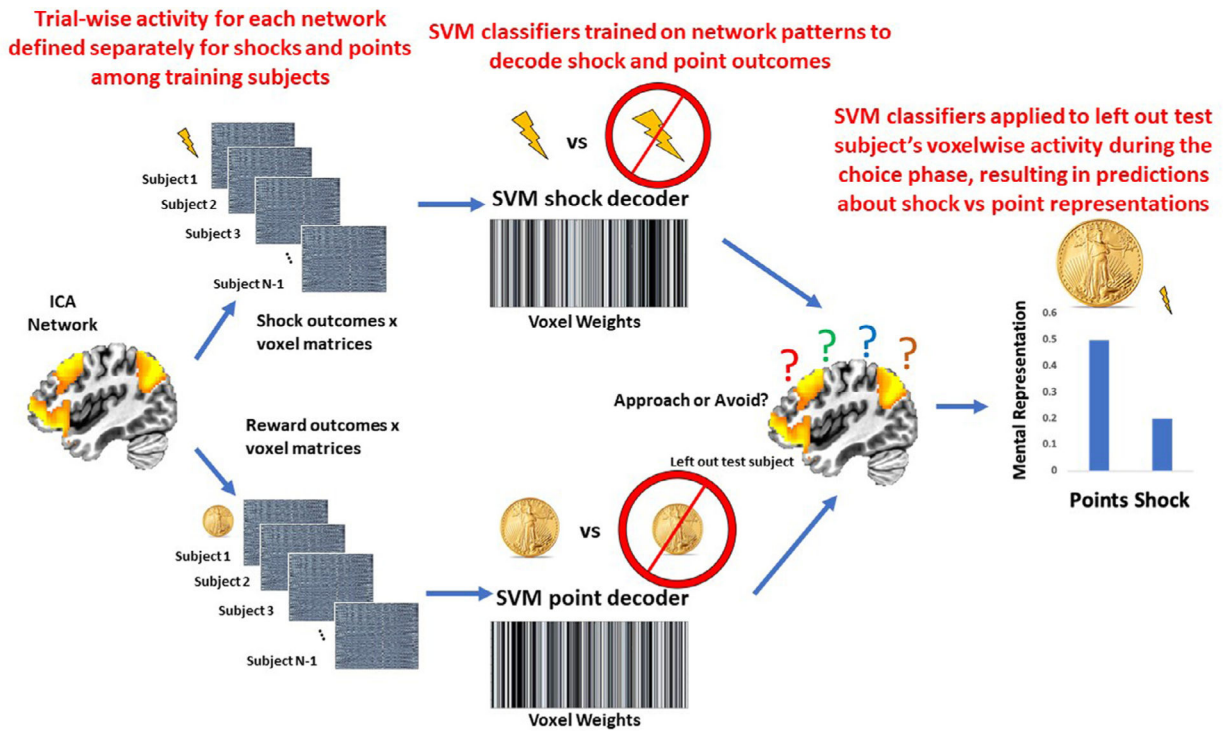


Fig. 3.

Graphical overview of the multivariate pattern analyses testing whether reward or threat representations at the time of choice predict subsequent decisions to approach or avoid. For each ICA network separately, trial \times voxel matrices of beta coefficients are created for all participants except one left out participant separately for point outcomes and shock outcomes during the task. Support vector machine classifiers are then trained on these data, resulting in a separate decoder for points and for shocks. Next, these point and shock decoders are applied to the trial \times voxel matrix of beta coefficients at the time choice for the participant that was left out of the training. This results in a prediction about the degree to which the point and shock representations are active at the time of choice, which can be compared to the participant's actual choice. This process is repeated until each participant has served as the left out test participant.

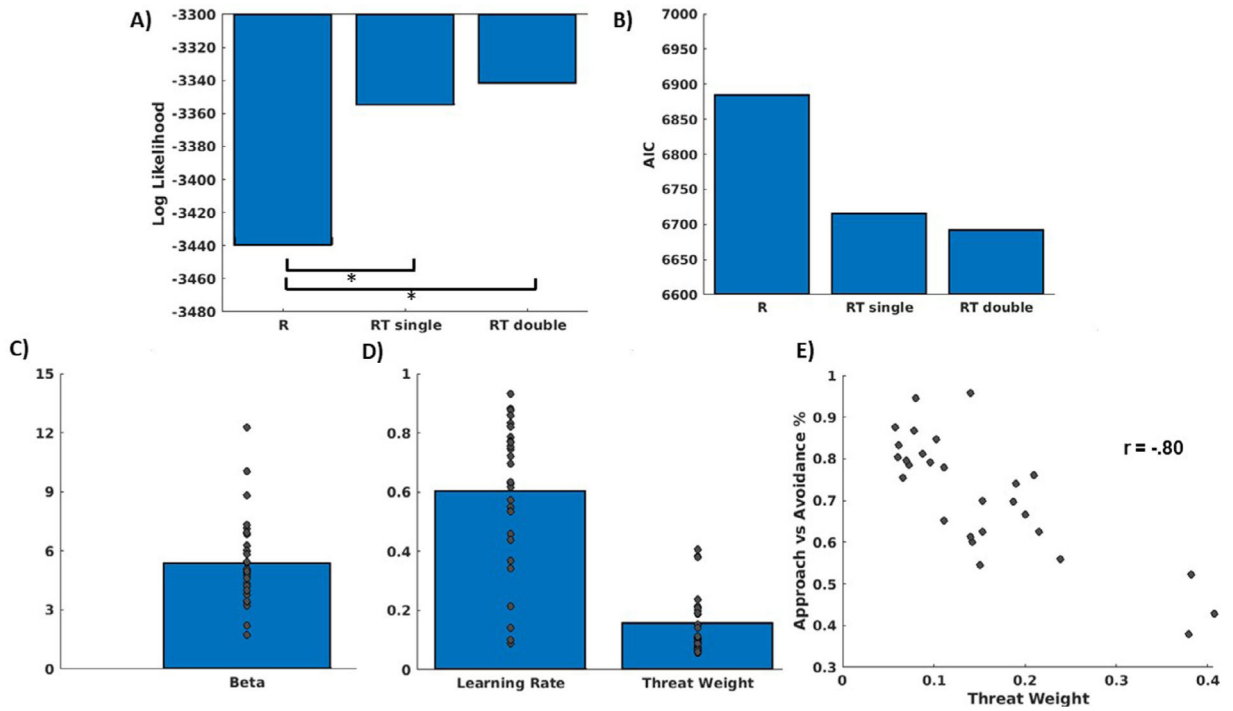


Fig. 4.

Computational modeling of behavior results. A) Log Likelihood comparisons between the three models, indicating better fit for the RT single model. B) Akaike Information Criterion (AIC) comparisons between three models, indicating better fit for the RT single model. Mean and distribution of the parameters (softmax beta, C, and learning rate and threat weight, D) from the best fitting model from A and B, the reward-threat single alpha model (RT single). E) scatter plot demonstrating strong inverse relationship between threat weight and percentage of high conflict trials where the individual chose to approach.

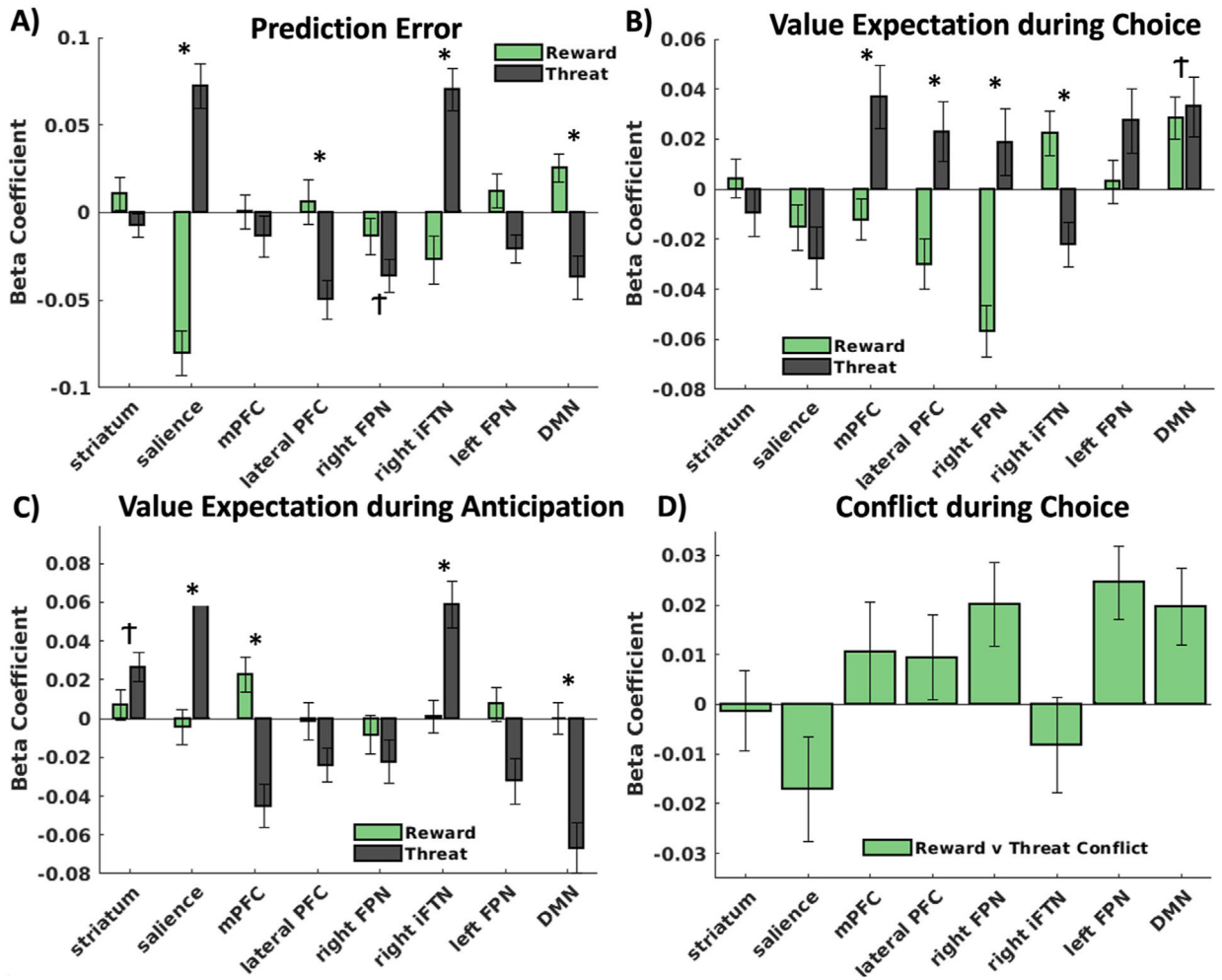


Fig. 5. Results of ICA network encoding of the computational model parameters for prediction error (A), value expectation during choice (B), value expectation during the anticipation of the outcome (C), and reward-threat conflict (D). * denotes significant differences between reward and threat encoding (corrected for multiple comparisons). † denotes significant overall encoding of the computational parameter (regardless of reward vs threat).

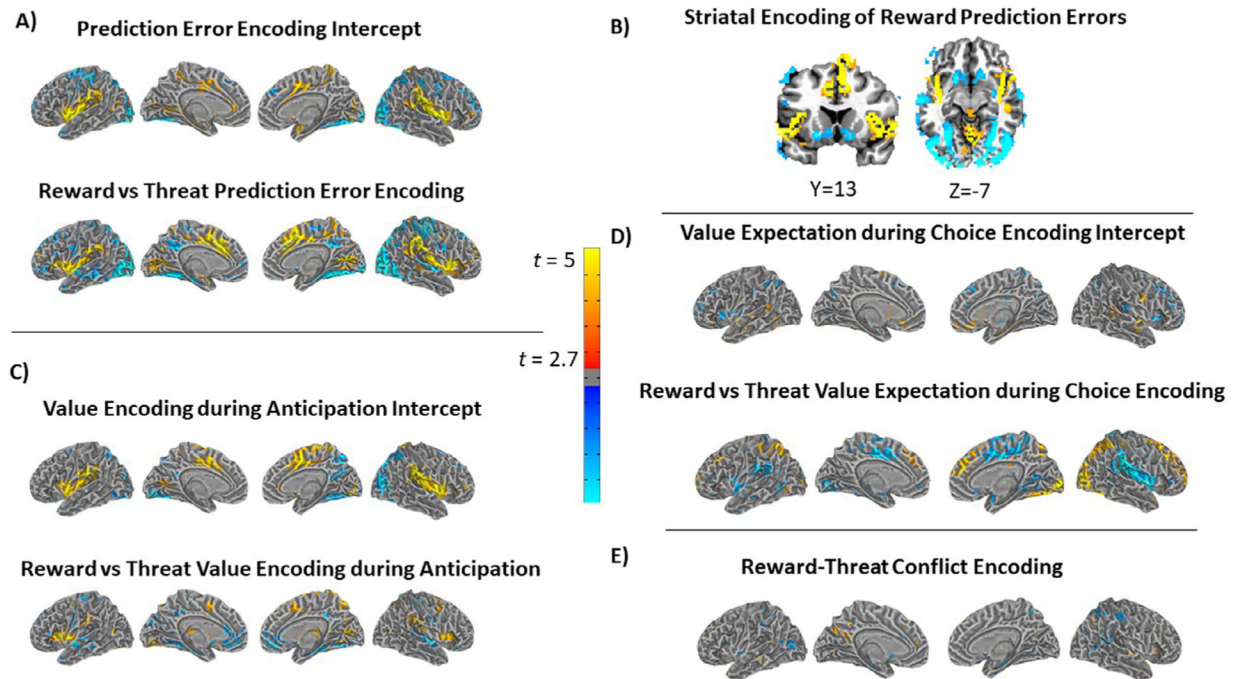


Fig. 6. Results from voxelwise analyses testing for encoding of the computational model parameters for prediction error (A and B), value expectation during anticipation (C), value expectation during choice (D), and reward-threat conflict.

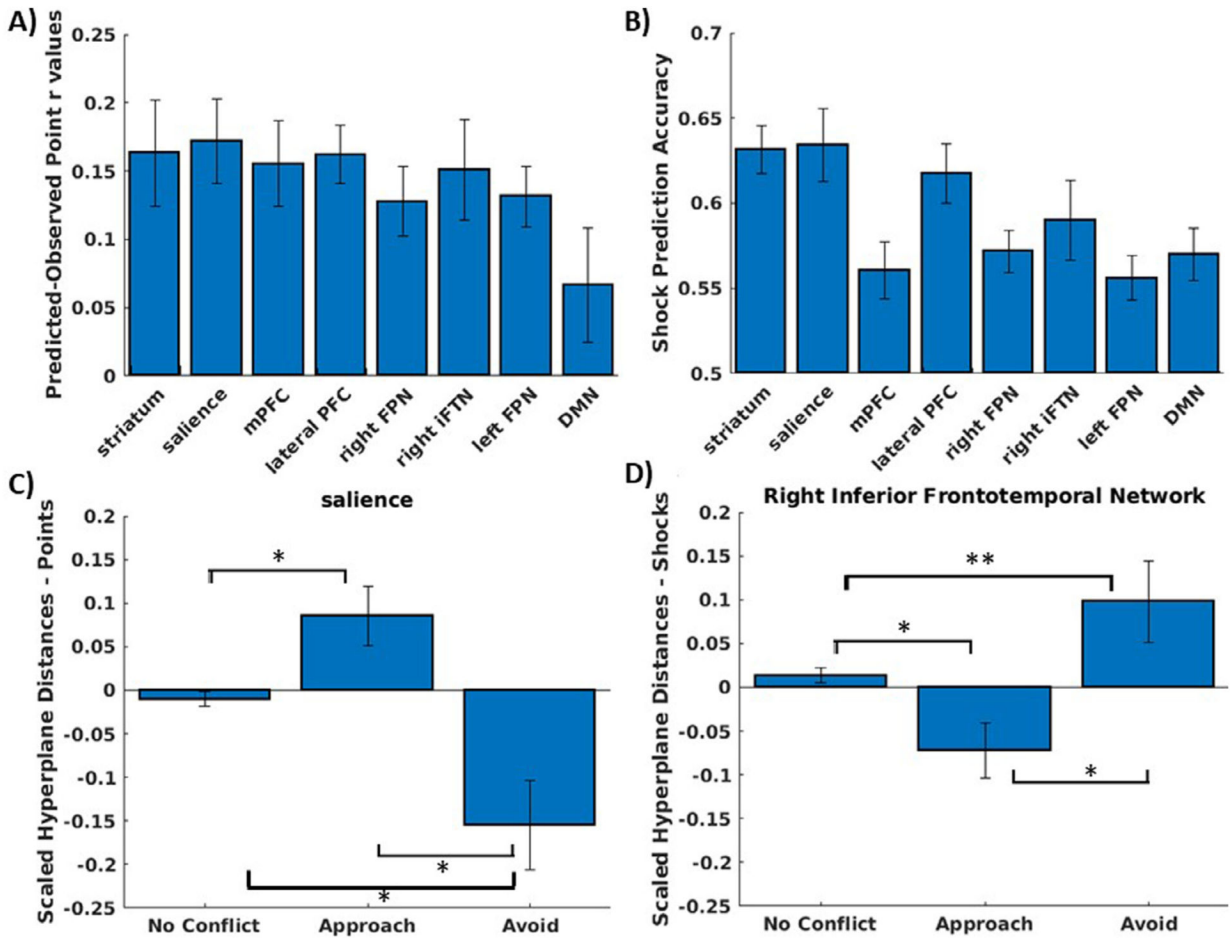


Fig. 7. A) mean and range of reward decoder accuracy for each ICA network, defined as correlation between predicted and observed points. B) mean and range of shock decoder accuracy for each ICA network, defined as the mean of specificity and sensitivity. C) z-scored hyperplane distances for the point decoder in the salience network, separated by no conflict trials, high conflict trials where the individual chose to approach, and high conflict trials where the individual chose to avoid. D) z-scored hyperplane distances for the shock decoder in the inferior frontotemporal network, separated by no conflict trials, high conflict trials where the individual chose to approach, and high conflict trials where the individual chose to avoid.

Table 1

Linear Mixed Effects Model Results for each ICA network and task phase.

Network	Task phase	Variable	t value	p value
Striatum				
Prediction Errors during Choice				
		Intercept	.2734	.78563
		Threat vs Reward Contrast	-1.9547	.056003
Value Expectation during Choice				
		Intercept	-0.45034	0.65433
		Threat vs Reward Contrast	-1.1853	.24128
Value Expectation during Anticipation				
		Intercept	3.1089	0.0030432
		Threat vs Reward Contrast	1.8868	0.064772
Reward-Threat Conflict during Choice				
		Intercept	-0.19689	0.8455
Saliency				
Prediction Errors during Choice				
		Intercept	-0.46717	0.64233
		Threat vs Reward Contrast	8.924	4.5508e-12
Value Expectation during Choice				
		Intercept	-2.6804	0.0098266
		Threat vs Reward Contrast	-0.87782	0.38408
Value Expectation during Anticipation				
		Intercept	4.6371	2.4165e-05
		Threat vs Reward Contrast	5.1991	3.4189e-06
Reward-Threat Conflict during Choice				
		Intercept	-1.9938	0.057189
Medial PFC				
Prediction Errors during Choice				
		Intercept	-0.89305	0.37594
		Threat vs Reward Contrast	-0.94191	0.35059

Network	Task phase	Variable	t value	p value	
Lateral PFC	Value Expectation during Choice	Intercept	1.5687	0.12278	
		Threat vs Reward Contrast	3.9056	0.00027248	
		Intercept	-1.5661	0.1234	
Lateral PFC	Value Expectation during Anticipation	Threat vs Reward Contrast	-5.2411	2.9472e-06	
		Intercept	1.0837	0.28884	
		Intercept			
Lateral PFC	Prediction Errors during Choice	Intercept	-2.6911	0.0095541	
		Threat vs Reward Contrast	-3.4262	0.0012033	
	Value Expectation during Choice	Intercept	-0.46279	0.64545	
		Threat vs Reward Contrast	3.6923	0.00053368	
	Value Expectation during Anticipation	Intercept	-1.973	0.05382	
		Threat vs Reward Contrast	-1.7447	0.08694	
Lateral PFC	Reward-Threat Conflict during Choice	Intercept	1.1479	0.2619	
		Intercept			
Right frontoparietal	Prediction Errors during Choice	Intercept	-3.5114	0.00093071	
		Threat vs Reward Contrast	-1.6919	0.096658	
	Value Expectation during Choice	Intercept	-2.346	0.022826	
		Threat vs Reward Contrast	4.6468	2.338e-05	
	Value Expectation during Anticipation	Intercept	-2.1612	0.035302	
		Threat vs Reward Contrast	-0.95914	0.34193	
	Lateral PFC	Reward-Threat Conflict during Choice	Intercept		
			Intercept		

Network	Task phase	Variable	t value	p value
Right inferior frontotemporal		Intercept	2.4703	0.020672
		Prediction Errors during Choice		
		Intercept	2.4721	0.016742
		Threat vs Reward Contrast	5.5416	1.0091e-06
Value Expectation during Choice		Intercept	-0.0045462	0.99639
		Threat vs Reward Contrast	-4.5818	2.9185e-05
Value Expectation during Anticipation		Intercept	4.1518	0.00012287
		Threat vs Reward Contrast	4.0385	0.00017774
Reward-Threat Conflict during Choice		Intercept	-0.90583	0.37367
		Left frontoparietal		
Prediction Errors during Choice		Intercept	-0.71031	0.48069
		Threat vs Reward Contrast	-2.7388	0.0084253
Value Expectation during Choice		Intercept	1.8514	0.069801
		Threat vs Reward Contrast	1.796	0.078303
Value Expectation during Anticipation		Intercept	-1.7527	0.085545
		Threat vs Reward Contrast	-2.8056	0.0070497
Reward-Threat Conflict during Choice		Intercept	3.6862	0.0011039
		Default mode		
Prediction Errors during Choice		Intercept	-0.80506	0.42445
		Threat vs Reward Contrast	-4.4273	4.9207e-05
Value Expectation during Choice		Intercept	4.3291	6.8373e-05

Network	Task phase	Variable	<i>t</i> value	<i>p</i> value
		Threat vs Reward Contrast	0.3161	0.75319
	Value Expectation during Anticipation			
		Intercept	-4.5255	3.5338e-05
		Threat vs Reward Contrast	-4.533	3.4446e-05
	Reward-Threat Conflict during Choice			
		Intercept	2.83	0.0090485

Note. Models also included covariates for head motion (fraction of TRs censored) and threat weight from the computational model. Bolded results indicate results that survived correction for multiple comparisons.