Grid-like entorhinal representation of an abstract value space during prospective decision making

Alexander Nitsch[1], Mona M. Garvert[1,2,3,4], Jacob L. S. Bellmund[1], Nicolas W. Schuck[2,3,5], Christian F. Doeller[1,6,7,8]

[1] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[2] Max Planck Research Group NeuroCode, Max Planck Institute for Human Development, Berlin, Germany

[3] Max Planck UCL Centre for Computational Psychiatry and Aging Research, Berlin, Germany

[4] Faculty of Human Sciences, Julius-Maximilians-Universität Würzburg, Würzburg, Germany

[5] Institute of Psychology, Universität Hamburg, Hamburg, Germany

[6] Kavli Institute for Systems Neuroscience, Centre for Neural Computation, The Egil and Pauline Braathen and Fred Kavli Centre for Cortical Microcircuits, Jebsen Centre for Alzheimer's Disease, Norwegian University of Science and Technology, Trondheim, Norway

[7] Wilhelm Wundt Institute for Psychology, Leipzig University, Leipzig, Germany

[8] Department of Psychology, Technical University Dresden, Dresden, Germany

Correspondence: nitsch@cbs.mpg.de or doeller@cbs.mpg.de

Abstract

Everyday decisions require us to predict how valuable different choice options will be in the future. Prior studies have identified a cognitive map in the hippocampal-entorhinal system that encodes relationships between states and enables prediction of future states, but does not inherently convey value during prospective decision making. Here, we investigated whether the entorhinal cortex integrates relational information about changing values by representing an abstract value space. To this end, we combined fMRI with a prospective decision making task that required participants to track and predict changing values of two choice options in a sequence. Such a sequence formed a trajectory through an underlying two-dimensional value space. Our results show that participants successfully integrated and extrapolated changes along the two value dimensions. Participants' choice behavior was explained by a prospective reinforcement learning model and the degree to which they updated values over time correlated with self-reported navigational abilities and preferences. Crucially, while participants traversed the abstract value space, the entorhinal cortex exhibited a grid-like representation, with the phase of the hexadirectional fMRI signal (i.e., the orientation of the estimated grid) being aligned to the most informative axis through the value space. A network of brain regions, including the ventromedial prefrontal cortex (vmPFC), tracked the prospective value difference between options and the occipital-temporal cortex represented the more valuable option. These findings suggest that the entorhinal grid system might support the

prediction of future values by representing a cognitive map, which might be used to generate lower-dimensional signals of the value difference between options and their identities for choices. Thus, these findings provide novel insight for our understanding of cognitive maps as a mechanism to guide prospective decision making in humans.

Introduction

Optimal decision making relies on predictions of future values associated with choice options. For example, if you were to invest in stocks, you would want to choose those stocks which are likely to be more valuable than others in the future. In particular, this implies that you should be able to predict if and when one stock becomes more valuable than another and choose accordingly to maximize long-term reward. Crucially, such prospective decision making requires an appropriate neural representation of the relation between changing and future values of choice options.

Prior studies established a role for parts of the ventromedial prefrontal cortex (vmPFC) and orbitofrontal cortex (OFC) as well as the ventral striatum in tracking the value difference between the chosen and the unchosen option during decision making (Bartra et al., 2013; Boorman et al., 2009; De Martino et al., 2013; FitzGerald et al., 2009; Hunt et al., 2012; Knutson et al., 2005; Levy & Glimcher, 2012; O'Doherty et al., 2001; Padoa-Schioppa & Assad, 2006; Pelletier & Fellows, 2019; Plassmann et al., 2007). Correct decisions in many previously used tasks depended primarily on updating values based on experience (Rescorla & Wagner, 1972). However, many decisions, such as in the introductory example of the stock market, require recognizing trends and extrapolating values into the future. In such scenarios, dorsal anterior cingulate cortex (dACC) has been implicated in comparing recent and past reward rates, allowing for trend-guided choices based on expected future rewards (Kolling et al., 2016; Wittmann et al., 2016).

Prediction of future values is enabled by an internal model, which represents transitions between states and reward contingencies in an environment or task. Reliance on an internal model has been referred to as model-based decision making, and can lead to distinct value computations found in the dorsomedial prefrontal cortex (dmPFC) (Daw et al., 2005, 2011; Doll et al., 2015). Moreover, the hippocampus has been implicated in model-based and value-based decision making (Bornstein & Daw, 2013; Gershman & Daw, 2017; Palombo et al., 2019; Schuck & Niv, 2019; Vikbladh et al., 2019; Wikenheiser & Schoenbaum, 2016; Wimmer & Shohamy, 2012). Interestingly, Vikbladh et al. (2019) found that the hippocampus serves as a common neural substrate for both model-based decision making and place memory in spatial navigation. A possible mechanism by which the hippocampus could support both model-based decision-making and spatial navigation is via the formation of cognitive maps.

Cognitive maps encode relationships between states in the world in a map-like format (Behrens et al., 2018; Bellmund et al., 2018; Epstein et al., 2017; O'Keefe & Nadel, 1978; Schuck et al., 2016; Stachenfeld et al., 2017; Tolman, 1948; Wilson et al., 2014). Neurally, cognitive maps are assumed to rely on the activity of spatially tuned cells in the hippocampal-

entorhinal system. For example, during spatial navigation, place cells in the hippocampus exhibit increased firing at a particular location within an environment (O'Keefe & Dostrovsky, 1971). Grid cells in the adjacent entorhinal cortex fire at multiple locations within an environment and these locations form a hexagonal grid (Hafting et al., 2005). Together, these cells enable self-localization and geometric computations supporting spatial navigation, e.g., the computation of distances and directions (Bush et al., 2015; Moser et al., 2017). Beyond spatial navigation, recent studies have shown hippocampal-entorhinal map-like and grid-like representations of more abstract information, e.g., in graph-like structures (Garvert et al., 2017) as well as in feature and concept spaces (Bao et al., 2019; Constantinescu et al., 2016; Tavares et al., 2015; Theves et al., 2019, 2020; Viganò et al., 2021). Therefore, hippocampal-entorhinal cognitive maps have been suggested to provide a more general mechanism for organizing information, allowing for adaptive decision making (Behrens et al., 2018; Bellmund et al., 2018; Garvert et al., 2023; Kaplan et al., 2017; Schiller et al., 2015). For example, two recent studies showed distance- and grid-like representations for novel inferences during decision making in a two-dimensional map of social hierarchies (Park et al., 2020, 2021).

In decision making, states in the world and values are usually considered different entities, i.e., values (rewards) are received after performing an action in a given state. However, it is conceivable that values constitute states themselves, which can be represented in a cognitive map. In line with this notion, Bongioanni et al. (2021) demonstrated first evidence for a grid-like representation of an abstract value space defined by reward magnitude and probability in macaques. While choice options in previous studies (Bongioanni et al., 2021; Park et al., 2020, 2021) were static with regard to their locations in the abstract space, an interesting question is whether the same map-like representation would code for values of options changing over time. By facilitating computations of directions of and distances between value changes over time, such a cognitive map could enable efficient prediction of future values for prospective decision making. This map could then be used to read out resulting values and generate lower-dimensional signals of the value difference between options and their identities for choices. First evidence for hippocampal neurons encoding position in a value space spanned by changing reward probabilities has been demonstrated in macaques (Knudsen & Wallis, 2021). However, it remains elusive whether an entorhinal grid-like representation would encode changing values during prospective decision making in humans.

Here, we aimed to investigate whether the entorhinal cortex integrates relational information about changing values during prospective decision making using a grid-like representation of an abstract value space. To address this question, we combined functional magnetic resonance imaging (fMRI) with a prospective decision making task which required participants to integrate values in an abstract two-dimensional value space. Our behavioral results show

that participants integrated and extrapolated changes along the two value dimensions to guide prospective choice, indicating they formed a map of the relationships between options. Crucially, while participants traversed the abstract value space along trajectories, the entorhinal cortex exhibited a grid-like representation, suggesting the formation of a cognitive map. A network of brain regions, including the ventromedial and dorsal prefrontal cortex, tracked not only the value difference between options during choices, but also particularly the prospective value component. Furthermore, the occipital-temporal cortex represented the more valuable option for decision making.

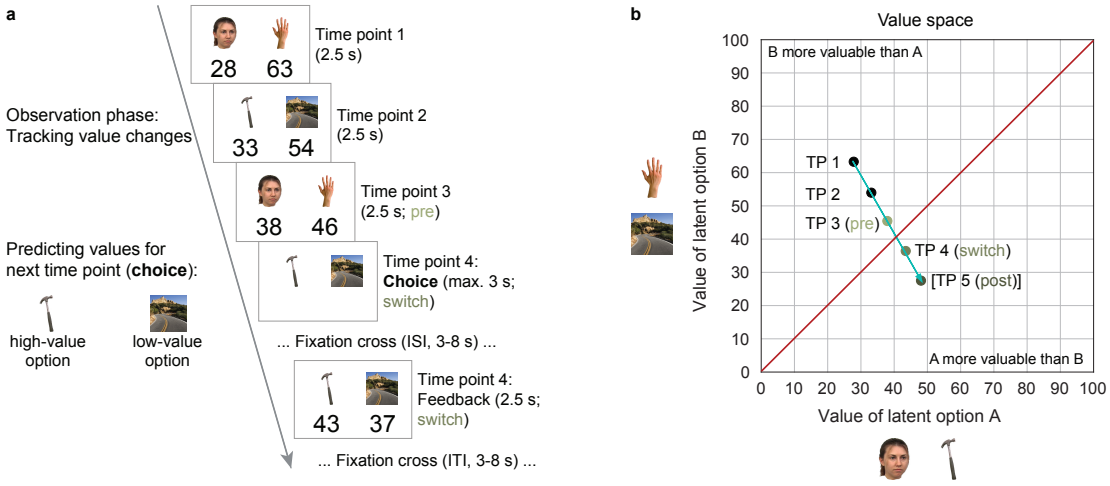


Fig. 1 | Design of prospective decision making task. Participants were instructed to track and predict changing values of two latent options A and B. Each latent option was signaled by one of two associated images, e.g., a face or a tool for option A and a hand or a scene for option B. Participants were instructed which images signaled the same option before the task. **a** Example trial, consisting of an observation phase and an active choice. During the observation phase, participants viewed the two options along with their changing values over a sequence of 3 time points (TPs). Across time points, the two value-congruent stimuli of a given option alternated. Subsequently, participants were asked to choose the option with the higher value at the next time point (choice time point). For their choice, participants had to consider how the values changed over time and accordingly how they will have changed towards the choice time point. Following their choice, they received feedback about the actual values at the choice time point. Note that in this example trial, the more valuable option changed from hand / scene (option B) in the beginning of the trial to face / tool (option A) at the choice time point. To facilitate fast tracking of the value changes across time points, options were displayed on the same side of the screen during the observation phase but sides were random at the choice time point. **b** A trial with its sequence of time points formed a trajectory through an underlying abstract two-dimensional value space, with the dimensions corresponding to the values associated with the two options. Each time point corresponded to a particular location in the value space, depicted by dots. The arrow in petrol depicts the trajectory through these locations. Trajectories crossing the 45°-diagonal of the space depicted in red involved a switch of the more valuable option. Choices sampled different time points across trajectories (trials). The first time point after the diagonal is referred to as the switch, the last time point before the diagonal as the pre and the second time point after the diagonal as the post time point. Note that the trial ended after the choice time point (**a**) and the post time point is added to the trajectory for illustration. Stimuli taken from publicly available stimulus datasets (Brady et al., 2008; Cichy et al., 2016; Kiani et al., 2007; Konkle et al., 2010; Kriegeskorte et al., 2008; Righi et al., 2012).

Results

Participants integrate and extrapolate value changes for prospective choices

We monitored whole-brain activity using fMRI while 46 participants performed a prospective decision making task (Fig. 1). The task required participants to maximize reward by tracking and predicting values (i.e., reward magnitudes) associated with two choice options. Each trial (Fig. 1a) consisted of an observation phase and an active choice. During the observation phase, participants viewed the two options along with their changing values over a sequence of time points. They were instructed to carefully track the value changes to be able to predict the options' values at the next time point. After 3-5 observed time points, participants were asked to choose the option with the higher value at the next time point (choice time point). Correct choices were translated into a monetary bonus for participants, which was based on the options' values.

More specifically, the two options were represented by the same four category-specific stimuli, which were mapped onto the two options (e.g., face/tool signaled option A, while hand/scene signaled option B). The stimulus mapping remained constant throughout the task and participants were informed about it before. Across time points, the two value-congruent stimuli of a given option alternated.

Crucially, a sequence of time points formed a trajectory through an underlying abstract two-dimensional value space, with the dimensions of the space corresponding to the values associated with the two options (Fig. 1b, Supplementary Fig. 1). In this space, the 45°-diagonal represented locations where the two options had the same values. Trajectories crossing the 45°-diagonal therefore involved a switch in which of the two options was more valuable. Tracking value changes over time, essentially recognizing the direction of and distances along a trajectory, allowed for prediction of future values and therefore detection of switches.

Participants' overall performance of the task, as indicated by choices of the more valuable option, was high (Fig. 2a, $M = 87.70$ %, $SD = 6.48$ %). If participants considered value changes over time for their choices, they should have detected switches of the more valuable option from one time point to the next. Indeed, they detected switches significantly more often than expected by chance (Fig. 2b; $t(45) = 10.82$, $p < .001$). Apart from the switch time point, the more valuable option was the same as at the preceding time point and participants could simply stay with that option. We therefore compared switch performance with the time points before (pre) and after (post) a switch. This comparison revealed a significant effect of time point (Fig. 2b; $F(2,90) = 35.93$, $p < .001$). Post-hoc pairwise tests indicated significantly reduced performance for both pre and switch compared to post (pre: $t(45) = -10.72$, $p < .001$; switch: $t(45) = -7.35$, $p < .001$) but no significant difference between pre and switch ($t(45) = -$

0.90, p = .40; all with α = 0.016, Bonferroni-corrected for three comparisons; controls for pre and post individually against chance: pre $t(45)$ = 11.02, p < .001, post $t(45)$ = 45.17, p < .001). Similarly, we observed a significant effect of time point on reaction times (Supplementary Fig. 2a; $F(2,90)$ = 60.65, p < .001; post hoc pairwise tests: pre-post: $t(45)$ = 9.85, p < .001, switch-post: $t(45)$ = 1.74, p = .09, pre-switch: $t(45)$ = 8.33, p < .001, post hoc pairwise tests with α = 0.016). This pattern of results suggests that participants successfully detected switches of the more valuable option and may even have over-extrapolated the value changes, leading to earlier switches than optimal.

As switches are induced by the 45°-diagonal of the value space, we tested more continuously how performance is influenced by the distance between the choice location and the diagonal using participant-specific logistic regressions. We found that the likelihood of correct choices increased with increasing distance of the choice location to the diagonal (Fig. 2c-d; $t(44)$ = 8.03, p < .001). As a control, we tested the same relationship using only choices in switch trajectories where locations lay inherently closer to the diagonal (Fig. 2c; $t(44)$ = 6.60, p < .001). This suggests that the closer choice locations are to the 45°-diagonal and hence the more similar the options' values become, the more difficult the choices become for participants.

Next, we investigated whether a reinforcement learning model which captured the prospective nature of the task, i.e., the value changes over time, fitted participants' choice behavior better than a model that did not. To this end, we modified a Rescorla-Wagner model (Rescorla & Wagner, 1972) so that it updated value estimates within a trial based on prediction errors and additionally value changes over time points:

$$V_{TP+1} = V_{TP} + \alpha*(O_{TP} + C_{TP} - V_{TP}) \text{ with } C_{TP} = O_{TP} - O_{TP-1},$$

whereby V_{TP} and V_{TP+1} are values at the current and next time points, respectively, O_{TP} is the outcome at the current time point, C_{TP} reflects how the value has changed from the previous to the current time point and α is the learning rate (free parameter of the model). In essence, this prospective Rescorla-Wagner model does not only update the expected value to the outcome just observed, but learns which outcome to expect given the past history of changes. We compared this to the original Rescorla-Wagner model which does not consider value changes over time points:

$$V_{TP+1} = V_{TP} + \alpha*(O_{TP} - V_{TP}).$$

As expected, the prospective Rescorla-Wagner model fitted the data better than the original Rescorla-Wagner model (Fig. 2e; model comparison per AIC: $t(45)$ = -8.71, p < .001, with α = 0.01, Bonferroni-corrected for five tests including alternative models). Initially, we constrained the learning rate of the prospective model to the range between 0 and 1, with 1 reflecting full

updating according to prediction errors and value changes. We observed a ceiling effect for the learning rate, with many participants having learning rates of 1 (M = 0.94, SD = 0.12). For this reason, we removed the upper bound of the learning rate and observed learning rates slightly above 1 on average, suggesting slight over-updating in line with the performance reduction at the pre time point described above (Supplementary Fig. 2f; M = 1.09, SD = 0.25; model comparison of unbound and bound model per AIC: $t(45) = -3.96$, $p < .001$). The learning rate correlated positively with performance at the switch time point (Supplementary Fig. 2g; $r(44) = .55$, $p < .001$) but negatively with performance at the pre time point (Supplementary Fig. 2h; $r(44) = -.44$, $p = .003$), reflecting the advantage and disadvantage of over-updating. In addition, we implemented a set of alternatives for the prospective model, e.g., with a separate learning rate for the change term or with a term for an expected prediction error (see Methods). The prospective model described above fitted the data better than all alternatives (Supplementary Fig. 2i; all $p < .001$). These modeling results confirm and extend our previous pre-switch-post performance analysis by showing that participants indeed extrapolated value changes for prospective choices, though to a slightly larger extent than optimal, presumably causing too early switches.

Lastly, we reasoned that if prospective decision making is supported by a cognitive map similar to spatial navigation, then participants with better navigational abilities may also perform better in our prospective decision making task. To investigate whether this is the case, we tested whether behavior in our task correlated with participants' self-reported navigational abilities and preferences (as measured by the Santa Barbara Sense of Direction Scale questionnaire, Hegarty (2002), completed in the last part of the study). We observed a significant positive correlation between the learning rate of the prospective Rescorla-Wagner model and self-reported navigational abilities and preferences (Fig. 2f; correlation with learning rate: $r(44) = .34$, $p = .02$; correlation with overall performance: $r(44) = .11$, $p = .45$; with $\alpha = 0.025$, Bonferroni-corrected for two tests). In addition, we wondered whether the prospective component of integrating and extrapolating values over time in our task relates to model-based decision making in the two-stage task, which assesses reliance on a model of state transition probabilities across two decision stages (Daw et al., 2011). Contrary to our expectations, we did not observe a significant correlation with model-based decision making in the two-stage task, potentially due to overall reduced model-based decision making in our sample (Supplementary Fig. 3a-c; correlation with learning rate: $r(44) = -.19$, $p = .20$; correlation with overall performance: $r(44) = .21$, $p = .15$; with $\alpha = 0.025$, Bonferroni-corrected for two tests).

Taken together, our behavioral results demonstrate that participants were able to integrate and extrapolate changes along the two value dimensions of the space to guide choice, suggesting they formed a map of the relationships between options.

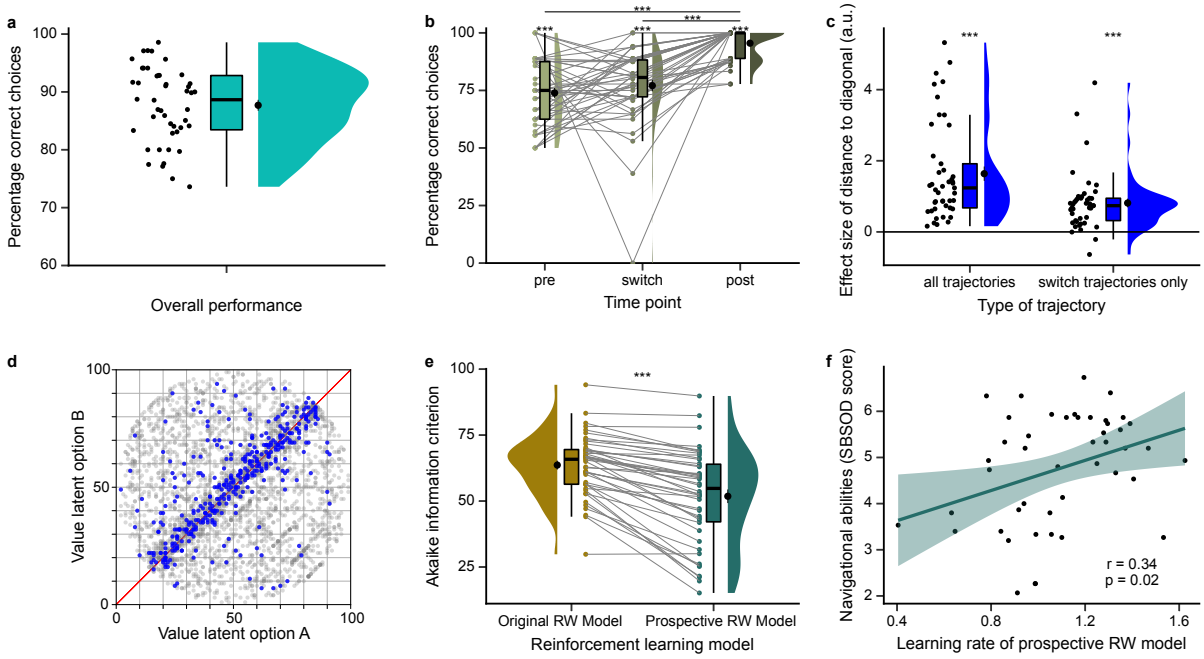


Fig. 2 | Participants integrate and extrapolate value changes for prospective choices. a Overall performance across all trajectories of the task. **b** Performance along different time points on switch trajectories: time point before the switch (pre), time point of the switch (switch) and time point after the switch (post). Performance at all time points was significantly above chance, suggesting that participants succeeded in detecting switches. Performance at pre and switch time points was significantly lower than at post time points. **c** Effect of the distance between the choice location and the 45°-diagonal of the value space on performance. Depicted are effect sizes estimated by a logistic regression, separately for including all trajectories (left) as well as including only switch trajectories as a control (right). In both cases, effect sizes are significantly positive, indicating that the likelihood of correct choices increased with a higher distance to the diagonal. **d** Visualization of the effect of the distance to the diagonal in **c**. Blue dots depict incorrect choice locations across participants, clustering around the 45°-diagonal. Grey dots depict correct choice locations. **e** Reinforcement learning model comparison. The Akaike information criterion (AIC) is significantly lower (better model fit) for the prospective Rescorla-Wagner (Prospective RW) model (right) compared to the Original RW model (left). **f** The learning rate α of the prospective Rescorla-Wagner model correlates significantly positively with participants' self-reported navigational abilities and preferences (Santa Barbara Sense of Direction Scale (SBSOD) questionnaire, Hegarty, 2002). Dots represent participants' data points; line represents linear regression line, with shaded regions as the 95% confidence interval. **a,b,c,e,f** Raincloud plots: dots represent participants' data points; boxplots show median and upper/ lower quartile with whiskers extending to the most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bar corresponds to mean ± SEM; distributions depict probability density function of data points. *** $p < .001$

Entorhinal cortex exhibits grid-like representation for value space

Our behavioral results suggest that participants formed a relational value map. Relationships between landmarks in physical space, as well as non-spatial relational structures are represented by entorhinal grid cells in a cognitive map. We hypothesized that the entorhinal cortex might also encode changing values using a grid-like representation. Such a neural representation would facilitate computations of directions of and distances between value changes over time and thereby enable efficient prediction of future values. Previous research has shown that the regular hexagonal firing pattern of grid cells in the entorhinal cortex translates to hexadirectional activity modulations during spatial navigation in fMRI (Doeller et al., 2010). In our prospective decision making task, a sequence of time points formed a trajectory through an underlying abstract value space (Fig. 1b). More specifically, participants moved along trajectories with directions ranging from 0°-350° in 10°-steps in each of the four task blocks (fMRI runs; Supplementary Fig. 1). If participants formed a cognitive map of changing values, akin to maps in physical space, then activity in the entorhinal cortex should show a hexadirectional modulation during this movement through the value space, with higher activity for trajectories aligned with the putative grid orientation (phase of the hexadirectional signal) than for trajectories misaligned with the putative grid orientation (Fig. 3a). To test this hypothesis, we implemented a cross-validation procedure, estimating the putative grid orientation using three of four task runs and testing for a hexadirectional modulation aligned to the orientation in the left-out run (based on Doeller et al., 2010; Nau et al., 2018).

In line with our hypothesis, we observed significant hexadirectional modulation of activity in the entorhinal cortex (Fig. 3b-c; small volume correction with $p_{FWE} < .05$ TFCE; MNI peak voxel coordinates: 18,-6,-26; peak voxel $t(45) = 4.17$, $p_{FWE} = .003$; one-sided test). At the whole-brain level, we observed no further regions surviving correction (see Supplementary Table 1 for whole-brain results with a liberal threshold of $p_{uncorr} < .001$). To visualize the hexadirectional modulation in the significant entorhinal cluster, we sorted trajectories according to the putative grid orientation and illustrate effects of aligned and misaligned 30°-bins (Fig. 3d). A complementary ROI analysis of the entorhinal cortex confirmed the hexadirectional (6-fold) effect and showed that the modulation of activity was specific to a 6-fold symmetry in line with grid-like responses, as there were no significant effects for control symmetries (Fig. 3e; ROI analysis, one-sided tests: 4-fold $t(45) = -0.21$, $p = .59$; 5-fold $t(45) = -0.02$, $p = .51$; 6-fold $t(45) = 2.91$, $p = .003$; 7-fold $t(45) = 1.34$, $p = .10$; 8-fold $t(45) = 1.42$, $p = .08$; control symmetries n.s.; with $\alpha = 0.01$, Bonferroni-corrected for five tests). There was no significant correlation between the magnitude of hexadirectional modulation and task performance (Supplementary Fig. 4g; $r(44) = -.08$, $p = .59$).

Furthermore, we performed exploratory analyses to investigate the relationship between the entorhinal grid system and the underlying value space. First, we wondered whether grid orientations would be anchored to a particular reference direction through the value space. We speculated that a direction of 45° constitutes a particularly informative reference direction because it indicates that values of both options change at the same rate and – given that it is parallel to the 45°-diagonal of the value space – that there will be no switch of the more valuable option. We thus examined whether grid orientations in the significant entorhinal cluster would cluster around 45°, which was indeed the case (Fig. 3f; V-Test for mean orientation of 45° across participants: $p = 0.01$). Secondly, we wondered whether the grid-like representation of the value space might be modulated by value (i.e., reward magnitude) itself. Recent evidence in rodents demonstrated restructuring of grid cells in response to reward locations during spatial navigation, with movement of grid fields towards reward locations and higher firing rates for grid fields closer to reward locations (Boccara et al., 2019; Butler et al., 2019). We therefore examined whether the magnitude of hexadirectional modulation differs between areas of the value space with relatively higher and lower values. To test this, we performed a median split of trajectories according to their mean value, i.e., contrasting trajectories in the lower left triangle of the space (low-value-area) with trajectories in the upper right triangle of the space (high-value-area). We note that this median split led to a substantial reduction of available trajectories per value condition and an unbalanced sampling of directions between the conditions, rendering this analysis less robust (Supplementary Fig. 4h; significant interaction between value condition and direction: $F(35,1575) = 3.56$, $p < .001$). Using the significant entorhinal cluster as ROI, we then repeated the cross-validated hexadirectional analysis separately for the two value conditions. The analysis suggested no difference in hexadirectional modulation between the two conditions (Supplementary Fig. 4i; $t(45) = -1.30$, $p = .19$). However, it is interesting to note that - contrary to our expectations based on the rodent literature - it suggested a hexadirectional modulation effect in low-value-areas but not in high-value-areas (Supplementary Fig. 4i; low-value: $t(45) = 2.06$, $p = .02$; high-value: $t(45) = 0.30$, $p = .38$; one-sided tests).

Taken together, these results provide evidence that the entorhinal cortex encoded the abstract value space using a grid-like representation, suggesting the formation of a cognitive map.

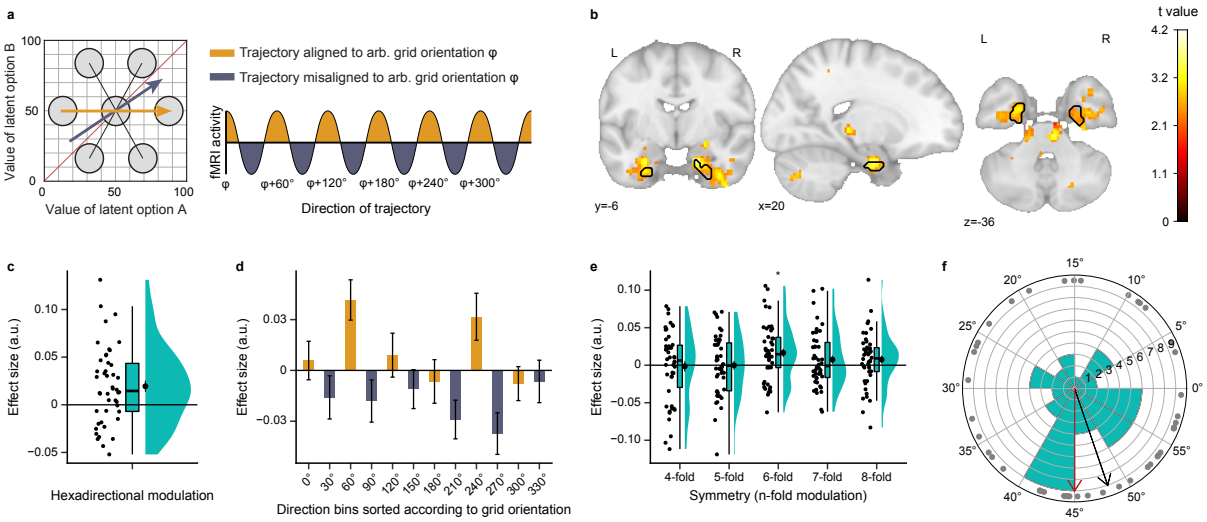


Fig. 3 | Entorhinal cortex exhibits grid-like representation for value space. a Logic of the grid analysis. Left: Schematic of a grid cell with its regular hexagonal firing pattern and an arbitrary orientation of φ = 0°, superimposed on the value space. Two example trajectories and their relation to the grid are shown: while the yellow example trajectory is aligned with the orientation of the grid, the blue example trajectory is misaligned. Right: The regular firing of grid cells translates to hexadirectional activity modulations in fMRI, with higher activity for trajectories aligned vs. misaligned with the grid orientation φ. **b** Grid-like hexadirectional modulation of activity, aligned with the mean grid orientation of the entorhinal cortex. For visualization, statistical image is thresholded at $p_{uncorr} < .01$. Voxels within the black outline are significant after correction for multiple comparisons using small volume correction in the entorhinal cortex ($p_{FWE} < .05$, TFCE). Statistical image is displayed on the MNI template. **c** Visualization of the hexadirectional effect sizes in the significant entorhinal cluster in **b** across participants. **d** Further visualization of the hexadirectional effect in the significant entorhinal cluster in **b**. Depicted are effect sizes of 30°-directional bins, sorted according to the mean grid orientation of the entorhinal cortex. Yellow bars depict aligned directions, blue bars depict misaligned directions (as in **a**). **e** Modulation of activity for different symmetries in the entorhinal cortex ROI (6-fold refers to the hexadirectional modulation of interest and is significant; 4-fold, 5-fold, 7-fold and 8-fold refer to controls, all n.s.). **e** Polar histogram of grid orientations in 60°-space of the significant entorhinal cluster across participants. Grey circles depict individual participants' orientations, bars of the histogram depict bins of 5°. Black arrow shows the circular mean of all participants' orientations. Red arrow highlights an orientation of 45°. Grid orientations cluster around 45°.

A network of brain regions tracks the prospective value difference during choices

To make a decision, representing values solely in a two-dimensional value map is not useful. Instead, values of the choice options also need to be mapped onto a single common scale for comparison. We thus tested whether neural signals track the value difference between the chosen and the unchosen option, especially in vmPFC based on previous literature. For this purpose, we modeled choice time points as a function of the chosen and unchosen values, derived from the prospective Rescorla-Wagner model, and contrasted these effects to test for a modulation by the value difference. We observed significant positive and negative modulation of neural activity by the value difference in a network of brain regions (see Fig. 4a for whole-brain effects; $p_{FWE} < .05$ TFCE-corrected; see Supplementary Table 2 for a list of significant clusters). Positive modulations reflected higher activity for a higher value difference and included amongst others vmPFC (MNI peak voxel coordinates: 3,42,-8; peak voxel statistics: $t(45) = 5.99$, $p_{FWE} < .001$), putamen, insular cortex, hippocampus, amygdala as well as motor and somatosensory cortex. Negative modulations reflected higher activity for a smaller value difference and included amongst others lateral parts of vPFC / OFC, dmPFC, thalamus and parietal cortex. These effects were still present when controlling for reaction time (Supplementary Fig. 5a). Furthermore, the value difference effect in the vmPFC cluster correlated significantly positively with task performance (Fig. 4c, $r(44) = .34$, $p = .02$; after exclusion of outlier: $r(44) = .33$, $p = .03$).

Moreover, we aimed to investigate whether neural signals would track particularly the prospective component of the value difference, i.e., the difference based on the prospective values at the choice time point rather than the non-prospective values of the preceding time point. For this purpose, we subtracted value estimates of the original Rescorla-Wagner model (non-prospective) from value estimates of the prospective Rescorla-Wagner model, thereby extracting particularly the prospective value component for each option. We then modeled choice time points as a function of the prospective components of the chosen and the unchosen option and contrasted these effects to test for a modulation by the prospective value difference. Again, we observed widespread significant positive and negative modulation of neural activity by the prospective value difference (Fig. 4d; $p_{FWE} < .05$ TFCE-corrected; see Supplementary Table 3 for a list of significant clusters). Many clusters overlapped with those tracking the original value difference. However, a cluster in vmPFC / OFC extended more dorsally and bordered ACC (MNI peak voxel coordinates: -7,52,-8; peak voxel statistics: $t(44) = 5.75$, $p_{FWE} < .001$). This prefrontal cluster was still present when controlling for reaction time (Supplementary Fig. 5b).

Taken together, these results demonstrate that a network of brain regions, including value regions such as vmPFC and dPFC, tracked not only the value difference between options during choices, but also particularly the prospective component of that value difference.

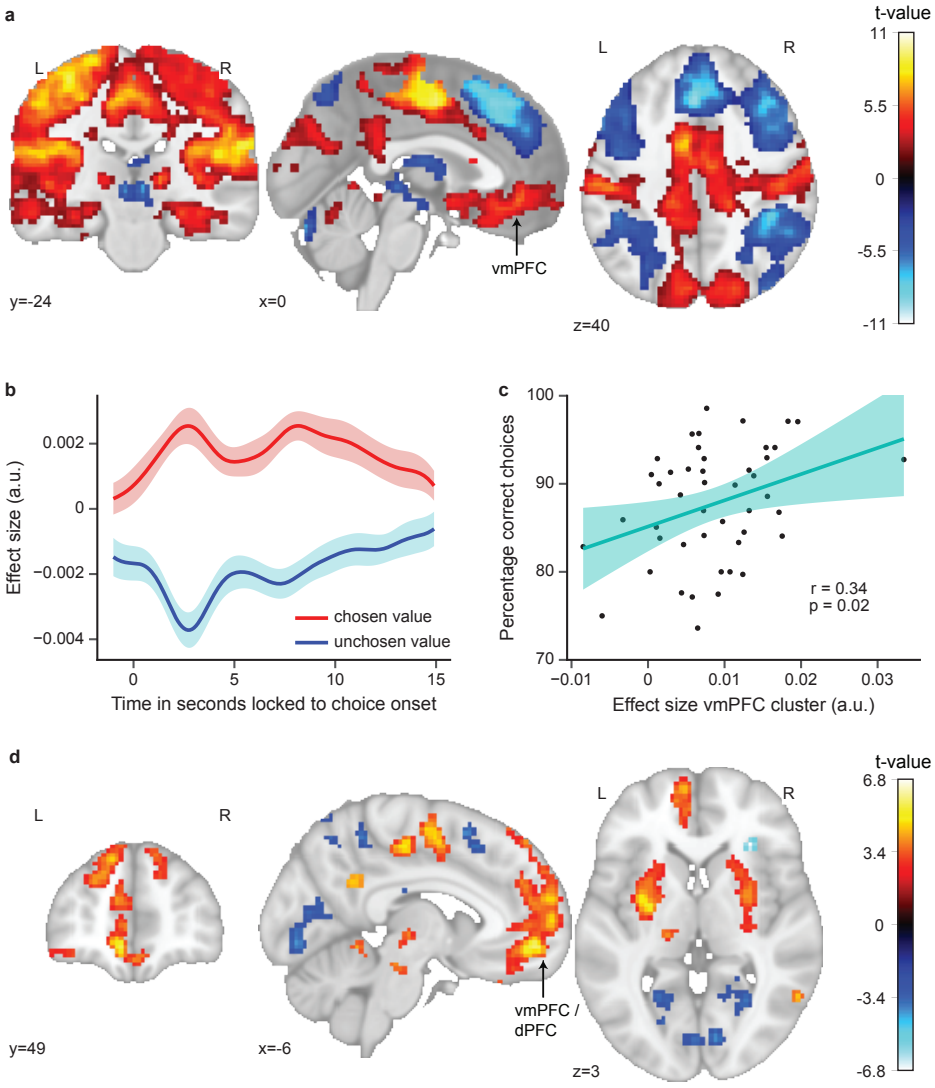


Fig. 4 | A network of brain regions tracks the prospective value difference during choices. a Modulation of activity by the difference between model-derived chosen vs. unchosen value during choices. Clusters depicted survive whole-brain correction (p_{FWE} < .05, TFCE). Statistical image is displayed on the MNI template. **b** Based on our expectation of a value difference effect in vmPFC, we visualize the effect in the vmPFC cluster by showing the time courses of the effect sizes of the chosen and unchosen value, time-locked to choice onset (choice onset at 0s). Lines represent the mean across participants, with shaded regions as the 95% confidence interval. **c** The value difference effect in vmPFC correlates significantly positively with performance. **d** Modulation of activity by the prospective component of the value difference during choices. The prospective component refers to the influence of values estimated by the prospective Rescorla-Wagner model over values estimated by the original (non-prospective) Rescorla-Wagner model. Clusters depicted survive whole-brain correction (p_{FWE} < .05, TFCE). Statistical image is displayed on the MNI template.

Occipital-temporal cortex represents the more valuable choice option

To make prospective decisions in the absence of direct experience, stimulus values need to be constructed based on past experience. We assumed that this construction process might involve the sensory representations of the respective choice options. We thus hypothesized that we might be able to index this construction process by assessing the stimulus representations at the choice point, with the idea that the identity of the option associated with the objectively higher value might be represented more strongly than that of the option associated with the objectively lower value.

To test this hypothesis, we leveraged neural responses to category-specific stimuli (faces, tools, scenes, body parts), which are known to activate category-selective regions of the occipital-temporal cortex. Using data from an independent picture viewing task (PVT) which took place before the prospective decision making task, we trained a decoder (support vector classifier) on occipital-temporal cortex voxels to distinguish neural activation patterns of the four category-specific stimuli (Fig. 5a). In the PVT, participants viewed a stream of pictures of the category-specific stimuli and performed a one-back cover task (performance: M = 94.60 %, SD = 9.87 %, Supplementary Fig. 6a). We first examined how well we could decode stimulus category within the PVT, before applying the decoder to the decision making task. For this purpose, we estimated trial-related activation patterns and implemented a cross-validation scheme, training the decoder on a subset of trials and predicting category labels on a subset of left-out test trials. Category decoding accuracy within the PVT was well above chance level (Supplementary Fig. 6b; M = 80.24 %, SD = 7.09 %, chance level = 25 %).

Subsequently, we tested our hypothesis that the high-value option would be represented more strongly than the low-value option, in particular in participants who performed the task well. We estimated trial-related activation patterns in the PVT and choice-time-point-related activation patterns in the prospective decision making task. We then trained a decoder using the independent PVT data and applied it to neural activation patterns of choice time points in the prospective decision making task (Fig. 5b). We extracted the probabilities which the decoder assigned to each of the four stimuli and computed two comparisons: First, we compared the probabilities assigned to the two stimuli presented on-screen during choice: probability of the high-value stimulus vs. the low-value stimulus. Secondly, we compared the probabilities assigned to the two value-congruent stimuli which were not presented on-screen during choice (but during the time point before): probability of the congruent high-value stimulus vs. the congruent low-value stimulus. As expected, we observed significantly higher probabilities for the high-value vs. the low-value stimulus, both when comparing on-screen stimuli (Fig. 5c-d; $t(45) = 8.92$, $p < .001$) as well as when comparing congruent off-screen stimuli (Fig. 5f-g; $t(45) = 7.54$, $p < .001$). Furthermore, the probability differences correlated

positively with task performance (correlation for on-screen stimuli: Fig. 5e, $r(44) = .38$, $p = .01$; correlation for congruent stimuli: Fig. 5h, $r(44) = .27$, $p = .07$). In addition, we tested whether the probability differences in favor of the high-value option depended on the distance between the choice location and the 45°-diagonal of the value space. We found that the probability difference high-value vs. low-value for on-screen stimuli increased significantly with increasing distance to the diagonal (Supplementary Fig. 6c; on-screen stimuli: $t(45) = 3.30$, $p = .002$; congruent stimuli: $t(45) = -1.00$, $p = .31$), presumably reflecting difficulty similar to the behavioral performance-distance effect. Contrary to our expectations, there was no significant correlation between the probability differences in favor of the high-value option and the magnitude of hexadirectional modulation (Supplementary Fig. 6d-e; on-screen stimuli: $r(44) = -.10$, $p = .54$; congruent stimuli: $r(44) = -.15$, $p = .34$).

We note that while we compared on-screen and off-screen congruent stimuli separately, the temporal proximity of their presentations during time points within a trajectory might render disentangling their effects difficult. To control for the temporal proximity to some extent, we repeated this analysis using only those choices which sampled the switch time point. In this case, the direction of the effect during choice (high-value vs. low-value, especially for the comparison of the congruent stimuli) should be different from the direction of the effect at the time point before the switch (pre). The probability difference effects and correlations with performance were still present in this control analysis (Supplementary Fig. 6f-k; on-screen stimuli: $t(45) = 5.42$, $p < .001$, $r(44) = .49$, $p = .003$ (correlation n.s. after outlier control); congruent stimuli: $t(45) = 6.17$, $p < .001$, $r(44) = .46$, $p < .001$).

Taken together, these results show that occipital-temporal cortex represented the more valuable option for decision making.

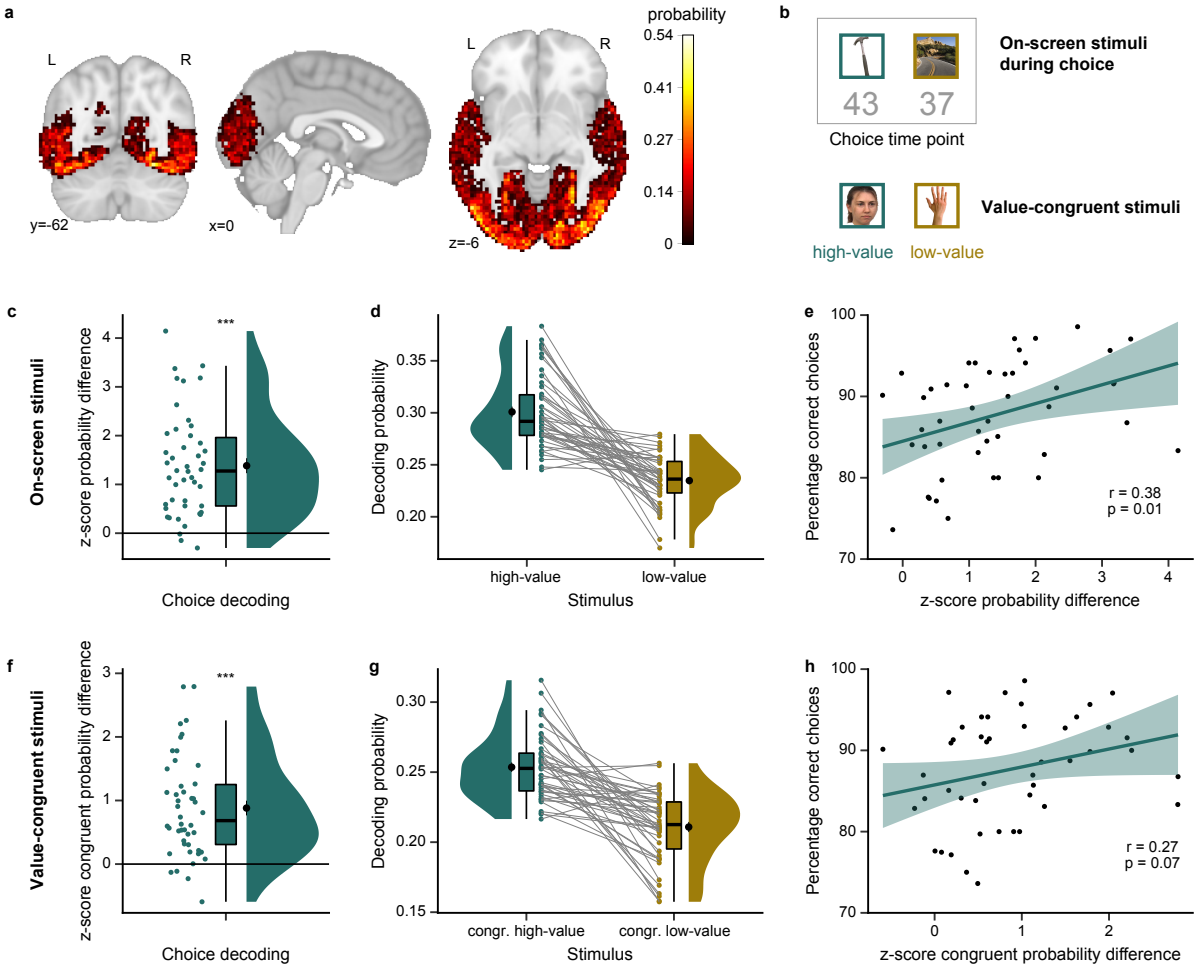


Fig. 5 | Occipital-temporal cortex represents the more valuable choice option. a Probability map of occipital-temporal cortex voxels used for the decoding analysis, displayed on the MNI template. Note that the decoding analysis was performed using participant-specific masks in native space, which were created based on anatomical and functional information (M = 2235 voxels, see Methods for details). This map depicts the probability of voxels in MNI space to be included in the mask across participants. **b** Example for stimuli at a choice time point. Top: Stimuli which are on-screen during the choice. Objective values of the stimuli at the choice time point are shown in grey for visualization (not shown to participants). Below: Value-congruent stimuli which are not presented on-screen during the choice. We compared decoding probabilities for the high- vs. low-value stimulus, separately for on-screen stimuli (**c-e**) and value-congruent off-screen stimuli (**f-h**). **c** Z-scores for the decoding probability difference for the on-screen high- vs. low-value stimuli based on decoding permutation test (see Methods). Occipital-temporal cortex represents the high-value stimulus significantly stronger than the low-value stimulus. **d** Visualization of the effect in **c**, showing the probabilities the decoder assigned to the stimuli (before the permutation test). **e** The high- vs. low-value difference score correlates significantly positively with performance. **f** Z-scores for the decoding probability difference for the value-congruent off-screen high- vs. low-value stimuli based on decoding permutation test (see Methods). Occipital-temporal cortex represents the congruent high-value stimulus significantly stronger than the congruent low-value stimulus. **g** Visualization of the effect in **f**, showing the probabilities the decoder assigned to the stimuli (before the permutation test). **h** The correlation between the congruent high- vs. low-value difference score and performance was not significant. **c,d,f,g** Raincloud plots: dots represent participants' data points; boxplots show median and upper/ lower quartile with whiskers extending to the most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bar corresponds to mean ± SEM; distributions depict probability density function of data points. **e,h** Correlation plots: dots represent participants' data points; line represents linear regression line, with shaded regions as the 95% confidence interval. Stimuli taken from publicly available stimulus datasets (Brady et al., 2008; Cichy et al., 2016; Kiani et al., 2007; Konkle et al., 2010; Kriegeskorte et al., 2008; Righi et al., 2012). *** $p < .001$

Discussion

Our capacity to predict future values of choice options is central to many decisions we face in everyday life. Understanding the mechanisms by which the brain enables prospective decision making is therefore of particular importance. In this study, we combined fMRI with a prospective decision making task to investigate how the brain represents relational information about changing values of choice options in an abstract value space. Participants integrated and extrapolated changes along the two value dimensions to guide prospective choice. Crucially, while participants traversed the abstract value space along trajectories, the entorhinal cortex exhibited a grid-like representation, suggesting the formation of a cognitive map. A network of brain regions, including vmPFC and dPFC, tracked the prospective value difference between options and occipital-temporal cortex represented the more valuable choice option.

Our finding of an entorhinal grid-like representation of an abstract value space dovetails with the broader idea of cognitive maps encoding abstract information (Behrens et al., 2018; Bellmund et al., 2018; Kaplan et al., 2017; Schiller et al., 2015) and research showing that vmPFC jointly encodes values and states (Moneta et al., 2023). Map-like representations of relationships between states enable prediction of future states (Stachenfeld et al., 2017). In spatial navigation and memory, the hippocampal-entorhinal system is involved in prospective mental simulations and imaginations of events and navigational goals (Bellmund et al., 2016; Brown et al., 2016; Horner et al., 2016; Nyberg et al., 2022; Pfeiffer & Foster, 2013; Schacter et al., 2007). In the context of prospective value-based decision making, predicting future states corresponds to predicting future values of choice options, such as in the introductory example of a stock market. Critically, value changes over time can be conceptualized as sequences through an abstract value space, allowing for prospective decision making by facilitating the computation of geometric distances and directions. In line with this, it is noteworthy that the degree to which participants updated values over time in our value space task correlated with self-reported navigational abilities and preferences during spatial navigation in everyday life.

Our results extend recent evidence for map-like representations of value spaces in macaques into human research. For example, Bongioanni et al. (2021) showed a grid-like representation in the macaque medial frontal cortex as a function of step-like transitions between static options in a space spanned by reward magnitude and probability. Knudsen and Wallis (2021) found that hippocampal neurons in macaques, similar to place cells during spatial navigation in a physical space, encode position in a value space spanned by changing reward probabilities. Here, we provide evidence for an entorhinal grid-like representation of a value space during prospective decision making in humans.

Interestingly, our exploratory analysis suggests that the entorhinal grid system adapts to properties of the value space. More specifically, our results suggested an anchoring of grid orientations around 45°. We speculate that a direction of 45° constitutes a particularly informative reference direction through our value space. This is because it indicates that values of both options change at the same rate and – given that it is parallel to the 45°-diagonal of the value space – that there will be no switch of the more valuable option. In line with this speculation, recent evidence in spatial navigation demonstrated anchoring of grid orientations to an informative axis in a virtual navigation arena which minimizes spatial uncertainty (Julian et al., 2018; Julian & Doeller, 2021; Navarro Schröder et al., 2020; Stensola et al., 2015). Our results suggest that grid orientations might anchor to an informative axis even in more abstract spaces. Furthermore, we wondered whether the grid-like representation of the value space might be modulated by value (i.e., reward magnitude) itself. Recent evidence in rodents demonstrated restructuring of grid cells in response to reward locations during spatial navigation, with movement of grid fields towards reward locations and higher firing rates for grid fields closer to reward locations (Boccara et al., 2019; Butler et al., 2019). While our results suggested no difference in the strength of grid-like representations between low- and high-value areas of the value space, they surprisingly pointed towards a grid-like representation of the low-value but not the high-value area. We note that these effects of an exploratory analysis should be interpreted with caution, given the absence of a clear difference in the grid-like signal. One could speculate that participants' subjective gain of correct choices might have been higher in low-value than in high-value areas because the received reward in high-value areas was high anyway, in agreement with notions of value distortions as value compression or diminishing utility (Juechems et al., 2021; Tversky & Kahneman, 1992). While the main goal of our study was to assess whether a grid-like representation encodes a value space in principle, future studies could aim to investigate modulations of such a grid-like representation by value itself.

It is conceivable to represent values during our task by two separate number lines, without necessarily integrating them into a two-dimensional space. Indeed, none of the participants reported having imagined the two-dimensional value space. However, our results, together with other studies demonstrating grid- and place-like representations of values (Bongioanni et al., 2021; Knudsen & Wallis, 2021), suggest indeed a neural representation of a two-dimensional space. Nevertheless, it is possible that different brain regions represent values differently, e.g., in a map-like format vs. combining them directly into a value difference or summary signal.

In light of this, while the entorhinal value map could support the prediction of future values by facilitating computations of directions of and distances between value changes over time,

other brain regions might read out the resulting values, map them onto a single common scale for comparison and thus generate a one-dimensional signal of the value difference used for decision making. In line with this notion and with previous literature in value-based decision making, a network of brain regions, including vmPFC, tracked the value difference between options during choices (Bartra et al., 2013; Boorman et al., 2009; De Martino et al., 2013; FitzGerald et al., 2009; Hunt et al., 2012; Knutson et al., 2005; Lee et al., 2014; Levy & Glimcher, 2012; O'Doherty et al., 2001; Padoa-Schioppa & Assad, 2006; Pelletier & Fellows, 2019; Plassmann et al., 2007). We observed both positive and negative modulations of activity by the value difference. While positive modulations might reflect the benefit of the chosen over the unchosen option, negative modulations could signal the relative value of the unchosen option as an alternative. Moreover, a vmPFC cluster extending more dorsally (dPFC) and bordering ACC tracked particularly the prospective component of the value difference. This prospective value difference effect is in line with reports of distinct model-based value correlates in dmPFC and reward rate tracking for trend-guided choice in neighboring dACC (Doll et al., 2015; Wittmann et al., 2016). Furthermore, the pattern of a value difference signal directly relevant for choices in vmPFC and particularly prospective value components extending into dPFC dovetails with reports of a functional gradient, with vmPFC encoding values for executable choices and dmPFC encoding abstractly modeled values (Nicolle et al., 2012). We speculate that the widespread involvement of brain regions in tracking values in our results might be explained by the high salience and relevance of values in the task. Ultimately, these value difference signals could be used to allocate attention to the more valuable option to guide eventual decision making in many foraging scenarios (Hall-McMaster et al., 2021; Hayden et al., 2011; Kolling et al., 2012; Wittmann et al., 2016). In line with this notion, occipital-temporal cortex represented the identity of the more valuable choice option. This result matches previous reports of prospective activation of states in the two-stage task (Doll et al., 2015).

Finally, we would like to point to possible limitations of our study. First, the direction and rate of value changes were constant within a trial and differed across trials of our task, based on the sampling of different trajectories. Future studies could add noise to the value changes or vary them in a more fine-grained manner within trials to investigate how a map-like representation of a value space translates to more ecological scenarios. Secondly, we did not find evidence for a correlation between the grid-like representation and performance of the prospective decision making task across participants. In light of this, previous research reported mixed results for across-participant grid-behavior correlations, ranging from positive to negative to no reports of correlations (Constantinescu et al., 2016; Horner et al., 2016;

Wagner et al., 2023). Future studies could opt for testing such relationships more fine-grained on an individual participant level.

In conclusion, our results provide evidence that the human entorhinal cortex uses a grid-like representation to integrate relational information about changing values in an abstract value space during prospective decision making, suggesting the formation of a cognitive map. This map might be used to generate lower-dimensional signals of the value difference between options and their identities for choices. Thus, our findings provide novel insight for our understanding of cognitive maps as a mechanism to guide prospective decision making in humans.

Methods

Participants

51 participants took part in this study. The sample size was determined by a power analysis using G*Power (Faul et al., 2007). This yielded a necessary sample size of 41 participants to achieve a statistical power of 80 % for a small-to-medium effect size (d = 0.4, α = 0.05, one-tailed t-test). Additionally, 10 participants were recruited to account for potential dropouts. All participants had normal or corrected-to-normal vision, no history of or current neurological or psychiatric disorders and were right-handed. Participants were recruited using the participant database of the Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.

For the data analysis, one participant was excluded due to missing fMRI data because of technical problems during data acquisition. Furthermore, four participants were excluded due to low performance of the prospective decision making task (performance criterion of 70% accuracy based on previous piloting). Thus, the final sample consisted of 46 participants (age: M = 28.15 years, SD = 4.77 years, range = 19-39 years; 25 female).

The study was approved by the ethics committee at the Medical Faculty at the University of Leipzig (421/19-ek) and all participants gave written informed consent prior to participation. Participants were reimbursed with a baseline fee of 10 € / h and could additionally earn a monetary bonus up to 10 € based on performance (see tasks for details).

Experimental procedure

Overview

The study consisted of three parts and lasted approximately three hours in total. The first part took place in a behavioral laboratory (approx. 45 min). Here, participants received instructions and training for the main task of the study, i.e., the prospective decision making task. In the second part (approx. 90 min), participants performed two tasks in the MRI scanner: First, they completed a picture viewing task (PVT) which served as an independent dataset to train a decoder for choice stimuli subsequently used in the prospective decision making task. Afterwards, they completed the prospective decision making task. In the third part (approx. 45 min), participants returned to a behavioral laboratory to complete two post-scanning tasks: the two-stage task (Daw et al., 2011) to study model-based vs. model-free decision making as well as the Santa Barbara Sense of Direction Scale (SBSOD) questionnaire (Hegarty, 2002) to assess navigational abilities and preferences.

Stimuli

Stimuli used for the picture viewing task and the prospective decision making task in the MRI scanner were category-specific pictures (faces, tools, scenes, body parts) which are known to

elicit neural responses in category-selective regions of the occipital-temporal cortex. Stimuli were taken from publicly available stimulus datasets (faces: Righi et al. (2012), Face images courtesy of Michael J. Tarr, Carnegie Mellon University, http://www.tarrlab.org/. Funding provided by NSF award 0339122; tools: Brady et al. (2008), https://konklab.fas.harvard.edu/#; scenes: Konkle et al. (2010), https://konklab.fas.harvard.edu/#; body parts: Cichy et al. (2016), Kiani et al. (2007), Kriegeskorte et al. (2008), http://userpage.fu-berlin.de/rmcichy/fusion_project_page/main.html). From these stimulus sets, three pictures of the categories faces, tools and scenes and one picture of the category body parts were pre-selected. From this preselection, one picture of each category was randomly chosen to create a set of four pictures for each participant.

Prospective decision making task

Participants performed a prospective decision making task which required them to maximize reward by tracking and predicting values (i.e., reward magnitudes) associated with two choice options. The values of the two options changed over a sequence of time points and participants' goal was to choose the more valuable option at the next time point.

The options were represented by four category-specific stimuli (a face, a tool, a scene, a body part). Participants were instructed that two of these stimuli each formed a pair such that stimuli within a pair yielded the same value at a given time (they were value-congruent across the entire task). For example, the stimuli face and tool might have formed a pair and the stimuli scene and body part might have formed a pair. Hence, the task comprised two latent options (i.e., options A and B), with two value-congruent stimuli per option. Value congruencies between the four pictures were counterbalanced across participants.

Each trial consisted of an observation phase and an active choice. During the observation phase, participants viewed the two options along with their changing values over a sequence of time points (TP). Pictures of the two options were displayed on the left and right sides of the screen, with their associated current values indicated by numbers underneath. Across time points, two aspects changed: First, which of the two value-congruent stimuli of a given option was shown on the screen alternated each time point (e.g., for option A TP1: face, TP2: tool, TP3: face, etc.). Participants were instructed that when a stimulus and its current value were shown on the screen, the other stimulus of the pair currently yielded the same value. Secondly, the values of the two latent options changed over time points. Participants were instructed to carefully track these changes to be able to predict the options' values at the next time point. Each time point was presented for 2.5 s and was directly followed by the next time point. During the observation phase (initial time points), each option stayed on the same side of the screen to facilitate fast tracking of the changes. However, across trials the sides of the options were counterbalanced and distributed randomly. After 3-5 observed time points, only two

pictures were presented and participants were asked to choose the option with the higher value at this future time point (choice time point). At the choice time point, the sides of the pictures on the screen (left / right) were random to prevent pure side-value associations. Participants were asked to indicate their choice by pressing the left or the right button on an MRI-compatible button box. Participants were given a maximum of 3 s to respond. After their choice (or the timeout), a fixation cross was presented at the center of the screen for an inter-stimulus interval sampled from a truncated exponential distribution (min = 3 s, max = 8 s, mu = 4 s, sampled mean = 4.1 s). Afterwards, a feedback screen was presented for 2.5 s, showing the pictures and their actual values at the choice time point. The value of the chosen option was highlighted in yellow. Lastly, a fixation cross was presented at the center of the screen for an inter-trial interval sampled from a truncated exponential distribution (min = 3 s, max = 8 s, mu = 4 s, sampled mean = 4.1 s).

The task comprised 144 trials. Half of the trials involved a choice as described above, with choices at the 4^{th}, 5^{th} or 6^{th} time point. The other half of the trials proceeded without any choice and consisted of six time points (passive trials). The purpose of including longer trials without any choice was to improve the estimation of trajectory-related fMRI signals for the planned analysis of hexadirectional signals (grid-like representation, see below). Choice and passive trials were intermixed randomly so that participants would always need to track the values in a given trial and be ready to indicate their choice.

Crucially, a sequence of time points in a trial formed a trajectory through an underlying abstract two-dimensional value space. The two dimensions of the space corresponded to the values associated with the two options (ranging from 0 to 100). Each time point with its current values of the two options corresponded to a particular location in the value space and a trial could therefore be conceptualized as movement along a trajectory through the space. Trajectories were sampled with directions (angles) ranging from 0°-350° in 10°-steps (0° referring to a value increase along the x-dimension of the space but no change in the y-dimension).

In this space, the 45°-diagonal represented locations where the two options had the same values. Trajectories crossing the 45°-diagonal therefore involved a switch in which of the two options was more valuable. Half of all trajectories in the task involved a switch while the other half did not (switch vs. non-switch trajectories). The switch time point was defined as the first time point after the 45°-diagonal. The switch time point appeared equally often at the 4^{th}, 5^{th} and 6^{th} time point across all switch trajectories (equal sampling both in choice and in passive trials / trajectories).

Time points (locations) along a trajectory were sampled equidistantly, i.e., the distance between two consecutive time points was the same within a given trajectory. The values

shown to participants during the task were rounded to integers. Depending on the direction (angle) of a trajectory, rounding could lead to differences of +/-1 in value changes between time points but it was ensured that this would not change the identity of the more valuable option at choice time points. Furthermore, the task included two types of trajectories with regard to the distance between two consecutive time points: trajectories with a relatively smaller distance of 6 (referred to as short-distance trajectories) and trajectories with a relatively larger distance of 10 (referred to as long-distance trajectories). For each distance type, one set of trajectories (36 directions, 0°-350°) was realized as choice trials and one set of trajectories as passive trials. Furthermore, for a given distance type each direction was once realized as a switch trajectory and once as a non-switch trajectory. The assignment of switch vs. non-switch to choice vs. passive trials was pseudorandom with the condition that in choice trials, each direction was realized as a switch trajectory at least once across both distance types. This ensured that a response from the participant was sampled for all directions 0°-350° involving a switch.

As noted above, choices occurred either at the 4^{th}, 5^{th} or 6^{th} time point. In non-switch trajectories, the 4^{th}, 5^{th} and 6^{th} time point equally often constituted the choice time point for each distance type. In switch trajectories, for each distance type half of the choice time points sampled the switch time point (18 trials). The other half sampled the time point before the switch (pre) and the time point after the switch (post) equally often (i.e., 50 % switch time point, 25 % pre time point, 25 % post time point; note that for 6 participants at the beginning of the study the balance between pre and post differed up to +/-3 trials).

The total of 144 trials (trajectories) was pseudorandomly distributed over four task blocks (fMRI runs) so that each block:

1. sampled all 36 directions ranging from 0°-350° in 10°-steps (hence, 36 trials per block),
2. sampled switch and non-switch trajectories equally often,
3. sampled choice and passive trajectories equally often,
4. sampled the switch time point as the choice time point equally often (one trial more in one block),
5. sampled short- and long-distance trajectories equally often
6. and for each distance type sampled choice and passive trajectories equally often.

In each block, the order of trajectories was randomized according to the direction.

In each block, trajectories were positioned within the circle created by a radius of 50 from the central point of the two-dimensional value space at [50,50]. Equal positioning of trajectories in

the relevant circular area of the space was achieved by a genetic algorithm. Its goal was to position trajectories so as to minimize the standard deviation of the number of time points (locations on trajectories) falling into the 10x10 sub-squares of the relevant circular area (for 2000 generations).

Each block lasted approx. 13 min (*M* = 12.74 min, *SD* = 0.15 min). It started with a fixation cross presented at the center of the screen for 10 s before the first trial. After the last trial, a fixation cross was presented at the center of the screen for 15 s, followed by a message informing the experimenter that the block finished and the MRI run could be stopped. After each block, participants received feedback about their performance in the given block. The feedback stated the number of correctly answered trials as well as the earned monetary bonus in the given block. More specifically, in each block a trial was randomly chosen for the bonus. If the answer in this trial was correct, the highest values across the entire trial were summed up and converted into a bonus (so that a value of 50 yielded 0.20 €). If the answer in this trial was false, no bonus was won. Participants were instructed about this bonus beforehand. Participants could take a short break before the next block.

Participants received instructions for the task and a training of 25 trials before performing the main task in the scanner. During training, incorrectly answered trials were repeated until answered correctly. For instructions and training only, a distinct set of stimuli of everyday objects from a publicly available stimulus dataset (Brady et al., 2008) was used.

At the end of the study, participants were asked about their strategies to solve the task and whether they imagined the underlying two-dimensional value space. None of the participants reported having imagined the two-dimensional value space.

The task was programmed in Python 3.7 using the PsychoPy package (Peirce et al. (2019); version 3.1.5; https://lindeloev.net/psychopy-course/) in Spyder (https://www.spyder-ide.org/; version 4.0.0b3) distributed via Anaconda (https://www.anaconda.com/; version 2019.03). The instruction was programmed using the Psycho Builder (Peirce et al. (2019); version 2020.2.3).

Picture Viewing Task (PVT)

Before the prospective decision making task, participants performed a picture viewing task (PVT) which served as an independent dataset to train a decoder for subsequent analyses. Participants viewed a stream of pictures of the category-specific stimuli which were later used as choice stimuli in the prospective decision making task. The PVT was participants' first exposure to these category-specific stimuli during the study.

To ensure that participants paid attention to the presentation of the stimuli, they performed a one-back task as a cover task. In each trial, a stimulus was presented for 2 s at the center of the screen. This was followed by a fixation cross at the center of the screen for an inter-trial

interval sampled from a truncated exponential distribution (min = 2 s, max = 8 s, mu = 3 s, sampled mean = 3.3 s). If the fixation cross was red, participants had to judge whether the stimulus in the next trial was the same as the preceding stimulus before the fixation cross (test trial). If the fixation cross was white, no judgement was required (regular trial). In test trials, participants had to indicate their judgement by pressing one of two buttons on an MRI-compatible button box if the stimulus was the same as the preceding one and the other button if it was different. Button contingencies (left vs. right button for which type of judgement) were counterbalanced and randomized across participants. Participants were instructed to press the button while the stimulus was presented (hence maximum response time of 2 s). After a test trial the task proceeded without direct trial-specific feedback.

The task consisted of 65 trials, with 14 regular trials per stimulus (+1 for one stimulus) and 2 test trials per stimulus. The sequence of trials was generated pseudorandomly so that every stimulus was preceded equally often by every other stimulus including self-repetitions (i.e., serial-order counterbalanced sequence; Brooks (2012)). Test trials were distributed pseudorandomly over the trial sequence so that every bin of 8 trials contained a test trial. Of the 2 test trials per stimulus, one trial was realized as a self-repetition trial (same-stimulus-judgement) and one as a non-self-repetition trial (different-stimulus-judgement).

The task lasted approx. 6 min (*M* = 6.18 min, *SD* = 0.02 min). The task started with a fixation cross presented at the center of the screen for 10 s before the first trial. After the last trial, a fixation cross was presented at the center of the screen for 15 s, followed by a message informing the experimenter that the task finished and the MRI run could be stopped. Afterwards, participants received feedback about their task performance. The feedback stated the number of correctly answered trials as well as the earned monetary bonus. Participants were instructed that for each correctly answered test trial they would earn a bonus of 0.15 €. Participants could take a short break after the task.

Serial-order counterbalancing of the trial sequence was performed in Matlab using a script by Brooks (2012). The task was programmed in Python 3.7 using the PsychoPy package (Peirce et al. (2019); version 3.1.5; https://lindeloev.net/psychopy-course/) in Spyder (https://www.spyder-ide.org/; version 4.0.0b3) distributed via Anaconda (https://www.anaconda.com/; version 2019.03).

Two-stage task

Participants performed the two-stage decision making task developed by Daw et al. (2011) to study model-based vs. model-free decision making. The task structure consisted of two decision stages and participants' goal was to maximize rewards obtained by decisions at the second stage. Stimuli were character symbols.

In each trial, decisions were made at two stages. At the first stage, participants had to choose between two stimuli by pressing one of two buttons on a keyboard. The chosen stimulus moved to the top of the screen. Below, one of two second-stage states was presented. The second stage consisted of two other stimuli and participants had to choose one of them by pressing one of two buttons on a keyboard. The second-stage decision was either rewarded (displayed by a coin) or not (displayed by a red X), presented on the screen for 1 s. Participants had a maximum of 3 s to indicate their decision.

Transitions to the two second-stage states depended probabilistically on the first-stage decision. One stimulus at the first stage led to one second-stage state with a higher probability of 70 % (common transition) while it led to the other second-stage state with a lower probability of 30 % (rare transition). This transition pattern was reversed for the other first-stage stimulus. At the second stage, reward probabilities were determined by a Gaussian process with a standard deviation of .025 and reflecting boundaries of .25 and .75.

Participants were instructed that one of the first-stage stimuli primarily lead to one second-stage state and vice versa and that this pattern would remain constant across the task. Furthermore, they were instructed that reward probabilities of second-stage stimuli could change and that collected rewards would be translated into a monetary bonus at the end of the task.

Participants performed a training with 50 trials and a distinct set of stimuli. The main task comprised 201 trials. At the end of the task, participants received feedback stating the earned monetary bonus based on their performance (bonus was calculated as 0.015 € per obtained reward). The task lasted approx. 30 min.

For this task, a PsychoPy-based Python script from a publicly available repository (Abraham Nunes, https://abrahamnunes.github.io/paradigms/) was used. The script was adapted in Python 3.7 using the PsychoPy package (Peirce et al. (2019); version 3.1.5) in Spyder (https://www.spyder-ide.org/; version 4.0.0b3) distributed via Anaconda (https://www.anaconda.com/; version 2019.03).

Santa Barbara Sense of Direction Scale (SBSOD)

Participants filled out the Santa Barbara Sense of Direction Scale (SBSOD) questionnaire (Hegarty, 2002) on a computer. This questionnaire measures navigational abilities and preferences and consists of 15 items (self-referential statements). Items were presented subsequently in the upper part of the screen, together with a 7-point rating scale underneath (1 = strongly agree, 7 = strongly disagree). Participants were asked to indicate their response by pressing the respective number on a keyboard and confirming their response with enter (response was self-paced).

To compute the score of the questionnaire, responses to positive items were reverse-coded so that a higher overall score reflected higher navigational abilities and preferences.

The task was programmed in Python 3.7 using the PsychoPy package (Peirce et al. (2019); version 3.1.5; https://lindeloev.net/psychopy-course/) in Spyder (https://www.spyder-ide.org/; version 4.0.0b3) distributed via Anaconda (https://www.anaconda.com/; version 2019.03).

MRI data acquisition

MRI data were recorded using a 3 Tesla Siemens Magnetom Prisma Fit scanner (Siemens, Erlangen, Germany) with a 32-channel head coil.

After a localizer scan, functional scans (fMRI) for the picture viewing task and the four runs of the prospective decision making task were acquired using T2*-weighted whole-brain gradient-echo echo planar imaging (GE-EPI) with multiband acceleration, sensitive to blood-oxygen-level-dependent (BOLD) contrast (Feinberg et al., 2010; Moeller et al., 2010). Settings of the fMRI sequence were as follows: TR = 1500 ms; TE = 22 ms; voxel size = 2.5 mm isotropic; field of view = 204 mm; flip angle = 70°; partial fourier = 0.75; bandwidth = 1794 Hz/Px; multi-band acceleration factor = 3; 69 slices interleaved; distance factor = 0 %; phase encoding direction = A-P. On average, 253 volumes were recorded for the PVT (*M* = 252.76 volumes, *SD* = 5.45 volumes) and 514 volumes per run of the prospective decision making task (*M* = 513.70 volumes, *SD* = 7.57 volumes).

After the second run of the prospective decision making task, field maps were acquired to measure and later correct for magnetic field inhomogeneities. Field maps were acquired using both opposite phase-encoded EPIs and a double echo sequence. Settings of the opposite phase-encoded EPIs were as follows: TR = 8000 ms; TE = 50 ms; voxel size = 2.5 mm isotropic; field of view = 204 mm; flip angle = 90°; partial fourier = 0.75; bandwidth = 1794 Hz/Px; multi-band acceleration factor = 1; 69 slices interleaved; distance factor = 0 %. Settings of the double echo sequence were as follows: TR = 620 ms; TE1 = 4.00 ms; TE2 = 6.46 ms; voxel size = 2.5 mm isotropic; field of view = 204 mm; flip angle = 60°; bandwidth = 412 Hz/Px; 69 slices interleaved; distance factor = 0 %.

At the end of the scanning session, a T1-weighted MPRAGE anatomical scan was acquired (TR = 2300 ms; TE = 2.98 ms; voxel size = 1 mm isotropic; field of view = 256 mm; flip angle = 9°; bandwidth = 240 Hz/Px; distance factor = 50 %).

To measure physiological noise signals during the fMRI runs, pulse oximeter data were recorded on participants' hands using a Siemens pulse sensor and the PhysioLog function of the multiband sequence.

Task stimuli were projected on a screen via a mirror attached to the head coil and behavioral responses were collected with an MRI-compatible button box.

Behavioral data analysis

We performed all behavioral analyses in Python 3.8 using Spyder (https://www.spyder-ide.org/; version 5.1.5) distributed via Anaconda (https://www.anaconda.com/; version 2020.11). Statistical analyses were based on the packages scipy (version 1.10.0) and statsmodels (version 0.13.2). T-tests and correlations tests were based on non-parametric permutation-based approaches to assess significance (10000 permutations). If not stated otherwise, we used an alpha level of .05 and two-sided tests.

Prospective decision making task performance and reaction times

We calculated performance in the prospective decision making task as the proportion of trials with a correct choice, defined as choice of the objectively more valuable option at the choice time point. We first assessed whether participants met our performance criterion of at least 70 % (based on previous piloting, see Participants) to be included in the final analysis sample. For this purpose, we left trials with missing responses labeled as incorrect (total of 50 trials with missing responses across participants). For further analyses, we labeled trials with missing responses as NaNs so that they were not considered in the analyses. Furthermore, we labeled trial scores as NaNs if both options had the same objective value at the choice time point (same value could happen due to constraints by the direction (angle) of the trajectory; M = 2.2 trials, SD = 1.19 trials across participants). We log-transformed reaction times.

In switch trajectories, we tested whether performance at the switch time point was better than expected by chance using a one-sample t-test against 50 % (and as controls also for the pre and post time point). Furthermore, we tested whether the time point in switch trajectories (pre, switch, post) influenced performance and reaction times using repeated measures ANOVAs and post-hoc pairwise tests (related-samples t-tests, with α = 0.016, Bonferroni-corrected for three comparisons).

To estimate the effect of the distance between the choice location and the 45°-diagonal of the value space on performance, we implemented a logistic regression for each participant predicting trial scores based on the distance. We then tested participant-specific effect sizes against 0 using a one-sample t-test on the group level. As a control, we repeated this analysis using only choices in switch trajectories where locations lay inherently closer to the diagonal. In both all and switch-trajectories-only analyses, one extreme outlier data point was excluded from the group level test (data point was 301.08 SD and 128.28 SD away from sample mean without that data point). We visualized the effect by showing correct and incorrect choice

locations in the value space. For reaction times, we tested the effect of the distance to diagonal using participant-specific linear regressions.

To test whether performance was influenced by short- vs. long-distance trajectories, we implemented a repeated measures ANOVA with the factors distance type and time point (pre, switch, post). We also analyzed performance for different directions of trajectories. We binned directions according to quadrants, reflecting whether values increased or decreased for both options or in opposite directions (Q1: 10-80°, Q2: 100-170°, Q3: 190-260°, Q4: 280-350° and cardinal directions of 0°,90°,180° and 270° as a separate bin). We tested whether the quadrant influenced performance using a repeated measures ANOVA.

Reinforcement learning model for prospective decision making task

We investigated whether a reinforcement learning model which captured the prospective nature of the task, i.e., the value changes over time, fitted participants' choice behavior better than a model that did not. To this end, we modified a Rescorla-Wagner model (Rescorla & Wagner, 1972). The Rescorla-Wagner model updates value estimates of choice options according to a prediction error, defined as the difference between the expected value and the received outcome. We modified the original Rescorla-Wagner model so that it updated value estimates within a trial based on prediction errors and additionally value changes over time points. We refer to this modified version as the prospective Rescorla-Wagner model. More specifically, each option's value within a trial was updated according to:

$$V_{TP+1} = V_{TP} + \alpha*(O_{TP} + C_{TP} - V_{TP}) \text{ with } C_{TP} = O_{TP} - O_{TP-1},$$

whereby V_{TP} and V_{TP+1} are values at the current and next time points, respectively, O_{TP} is the outcome at the current time point, C_{TP} reflects how the value has changed from the previous to the current time point and α is the learning rate (free parameter of the model). Value estimates of both options were translated into choices by computing the probability of each option's choice using a softmax function:

$$P_A = \frac{e^{\beta * V_A}}{e^{\beta * V_A} + e^{\beta * V_B}},$$

with P_A as the probability of choosing option A, e as the exponential, V_A and V_B as the values of options A and B (values divided by 100) and β as inverse temperature indicating the determinacy of choices (free parameter of the model). In each trial, values were initialized with the objective values of the first time point, outcomes of the second time point were received and value predictions were made for the following time points.

We fitted this prospective Rescorla-Wagner model to each participant's choice data and searched for the best-fitting estimates of the free parameters α and β by minimizing the

negative log-likelihood of the model. Parameter estimates were initially bound to ranges [0,1] for α and [0,100] for β. As we observed a ceiling effect for α, we removed its upper bound to allow estimates greater than 1.

We compared the fit of this prospective Rescorla-Wagner model to the fit of the original Rescorla-Wagner model. The original Rescorla-Wagner model does not consider value changes over time points (no prospective component):

$$V_{TP+1} = V_{TP} + \alpha*(O_{TP} - V_{TP}).$$

Notations, translation of value estimates into choice probabilities using a softmax function and model fitting were the same as described above. Parameter estimates of the original Rescorla-Wagner model were bound to ranges [0,1] for α and [0,100] for β. We compared the fits of the prospective and the original Rescorla-Wagner model by testing for a difference in the Akaike Information Criterion (AIC) using a related-samples t-test (with $\alpha = 0.01$, Bonferroni-corrected for five tests including alternative models, see below). We extracted parameter estimates of the winning model (prospective Rescorla-Wagner model). As a control, we correlated the learning rate α with performance at the switch and the pre time point using Pearson correlations.

In addition to the prospective Rescorla-Wagner model described above, we implemented four alternative control models which similarly aimed to capture the prospective nature of the task:

1. Prospective control model 1: Similar to prospective Rescorla-Wagner model described above, but C_{TP} as the option's value change is updated itself across time points with its own learning rate:

   $V_{TP+1} = V_{TP} + \alpha*(O_{TP} + C_{TP\_exp} - V_{TP})$

   with $C_{TP} = O_{TP} - O_{TP-1}$

   and $C_{TP\_exp} = C_{TP}$ for the first update within a trial, and $C_{TP\_exp} = C_{TP\_exp} + \alpha_C*C_{TP}$ afterwards.

2. Prospective control model 2: Value update with standard prediction error and an additional parameter for the value change:

   $V_{TP+1} = V_{TP} + \alpha*(O_{TP} - V_{TP}) + \delta*(O_{TP} - O_{TP-1})$

3. Prospective control model 3: Value update with standard prediction error and expected prediction error, similar to expected prediction error models in Wittmann et al. (2016):

   $V_{TP+1} = V_{TP} + \alpha*PE + PE_{exp}$

   $PE = O_{TP} - V_{TP}$

PE_{exp} = PE for the first update within a trial, $PE_{exp} = PE_{exp} + \alpha*(PE - PE_{exp})$ afterwards

4. Prospective control model 4: Similar to prospective control model 3, but PE_{exp} is updated with its own learning rate:

   $V_{TP+1} = V_{TP} + \alpha*PE + PE_{exp}$

   $PE = O_{TP} - V_{TP}$

   PE_{exp} = PE for the first update within a trial, $PE_{exp} = PE_{exp} + \alpha_{PE}*(PE - PE_{exp})$ afterwards

To allow similar parameter fits as for the prospective Rescorla-Wagner model described above, we removed the upper bound of 1 for learning rates of these control models. The prospective Rescorla-Wagner model described above fitted the data better than any of the control models (test for difference in AIC using related-samples t-tests; PC1: $t(45) = -3.58$, $p < .001$; PC2: $t(45) = -4.30$, $p < .001$; PC3: $t(45) = -6.93$, $p < .001$; PC4: $t(45) = -5.92$, $p < .001$, with $\alpha = 0.0125$, Bonferroni-corrected for four comparisons).

Picture viewing task performance

We calculated performance in the one-back cover task of the picture viewing task as the proportion of correctly answered test trials. For this purpose, we labeled trials with missing responses as incorrect. For two participants at the beginning of the study, button presses were not registered due to a technical mistake (except for the first test trial). Therefore, we could not assess performance for these participants. However, we still used their fMRI data of the picture viewing task for the fMRI analysis as the purpose of the one-back cover task was only to ensure participants' attention to the stimuli. In addition, two participants reported that they confused the buttons for the two response types (button contingencies: left or right button for same or different stimulus judgement). Indeed, their responses matched exactly the opposite pattern of all correct trial-wise responses. For this reason, we reverse-coded their responses to calculate their performance.

Two-stage task: model-based decision making

Analogously to Daw et al. (2011), we tested whether the probability of repeating a first-stage choice depended on the reward and the transition type in the preceding trial. For this purpose, we labeled each trial as 1 if participants chose the same first-stage stimulus as in the preceding trial and as 0 if not. We calculated stay percentages for the factors reward (received or not) and transition type (common or rare). Across participants, we tested whether stay percentages were influenced by reward, transition type and their interaction using a repeated measures ANOVA.

In addition, we fitted each participant's choice data using the hybrid reinforcement learning model as described in Daw et al. (2011). This model learns values by both model-based and

model-free decision algorithms. Both values are weighted by a free parameter indicating the influence of model-based values on choices (ranging from 0 for model-free to 1 for model-based). In addition, the model contains separate learning rates and inverse temperatures for the two stages as well as a perseverance parameter and an eligibility trace. We fitted this model using its implementation in the hBayesDM package (Ahn et al. (2017); version 1.1.1; model: ts_par7).

We correlated estimates of the model-based parameter of the two-stage task with the learning rate and overall performance of the prospective decision making task using Pearson correlations (with $\alpha = 0.025$, Bonferroni-corrected for two tests).

Santa Barbara Sense of Direction Scale (SBSOD) correlations

We correlated scores of the SBSOD with the learning rate and overall performance of the prospective decision making task using Pearson correlations (with $\alpha = 0.025$, Bonferroni-corrected for two tests).

MRI preprocessing

We performed all MRI analyses (preprocessing and main analyses) in Python 3.8 using Spyder (https://www.spyder-ide.org/; version 5.1.5) distributed via Anaconda (https://www.anaconda.com/; version 2020.11). MRI analyses were mainly based on the packages nilearn (version 0.9.0), nibabel (version 3.2.1), scikit-learn (version 1.0.1) as well as FSL (version 6.0.3), ANTS (version 2.3.5) and tools stated below. Statistical analyses were based on the package scipy (version 1.10.0) and statsmodels (version 0.13.2). T-tests and correlations tests were based on non-parametric permutation-based approaches to assess significance (10000 permutations). If not stated otherwise, we used an alpha level of .05 and two-sided tests.

Conversion of MRI data to the Brain Imaging Data Structure (BIDS) standard

We converted DICOM files of the MRI scanner to NIfTI files and reorganized them according to the BIDS standard (K. J. Gorgolewski et al., 2016) using the tool dcm2bids (version 2.1.6, https://unfmontreal.github.io/Dcm2Bids/). Furthermore, we removed facial structure in the anatomical scan using the tool pydeface (version 2.0.0, https://github.com/poldracklab/pydeface) to further anonymize the data. For the functional task runs, we created event files specifying the details of the task events during the given fMRI run.

Preprocessing by fMRIPrep

Results included in this manuscript come from preprocessing performed using *fMRIPrep* 20.2.6 (Esteban et al., 2018; Esteban et al., 2022; RRID:SCR_016216), which is based on *Nipype* 1.7.0 (Esteban et al., 2022; Gorgolewski et al., 2011; RRID:SCR_002502).

Anatomical data preprocessing

A total of 1 T1-weighted (T1w) images were found within the input BIDS dataset. The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al., 2010), distributed with ANTs 2.3.3 (Avants et al., 2008, RRID:SCR_004757), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a *Nipype* implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, RRID:SCR_002823, Zhang et al., 2001). Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, RRID:SCR_001847, Dale et al., 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (RRID:SCR_002438, Klein et al., 2017). Volume-based spatial normalization to two standard spaces (MNI152NLin2009cAsym, MNI152NLin6Asym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following templates were selected for spatial normalization: *ICBM 152 Nonlinear Asymmetrical template version 2009c* [Fonov et al., 2009, RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym], *FSL's MNI ICBM 152 non-linear 6th Generation Asymmetric Average Brain Stereotaxic Registration Model* [Evans et al., 2012, RRID:SCR_002823; TemplateFlow ID: MNI152NLin6Asym],

Functional data preprocessing

For each of the 5 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. A B0-nonuniformity map (or *fieldmap*) was estimated based on two (or more) echo-planar imaging (EPI) references with opposing phase-encoding directions, with 3dQwarp Cox & Hyde (1997) (AFNI 20160207). Based on the estimated susceptibility distortion, a corrected EPI (echo-planar imaging) reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration (Greve & Fischl, 2009). Co-registration was configured with six degrees of freedom. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9, Jenkinson et al., 2002). BOLD runs were slice-time corrected to 0.708s (0.5 of slice acquisition range 0s-1.42s) using 3dTshift from AFNI 20160207 (Cox & Hyde (1997), RRID:SCR_005927). The BOLD time-series were resampled onto the following surfaces

(FreeSurfer reconstruction nomenclature): *fsnative*, *fsaverage*. The BOLD time-series (including slice-timing correction when applied) were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. These resampled BOLD time-series will be referred to as *preprocessed BOLD in original space*, or just *preprocessed BOLD*. The BOLD time-series were resampled into standard space, generating a *preprocessed BOLD run in MNI152NLin2009cAsym space*. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Automatic removal of motion artifacts using independent component analysis (ICA-AROMA, Pruim et al. (2015)) was performed on the *preprocessed BOLD on MNI space* time-series after removal of non-steady state volumes and spatial smoothing with an isotropic, Gaussian kernel of 6mm FWHM (full-width half-maximum). Corresponding "non-aggresively" denoised runs were produced after such smoothing. Additionally, the "aggressive" noise-regressors were collected and placed in the corresponding confounds file. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al. (2014)) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al. (2002)). FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* (following the definitions by Power et al. (2014)). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (*CompCor*, Behzadi et al. (2007)). Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. (2007) in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's *aseg* segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the *k* components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the

correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al., 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e., head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964). Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

Many internal operations of *fMRIPrep* use *Nilearn* 0.6.2 (Abraham et al. (2014), RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIPrep*'s documentation (https://fmriprep.org/en/latest/workflows.html).

MRI data quality checks

We assessed fMRI data quality based on measures of head motion as a potential source for noise and artifacts. We investigated framewise displacement for each run and participant and marked each volume as an outlier if it exceeded a threshold of 0.5 mm (criterion used by fMRIPrep). On average, motion was relatively low (mean framewise displacement across participants: M = 0.15 mm, SD = 0.04 mm, range = 0.07 – 0.25 mm; mean percentage of outlier volumes: M = 0.91 %, SD = 1.34 %, range = 0 – 6.09 %; all participants below our criterion of max. 10 % outlier volumes for inclusion in the main data analyses). To control for head motion, we included motion parameters as confounds in first-level GLMs (see below).

Additionally, we assessed MRI data quality using the tool MRIQC (version 0.16.1) which calculates a set of image quality metrics for both functional and anatomical image data.

Region of interest (ROI) definition

For our hypothesis of a grid-like representation in the entorhinal cortex, we used participant-specific bilateral entorhinal cortex masks created by FreeSurfer segmentations of the participants' anatomical images during preprocessing with fMRIPrep (FreeSurfer labels 1006 & 2006, M = 269 voxels, SD = 42 voxels). For small volume correction within the entorhinal cortex on the group level, we combined both participant-specific anatomy and MNI standard atlas labeling. For this purpose, we first transformed the participant-specific masks to MNI standard space and created the union of all masks across participants. We then intersected this union mask with the entorhinal cortex mask of the Juelich Histological Atlas provided by

FSL and thresholded at 50% probability. Finally, we intersected this mask with the whole-brain group mask comprising only voxels shared across participants (resulting mask used for small volume correction: 411 voxels; Supplementary Fig. 4a).

For our choice decoding hypothesis, we leveraged neural responses to category-specific stimuli (faces, tools, scenes, body parts) in category-selective regions of the occipital-temporal cortex. We created participant-specific occipital-temporal ROI masks as follows. First, we thresholded occipital and temporal lobe probability masks of the MNI Structural Atlas provided by FSL (version 6.0.3) at a threshold of 25 % and created their union. We then transformed this MNI-based mask to each participants' native space using ANTS (version 2.3.5) and resampled it to the resolution of the functional data based on transformation files created during preprocessing with fMRIPrep. We intersected these with participant-specific gray matter masks. For this purpose, we thresholded gray matter probability masks created by fMRIPrep's segmentation of the anatomical image at a threshold of 50 % and resampled them to the functional resolution. In the decoding analysis, we used these participant-specific gray matter occipital-temporal masks for additional feature selection based on univariate stimulus-category effects in the PVT training data (see below). The final masks used for choice decoding comprised 2235 voxels on average (*SD* = 173 voxels; Fig. 5a).

fMRI data analysis

General set-up of first level general linear models (GLMs)

For our fMRI data analyses, we used both univariate and multivariate approaches. For both approaches, we modeled the fMRI data using event-related GLMs. In the following, we briefly describe commonalities of GLMs across analyses.

We implemented run-wise first level GLMs using the FirstLevelModel class of the nilearn package. GLMs were computed within a brain mask (either in participants' native space or in MNI standard space, stated for each analysis below). To create a common brain mask for all runs, we resampled the anatomical brain mask in native or MNI space created during preprocessing with fMRIPrep to the resolution of the functional data. Task-related regressors in the GLMs were convolved with the Glover haemodynamic response function (HRF). Temporal autocorrelation in the fMRI data was accounted for using an autoregressive AR(1) model. For univariate analyses, the data were spatially smoothed with a 6 mm full-width at half maximum Gaussian filter (FWHM). For the multivariate choice decoding analysis, no smoothing was applied to preserve differences between voxels.

All GLMs included the following regressors for task-related events of no interest: two regressors for left and right button presses with a stick duration as well as a regressor modeling the end-of-block notification screen at the end of a run. To control for noise signals

in the fMRI data, the GLMs included 37 confound regressors estimated during preprocessing with fMRIPrep. Following the denoising strategy proposed by Satterthwaite et al. (2013), these confounds included 24 motion parameters (6 basic translation / rotation parameters, 6 temporal derivatives of these and 12 quadratic terms of the basic parameters and their derivatives) as well as 12 global signal parameters (3 basic average CSF, WM and global signal parameters, 3 temporal derivatives of these and 6 quadratic terms of the basic parameters and their derivatives). Additionally, the confounds included framewise displacement as a summary metric of frame-to-frame head motion. Furthermore, the GLMs included discrete cosine-basis regressors estimated by fMRIPrep to account for temporal low-frequency signal drifts.

Analysis of hexadirectional signals (grid-like representation)

To investigate whether the entorhinal cortex encodes the abstract value space using a grid-like representation, we implemented the hexadirectional analysis approach by Doeller et al. (2010). Grid cells in the entorhinal cortex are characterized by their regular hexagonal firing pattern which translates to hexadirectional activity modulations during navigation in fMRI, with higher activity for navigation in directions aligned with the putative grid orientation (phase of the hexadirectional signal) than for trajectories misaligned with the putative grid orientation. The analysis consists of two steps: In the first step, the grid orientation is estimated and in the second step the prediction of hexadirectional modulation according to the grid orientation is tested using independent data. Here, we tested for such a hexadirectional modulation as a function of trajectories through our value space. We implemented a cross-validation procedure, estimating the putative entorhinal grid orientation using three of four task runs and testing for a hexadirectional modulation aligned to the orientation in the left-out test run (based on Doeller et al., 2010; Nau et al., 2018).

We implemented this cross-validation procedure on fMRI data in participants' native space to enable estimations of grid orientations in participant-specific entorhinal cortex ROIs.

In the estimation set (three of four runs, GLM1), the GLM for each run included a main effect regressor modeling trajectories including all time points and a main effect regressor modeling feedback periods. The regressors were modeled with the actual onset and durations of the events during the task. The trajectory regressor was accompanied by two parametrically modulated regressors. These modulations reflected the sine and cosine of the direction (angle) θ of the trajectory with 60° (6-fold) periodicity ($\sin(6*\theta_t)$ and $\cos(6*\theta_t)$). Values for both regressors were demeaned. Effect sizes of the regressors were averaged across runs of the estimation set (fixed effects). We then used the effect sizes of the sine (β_{sin}) and cosine (β_{cos}) regressors to estimate the grid orientation in 60°-space (range [0,60°]) in each voxel of the entorhinal cortex (see ROI definition) as follows:

$$\Theta = \frac{arctan\left(\frac{\beta_{sin}}{\beta_{cos}}\right)}{6}$$

Subsequently, we calculated the mean orientation across voxels of the entorhinal cortex with a weighting of the voxels by their amplitude of the hexadirectional modulation ($\sqrt{\beta_{sin}^2 + \beta_{cos}^2}$) (Stangl et al., 2017). For this purpose, we first transformed voxel orientations back to 360°-space to allow for calculations of trigonometric functions (multiplication by 6). We then transformed these orientations and the amplitudes from polar to cartesian coordinates and took the mean separately for both dimensions. Afterwards, we transformed the mean back to polar coordinates and subsequently transformed the mean orientation back to 60°-space.

In the independent test set (left-out run, GLM2), the GLM included a main effect regressor modeling trajectories including all time points and a main effect regressor modeling feedback periods. The regressors were modeled with the actual onset and durations of the events during the task. The trajectory regressor was accompanied by a parametrically modulated regressor reflecting a six-fold (hexadirectional) sinusoidal modulation based on the mean entorhinal grid orientation ($\cos(6*(\theta_t - \Theta))$). Values for the regressor were demeaned. Effect sizes of the parametric cosine regressor were averaged across the four cross-validation folds (fixed effects) to obtain an overall effect size.

For group level statistics, we first transformed effect size images of the parametric cosine regressor to MNI standard space. We then performed an analysis with small volume correction based on our a priori ROI of the entorhinal cortex (see ROI definition). Additionally, we performed a whole-brain analysis based on a whole-brain group mask comprising only voxels shared across participants. We tested significance across participants using non-parametric permutation testing implemented in FSL Randomise with 10000 permutations. We used one-sided tests as the predicted direction of the hexadirectional effect is inherently positive (higher activity for navigation in directions aligned vs. misaligned with the grid orientation). We used threshold-free cluster enhancement and corrected for multiple comparisons with family-wise error rate ($p_{FWE} < 0.05$) within the small volume correction mask and whole-brain. For exploration of whole-brain effects at an uncorrected threshold of $p < .001$, we extracted cluster information using nilearn and respective brain region labels of the Harvard-Oxford Cortical Structural Atlas, Harvard-Oxford Subcortical Structural Atlas and Juelich Histological Atlas using FSL atlasquery.

To visualize the hexadirectional effect in the significant entorhinal cortex cluster, we implemented an additional GLM for the test set (left-out run, GLM2) by binning trajectories

based on directions. To this end, we sorted trajectories into bins of 30° based on the mean entorhinal grid orientation (+/- 15° of the grid orientation and multiples of 60°). This resulted in 12 trajectory bin regressors, 6 reflecting trajectories aligned and 6 misaligned with the grid orientation. In this GLM, we therefore modeled trajectories using the 12 bin regressors and a main effect regressor for all trajectories capturing the mean. Effect size images were averaged across the four cross-validation folds (fixed effects) and transformed to MNI standard space. We extracted the mean effect size of each trajectory bin in the significant cluster.

In control analyses, we investigated the relationship of the hexadirectional effect with the spatial and temporal stability of voxel-wise grid orientations in the significant entorhinal cortex cluster. Spatial stability refers to similarity of orientations across voxels within the significant cluster. To investigate spatial stability, we first transformed effect size images of the sine (β_{sin}) and cosine (β_{cos}) regressors of the estimation GLM (GLM1) based on all runs (to increase power) to MNI standard space. We then estimated voxel orientations as described above. For each participant, we tested deviation from a uniform distribution of voxel orientations in the significant cluster using a Rayleigh test for non-uniformity of circular data (implemented in the package pycircstat, version 0.0.2, https://github.com/circstat/pycircstat). Across participants, we calculated a Pearson correlation between the Rayleigh z-statistic and the hexadirectional effect. To control for similarity of voxels introduced by smoothing, we additionally investigated spatial stability using unsmoothed data. Temporal stability refers to similarity of orientations within a voxel across time. To investigate temporal stability, we followed the logic of the cross-validation procedure described above and additionally estimated orientations in the left-out test run. For each voxel and for each cross-validation fold, we calculated the orientation difference between the estimation and the test set. Subsequently, we averaged orientation differences across folds and classified voxels as stable if their mean orientation difference was within 15°. Across participants, we tested whether the percentage of stable voxels was different from 50% using a one-sample t-test. Furthermore, we calculated a Pearson correlation between the percentage of stable voxels and the hexadirectional effect.

In addition to the small volume correction analysis, we conducted a complementary ROI analysis based on participants' individual entorhinal Freesurfer masks. In this ROI analysis, we also investigated the specificity of a hexadirectional (6-fold) modulation of activity in line with grid cell firing by performing control analyses for a four-, five-, seven- and eight-fold modulation (same cross-validation procedure as described above). On the group level, we tested whether effect sizes were different from 0 using one-sample t-tests (with α = 0.01, Bonferroni-corrected for five tests).

We assessed the relationship between the hexadirectional effect in the significant entorhinal cluster and overall task performance using a Pearson correlation.

Lastly, we performed exploratory analyses to investigate the relationship between the entorhinal grid system and the underlying value space. First, we investigated clustering of orientations in the significant entorhinal cluster. To this end, we estimated each participant's mean orientation in the significant cluster as described above (based on all runs to increase power). We tested whether orientations across participants cluster around 45° using a V-Test (implemented in the package astropy, version 5.0, The Astropy Collaboration et al. (2022)). Secondly, we tested whether the magnitude of the hexadirectional modulation differed between high- and low-value-areas of the value space. For this purpose, we performed a median split of trajectories according to their mean value. This meant contrasting trajectories in the lower left triangle of the space (low-value-area) with trajectories in the upper right triangle of the space (high-value-area). We note that this median split led to a substantial reduction of available trajectories per value condition and an unbalanced sampling of directions between the conditions, rendering this analysis less robust. To examine sampling of directions, we counted the frequency of directions per condition per participant and tested for differences using a repeated measures ANOVA across participants with the factors direction and value condition. We repeated the cross-validated hexadirectional analysis described above, with two changes: First, we based this analysis on the significant entorhinal cluster, both for the estimation of the grid orientation (GLM1) and for testing the hexadirectional effect on the group level (ROI analysis). Note that this analysis is still unbiased as we were interested in the difference of the hexadirectional effect between value conditions. Secondly, the GLMs for the estimation and test set estimated effects separately for the value conditions. More specifically, the estimation set (GLM1) included separate main effect and sine- and cosine-parametrically modulated regressors for each value condition and the grid orientation was estimated separately for each value condition. Analogously, the independent test set (left-out run, GLM2) included separate main effect and cosine-parametrically modulated regressors for each value condition. We averaged effect sizes across voxels of the ROI (significant cluster of the overall hexadirectional effect). Across participants, we tested for a difference between value conditions using a related-samples t-test as well as for individual effects using one-sample t-tests (one-sided).

Value difference analysis

To investigate whether fMRI activity is modulated by the value difference between options during choices, we implemented a GLM with three main effect regressors: one regressor modeled the observation phase (initial time points) of the trajectories, one regressor modeled choice time points and one regressor modeled feedback periods. The regressors were modeled with the actual onset and durations of the events during the task. The choice time point regressor was accompanied by two parametrically modulated regressors. These

modulations reflected the value of the chosen option and the value of the unchosen option, as estimated by the prospective Rescorla-Wagner model. Values for both regressors were demeaned so that they were orthogonal to the main effect regressor. We then contrasted the estimated effect sizes of the chosen value vs. the unchosen value regressor [1, -1] to test for a modulation of activity by the value difference. Contrasts were averaged across runs (fixed effects).

To investigate whether fMRI activity is modulated specifically by the prospective component of the value difference, we changed the two parametrically modulated regressors for the choice time points as follows: One regressor reflected the difference in the value estimate of the chosen option between the prospective Rescorla-Wagner model and the original Rescorla-Wagner model (non-prospective). Analogously, the other regressor reflected the difference in the value estimate of the unchosen option between the prospective Rescorla-Wagner model and the original Rescorla-Wagner model (non-prospective). We excluded one participant from this analysis because the value estimates of the two models were very similar (mean difference = 0.16, participant with lowest learning rate in the prospective Rescorla-Wagner model).

In a control analysis, we added an additional parametrically modulated regressor for choice time points reflecting reaction time. Reaction times were log-transformed and demeaned.

We computed these GLMs on fMRI data in MNI standard space. For group level statistics, we performed whole-brain analyses based on a whole-brain group mask comprising only voxels shared across participants. We tested the significance of contrasts across participants using non-parametric permutation testing implemented in FSL Randomise with 10000 permutations. We used threshold-free cluster enhancement and corrected for multiple comparisons with family-wise error rate ($p_{FWE} < 0.05$). We extracted cluster information using nilearn and respective brain region labels of the Harvard-Oxford Cortical Structural Atlas, Harvard-Oxford Subcortical Structural Atlas and Juelich Histological Atlas using FSL atlasquery.

Furthermore, we visualized the value difference effect in the significant vmPFC cluster by showing time courses of the effects of the chosen and unchosen value time-locked to the onset of the choice time points. For this purpose, we extracted the preprocessed fMRI time series of voxels in the vmPFC cluster. Analogously to general first-level modeling, we spatially smoothed (6 mm FWHM) and cleaned the data by regressing out confounds and temporal low-frequency signal drifts. We z-scored each voxel's time series, averaged them across voxels of the cluster and interpolated the signal (cubic spline interpolation). For each choice time point, we extracted the cluster signal in a time window of 16s, time-locked to 1s before onset of the choice time point in steps of 0.1s until 15s after onset. Subsequently, we ran a

linear regression across choice time points of a run for each time step (in steps of 0.1s), with the regressors chosen value, unchosen value, trial number and an intercept. Value and trial regressors were demeaned beforehand. We extracted effect sizes of the chosen and unchosen value regressor for all time steps and averaged them across runs for a given participant. Lastly, we averaged these time courses across participants for visualization.

We tested for a relationship between the value difference effect in the vmPFC cluster and task performance using a Pearson correlation.

Choice decoding analysis

To investigate whether occipital-temporal cortex represents the high-value option more strongly than the low-value option during choices, we implemented the following decoding analysis. Using independent data from the picture viewing task (PVT) which took place before the prospective decision making task, we trained a decoder (support vector classifier) on occipital-temporal cortex voxels to distinguish neural activation patterns of the four category-specific stimuli (faces, tools, scenes, body parts). We then applied this decoder to neural activation patterns of choice time points in the prospective decision making task. We performed this analysis in participants' native space.

To estimate neural activation patterns of stimuli in the PVT training data, we implemented a Least-Squares Separate GLM approach. More specifically, we ran 57 single-trial-GLMs, one for each regular trial of the task. Each GLM included one regressor modeling the trial of interest and one regressor modeling all other regular trials. Test trials were modeled in a separate regressor. The regressors were modeled with the actual onset and durations of the events during the task. We used z-scores of the trial regressors for the next steps (56 z-scores, the first trial was discarded to allow for balanced sampling of stimulus categories: 14 trials per category).

Based on the PVT training data, we created the final participant-specific ROI masks used for the decoding analysis. We combined the predefined anatomical gray matter occipital-temporal masks with the functional PVT data to select category-stimuli-responsive voxels (features). To this end, we extracted trial-wise z-scores for each voxel within the predefined anatomical mask. We z-standardized them across trials and performed univariate feature selection by computing ANOVA F-values between each feature and the trial labels. We selected those 20 % of the voxels with the highest F-values. The resulting masks were used for the decoding analysis in the next steps.

As a control, we first examined how well we could decode stimulus category within the PVT, before applying the decoder to the decision making task. For this purpose, we extracted trial-wise z-scores for each voxel within the decoding ROI mask. We implemented a 7-fold cross-

validation scheme with 8 left-out test trials (2 trials per category) and 48 training trials. We trained a decoder to distinguish neural activation patterns of the four category-specific stimuli and to predict labels for the left-out test trials (support vector classifier, regularization parameter C = 1.0, kernel = rbf). We z-standardized trial-wise z-scores within the training set and applied the standardization parameters to the test trials. We assessed accuracy as the proportion of correctly predicted trials, averaged over cross-validation folds. Note that we conducted this analysis using the decoding ROI mask based on previous univariate feature selection across all trials of the PVT (see above). We did so as this was the mask used for the following across-task decoding (PVT to prospective decision making task) and the stimulus category decoding within the PVT served only as a control.

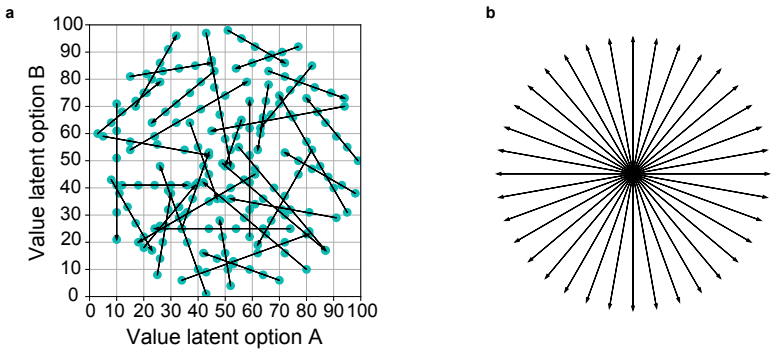
Next, we aimed to investigate stimulus representations during choices in the prospective decision making task. To first estimate neural activation patterns of choice time points in the prospective decision making task, we implemented the following GLM. One regressor modeled the observation phase (initial time points) of the trajectories and one regressor modeled feedback periods. Each choice time point (18 choice time points per run) was modeled in a separate regressor. The regressors were modeled with the actual onset and durations of the events during the task. We used z-scores of the choice time point regressors as test data for the across-task decoding analysis.

For each voxel within the decoding ROI mask, we extracted trial-wise z-scores of the PVT as training data and choice z-scores of the prospective decision making task as test data. We z-standardized the data run-wise. We then trained a decoder to distinguish neural activation patterns of the four category-specific stimuli based on the PVT data (support vector classifier, regularization parameter C = 1.0, kernel = rbf, probability = True to enable probability estimates). Subsequently, we applied this decoder to the neural activation patterns of choices in the prospective decision making task. More specifically, we extracted the probabilities which the decoder assigned to each of the four stimuli and computed two difference scores for each choice. First, we compared the probabilities assigned to the two stimuli presented on-screen during choice: probability of the stimulus with the objectively higher value vs. probability of the stimulus with the objectively lower value. Secondly, we compared the probabilities assigned to the two value-congruent stimuli which were not presented on-screen during choice (but during the time point before): probability of the congruent high-value stimulus vs. the congruent low-value stimulus. To compare these difference scores against chance level performance of the decoder, we implemented a permutation test, repeating this procedure 1000 times with randomly permuted trial labels in the PVT training data. For each choice, we then converted the original difference scores to z-scores based on the null distribution generated by the permutations. Lastly, we averaged z-scores across choices to obtain two summary scores per

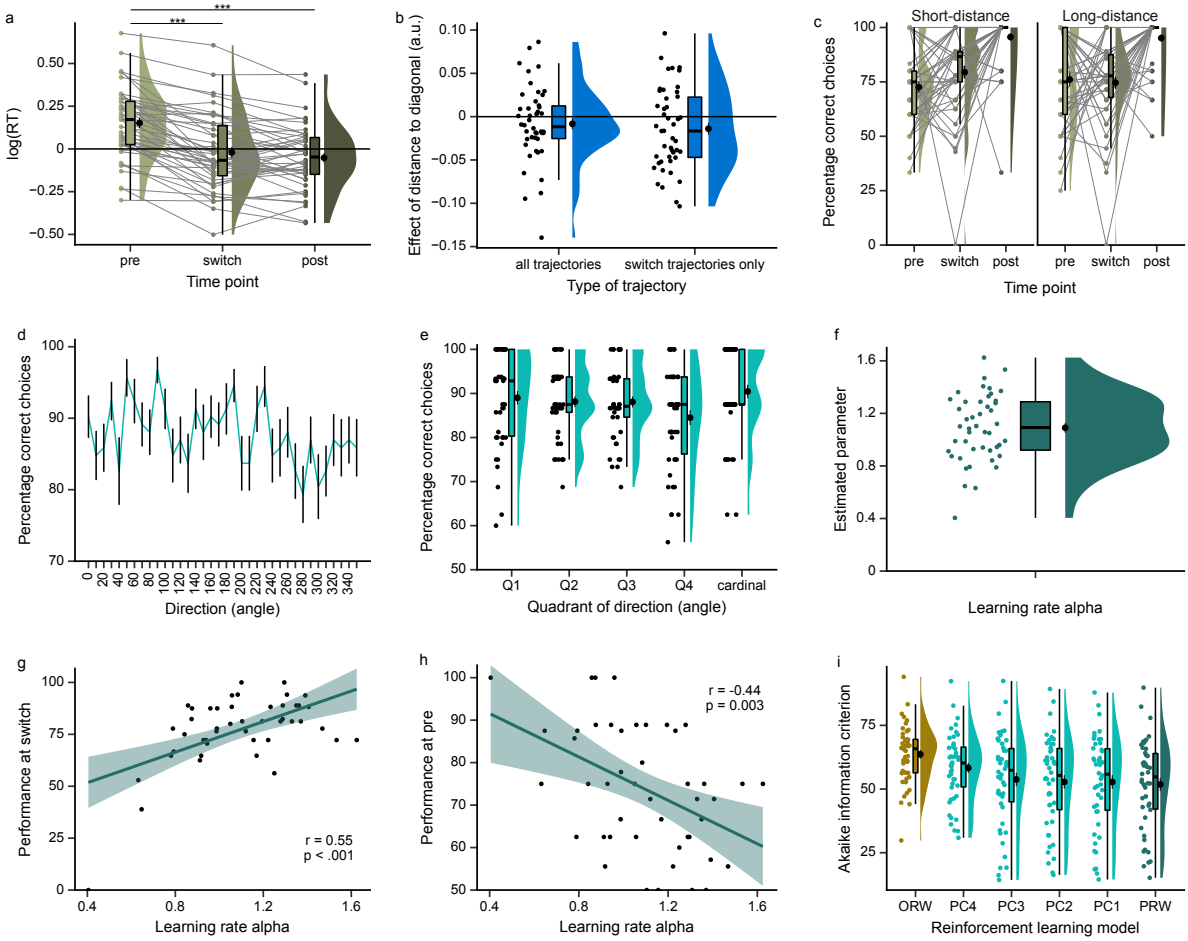
participant. On the group level, we tested participant-specific z-scores against 0 using one-sample t-tests. Furthermore, we calculated Pearson correlations between the z-scores and task performance. In addition, we tested whether the probability differences in favor of the high-value option depended on the distance between the choice location and the 45°-diagonal of the value space using participant-specific linear regressions. We then tested participant-specific effect sizes against 0 using a one-sample t-test on the group level. Furthermore, we calculated Pearson correlations between the probability differences in favor of the high-value option and the hexadirectional modulation effect in the significant entorhinal cluster.

When comparing on-screen and congruent off-screen stimuli separately, the temporal proximity of their presentations during time points within a trajectory might render disentangling their effects difficult. To control for the temporal proximity to some extent, we repeated the analysis using only those choices which sampled the switch time point as a control. In this case, the direction of the effect during choice (high-value vs. low-value, especially for the comparison of the congruent stimuli) should be different from the direction of the effect at the time point before the switch (pre).

Supplementary figures

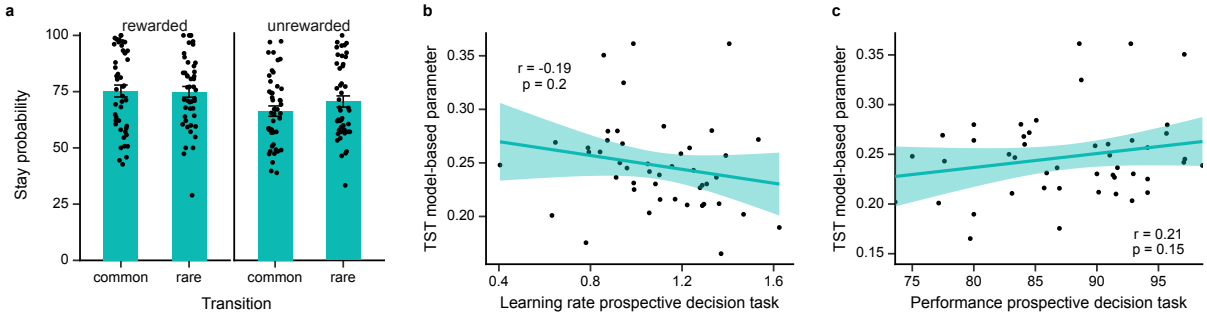


Supplementary Fig. 1 | Trajectories through the value space. a Example of trajectories through the value space in a task block. Arrows depict trajectories, and dots along trajectories depict the time points. **b** In each task block, directions of trajectories were sampled homogeneously from 0°-350° in 10°-steps.

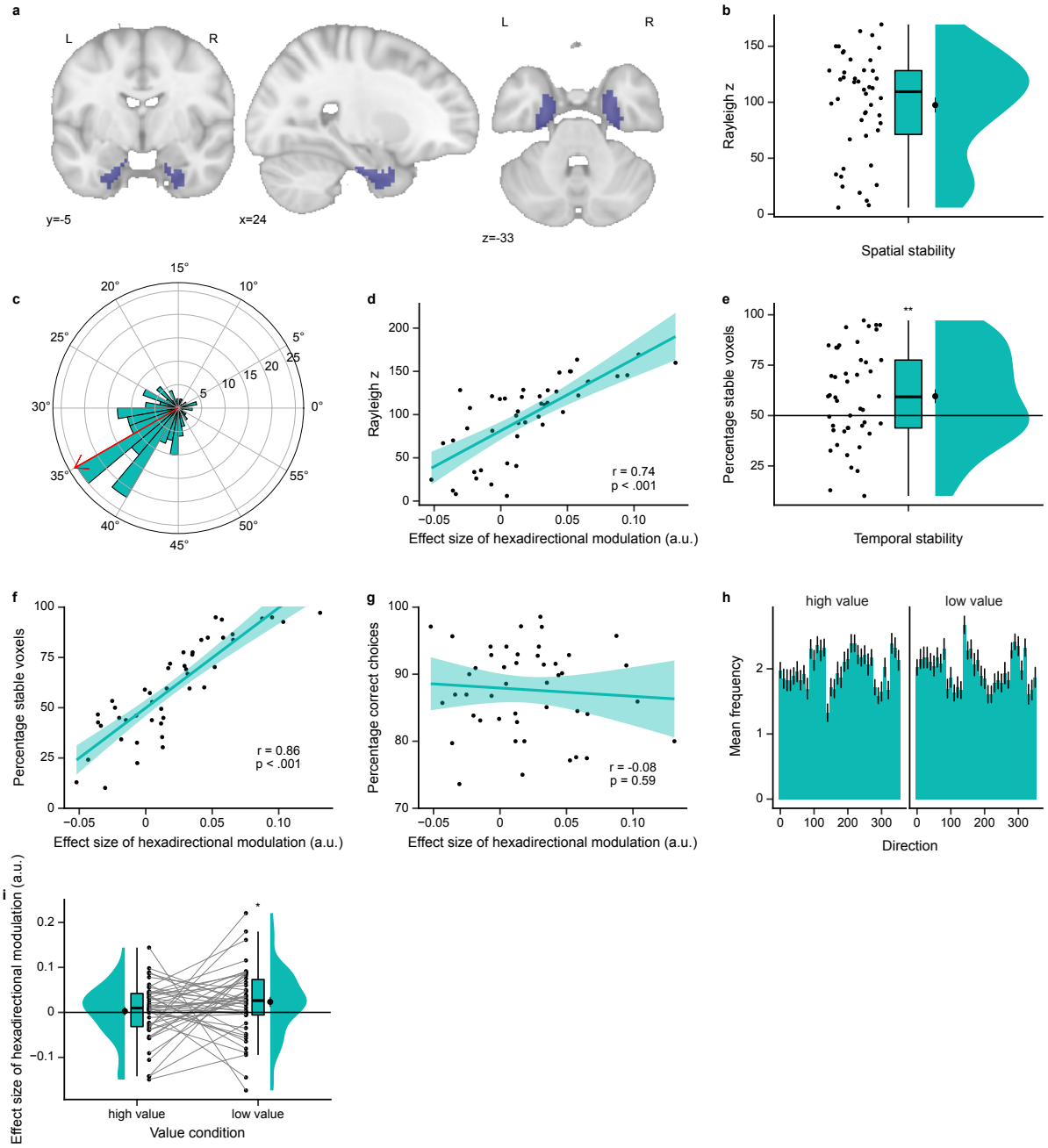


Supplementary Fig. 2 | Participants integrate and extrapolate value changes for prospective choices. a Reaction times (log) along different time points on switch trajectories: time point before the switch (pre), time point of the switch (switch) and time point after the switch (post). There was a significant effect of time point on reaction times ($F(2,90) = 60.65$, $p < .001$; post hoc pairwise tests: pre-post: $t(45) = 9.85$, $p < .001$, switch-post: $t(45) = 1.74$, $p = .09$; pre-switch: $t(45) = 8.33$, $p < .001$; with $\alpha = 0.016$, Bonferroni-corrected for three comparisons). **b** Effect of the distance between the choice location and the diagonal of the value space on reaction times. Depicted are effect sizes estimated by a linear regression, separately for including all trajectories (left) as well as including only switch trajectories as a control (right). There were no significant effects (all trajectories: $t(45) = -1.29$, $p = .20$; switch trajectories: $t(45) = -2.01$, $p = .051$). **c** Performance along different time points on switch trajectories (analogously to **a**), separately for the two types of trajectories with regard to the distance between two consecutive time points (see Methods): trajectories with a relatively smaller distance between time points (short-distance trajectories) or a relatively larger distance (long-distance trajectories). We reasoned that variation in the distance between time points would place different demands on extrapolating values. Previous piloting work indicated performance differences between these two distance levels, with higher performance at the switch and lower performance at the pre time point in short-distance compared to the long-distance trajectories. Contrary to our expectations based on previous piloting, we did not observe a significant interaction between distance type and time points (interaction effect: $F(2,90) = 1.40$, $p = .25$; main effect of distance type: $F(1,45) = 0.16$, $p = .69$; direct comparison of the switch time point: $t(45) = 1.69$, $p = .097$). **d** Performance for different directions (angles) of trajectories. Line depicts mean performance and error bars correspond to the standard error of the mean. **e** Performance for different quadrants of directions (angles). Q1 refers to directions 10°-80°, Q2 to 100°-170°, Q3 to 190°-260°, Q4 to 280°-350° and cardinal refers to directions 0°, 90°, 180° and 270°. We observed an effect of quadrant, potentially driven by slightly higher performance for cardinal directions and slightly lower performance in Q4 ($F(4,180) = 2.93$, $p = .02$). **f** Estimated parameters of the learning rate α of the prospective Rescorla-Wagner model. A learning rate of 1 reflects full updating of values according to prediction errors and value changes, learning rates above 1 hence suggest slight over-updating. **g** Significant positive correlation between the learning rate α of the prospective Rescorla-Wagner model and performance at the switch time point ($r(44) = .55$, $p < .001$; after exclusion of outlier with performance=0: $r(44) = .41$, $p = .005$). **h** Significant negative correlation between the learning rate α of the prospective Rescorla-Wagner model and performance at the pre time point. **i** Reinforcement learning model comparison for all models including alternative control models. ORW refers to the original Rescorla-Wagner (RW) model and PRW refers to the prospective Rescorla-Wagner (RW) model as depicted in Fig. 2 of the main text. PC

refers to prospective control models and the numbers correspond to numbers of these control models in the Methods section. Depicted is the Akaike information criterion (AIC), with lower values indicating better model fit. Models are ordered according to their mean AIC from left to right. The prospective Rescorla-Wagner model (PRW) fitted the data better than any control model (all $p < .001$, see Methods). **a,b,c,e,f,i** Raincloud plots: dots represent participants' data points; boxplots show median and upper/ lower quartile with whiskers extending to the most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bar corresponds to mean ± SEM; distributions depict probability density function of data points. **g,h** Correlation plots: dots represent participants' data points; line represents linear regression line, with shaded regions as the 95% confidence interval. *** $p < .001$

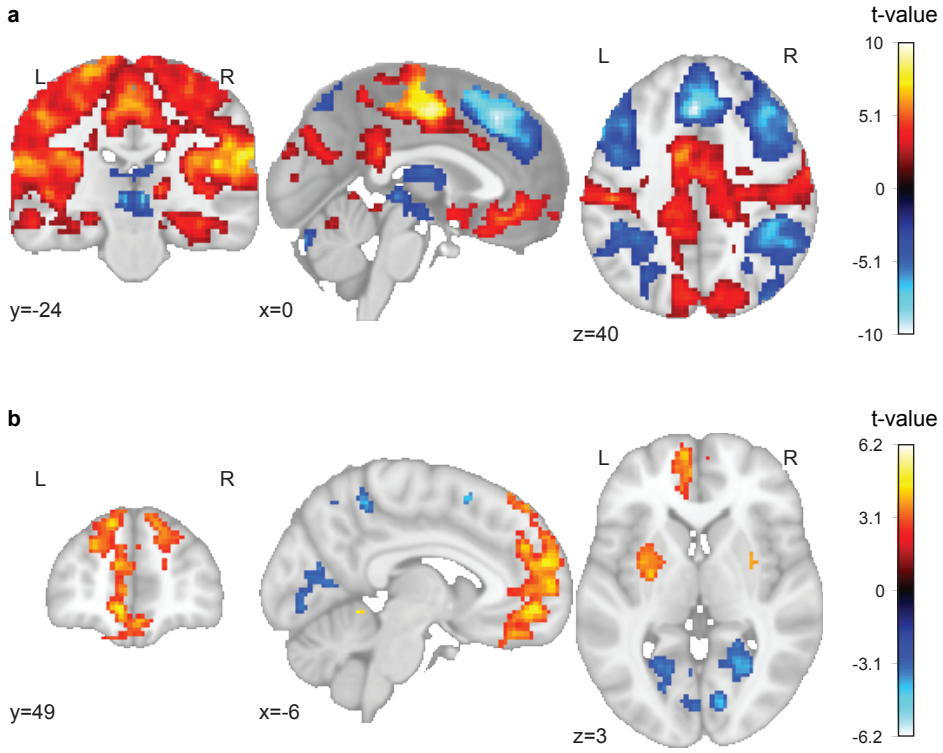


Supplementary Fig. 3 | Correlation of the prospective decision making task with the two-stage task. a Two-stage task: Probability of repeating a first-stage choice (staying) as a function of reward and transition type in the preceding trial. As expected, we observed a significant effect of reward ($F(1,45) = 20.50$, $p < .001$) and a significant interaction of reward and transition type ($F(1,45) = 4.38$, $p = .04$; main effect transition type: $F(1,45) = 3.37$, $p = .07$). In addition, we fitted choice behavior with the hybrid reinforcement learning model as described in Daw et al. (2011) to extract a model-based parameter estimate per participant (y-axis in **b-c,** parameter ranging from 0=model-free to 1=model-based). To our surprise, model-based behavior in our sample was less pronounced than typically observed, given the relatively smaller interaction effect of reward and transition type as well as lower estimates of the model-based parameter compared to Daw et al. (2011). We speculate that this might be partly attributed to participants' exhaustion after the MRI session when performing the task at the end of the study. **b-c** We did not observe any significant correlations between the model-based parameter and the learning rate or performance in our prospective decision making task (with $\alpha = 0.025$, Bonferroni-corrected for two tests). We speculate that the restricted range of the model-based parameter might have contributed to the lack of a correlation.

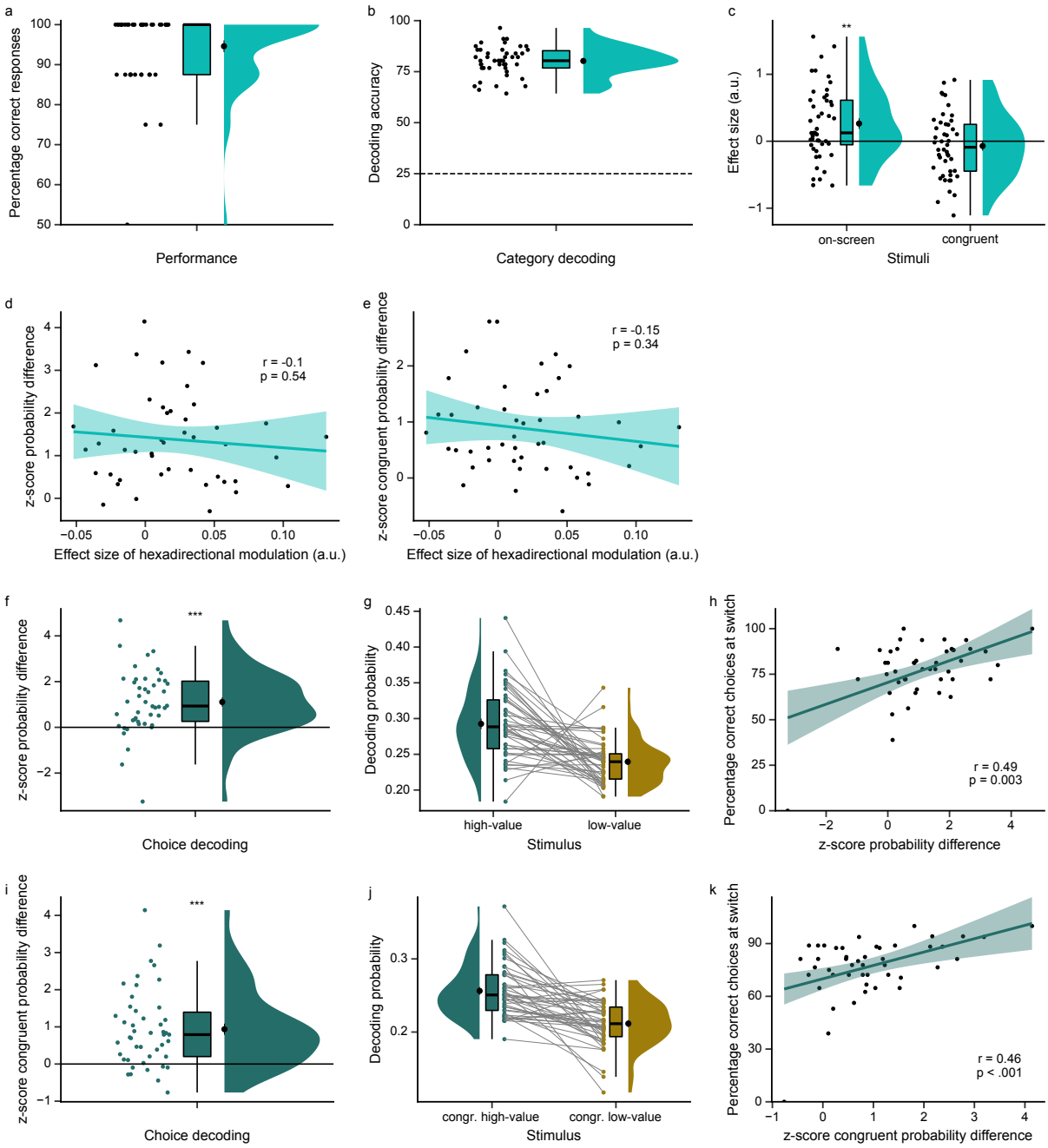


Supplementary Fig. 4 | Entorhinal cortex exhibits grid-like representation for value space. a Entorhinal cortex mask used for small volume correction on the group level. Mask is displayed on the MNI template. **b-f** Control analyses to examine the stability of grid orientations in the significant entorhinal cluster. The magnitude of the hexadirectional modulation should depend on both spatial and temporal stability of grid orientations. **b** Regarding spatial stability, we observed significant clustering of orientations across voxels within-participant. Depicted are the Rayleigh z statistics of the Rayleigh test for non-uniformity of circular data. The test indicated significant deviation from uniformity of voxel orientations in all participants ($p < .05$ in all participants). **c** To illustrate this spatial stability, this plot shows a polar histogram of voxel orientations in the significant entorhinal cluster of an example participant (median Rayleigh z statistic). Red arrow denotes the mean orientation. **d** Across participants, this spatial stability correlated significantly positively with the magnitude of the hexadirectional modulation ($r(44) = .74$, $p < .001$). The effects in **b** and **d** were still present when analyzing spatial stability in unsmoothed data (Rayleigh test for non-uniformity of circular data: $p < .05$ in 41 of 46 participants; correlation: $r(44) = .85$, $p < .001$). **e** Regarding temporal stability, we assessed the percentage of voxels with an orientation difference less than 15° between the estimation set and the left-out test run. Across participants, we observed significant temporal stability (percentage of stable voxels > 50% in 27 of 46 participants, $t(45) = 2.78$, $p = .01$). **f** Across participants, this temporal stability correlated significantly positively with the magnitude of the hexadirectional modulation ($r(44) = .86$, $p < .001$). **g** The magnitude of the hexadirectional modulation does not correlate significantly with performance of the task. **h** Sampling of directions after median split of trajectories according to their mean value. **i** Hexadirectional modulation in the

significant entorhinal cluster, separately for the high- and low-value condition. While the analysis suggested no difference between the two conditions, it suggested a hexadirectional modulation effect only in the low-value condition. **b,e,i** Raincloud plots: dots represent participants' data points; boxplots show median and upper/ lower quartile with whiskers extending to the most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bar corresponds to mean ± SEM; distributions depict probability density function of data points. **d,f,g** Correlation plots: dots represent participants' data points; line represents linear regression line, with shaded regions as the 95% confidence interval. * $p < .05$, ** $p < .01$



Supplementary Fig. 5 | A network of brain regions tracks the prospective value difference during choices. **a** Control analysis: Modulation of activity by the difference between model-derived chosen vs. unchosen value during choices after controlling for reaction time. Clusters depicted survive whole-brain correction (p_{FWE} < .05, TFCE). Statistical image is displayed on the MNI template. **b** Control analysis: Modulation of activity by the prospective component of the value difference during choices after controlling for reaction time. The prospective component refers to the influence of values estimated by the prospective Rescorla-Wagner model over values estimated by the original (non-prospective) Rescorla-Wagner model. Clusters depicted survive whole-brain correction (p_{FWE} < .05, TFCE). Statistical image is displayed on the MNI template.



Supplementary Fig. 6 | Occipital-temporal cortex represents the more valuable choice option. a Performance in the one-back cover task of the picture viewing task (PVT). **b** Decoding accuracy for stimulus category decoding (faces, tools, scenes, body parts) within the PVT, using a leave-trials-out cross-validation procedure. The dashed line at 25 % denotes chance level performance of the decoder. Stimulus category decoding accuracy was well above chance. **c** Effect of the distance between the choice location and the 45°-diagonal of the value space on the decoding probability difference high- vs. low-value option. Depicted are effect sizes estimated by a linear regression, separately for on-screen (left) and for off-screen congruent stimuli (right). For on-screen stimuli, effect sizes are significantly positive, indicating that the probability difference in favor of the high-value option increased with a higher distance to the diagonal. **d** The probability difference in favor of the high-value option for on-screen stimuli does not correlate significantly with the magnitude of the hexadirectional modulation. **e** The probability difference in favor of the high-value option for congruent stimuli does not correlate significantly with the magnitude of the hexadirectional modulation. **f-k** Control analysis for choice decoding using only those choices which sampled the switch time point. Logic of the figures is the same as in Fig. 5 of the main text. **f** Z-scores for the probability difference for the on-screen high- vs. low-value stimuli based on decoding permutation test (see Methods). Occipital-temporal cortex represents the high-value stimulus significantly stronger than the low-value stimulus. **g** Visualization of the effect in **f**, showing the probabilities the decoder assigned to the stimuli (before the permutation test). **h** Correlation between the high- vs. low-value difference score and performance at the switch time point ($r(44) = .49$, $p = .003$; after exclusion of two outliers: $r(42) = .17$, $p = .28$). **i** Z-scores for the probability

difference for the congruent high- vs. low-value stimuli based on decoding permutation test (see Methods). Occipital-temporal cortex represents the congruent high-value stimulus significantly stronger than the congruent low-value stimulus. **j** Visualization of the effect in **g**, showing the probabilities the decoder assigned to the stimuli (before the permutation test). **i** The congruent high-value vs. low-value difference score correlates significantly positively with performance at the switch time point ($r(44) = .46$, $p < .001$; after exclusion of two outliers: $r(42) = .34$, $p = .02$). **a,b,c,f,g,i,j** Raincloud plots: dots represent participants' data points; boxplots show median and upper/lower quartile with whiskers extending to the most extreme data point within 1.5 interquartile ranges above/below the upper/lower quartile; black circle with error bar corresponds to mean ± SEM; distributions depict probability density function of data points. **d,e,h,k** Correlation plots: dots represent participants' data points; line represents linear regression line, with shaded regions as the 95% confidence interval. ** $p < .01$, *** $p < .001$

Supplementary tables

Supplementary Table 1: Exploratory whole-brain clusters of the hexadirectional modulation effect

No clusters survived family-wise error correction in the whole-brain analysis. This table lists clusters at an uncorrected threshold of $p < .001$ with a minimum of 10 voxels. Table lists MNI coordinates (X, Y, Z), statistical T values and atlas labels of peak voxels of the clusters. Atlas labels are based on the Harvard-Oxford Cortical Structural Atlas (HOCSA), Harvard-Oxford Subcortical Structural Atlas (HOSSA) and Juelich Histological Atlas (JHA) provided by FSL. If no label was found for a given atlas, the atlas is not listed. Subclusters are denoted by letters after the cluster ID.

Cluster ID	X	Y	Z	T value	Cluster Size (voxels)	Atlas label
1	-9.0	-11.0	9.0	4.25	12	HOSSA: 100% Left Thalamus
						JHA: 1% WM Corticospinal tract L
2	18.0	-6.0	-26.0	4.17	30	HOCSA: 66% Parahippocampal Gyrus, anterior division
						HOSSA: 66% Right Cerebral Cortex, 24% Right Hippocampus, 8% Right Amygdala, 2% Right Cerebral White Matter
						JHA: 68% GM Hippocampus entorhinal cortex R, 27% GM Hippocampus subiculum R, 16% GM Hippocampus hippocampal-amygdaloid transition area R, 5% GM Hippocampus cornu ammonis R, 2% GM Amygdala_superficial

						group R, 1% GM Amygdala_laterobasal group R
3	6.0	-26.0	-38.0	3.95	12	HOSSA: 100% Brain-Stem
4	-27.0	-3.0	-28.0	3.91	26	HOCSA: 9% Parahippocampal Gyrus, anterior division, 1% Temporal Pole
						HOSSA: 50% Left Amygdala, 27% Left Cerebral White Matter, 12% Left Cerebral Cortex, 11% Left Hippocampus
						JHA: 92% GM Amygdala_laterobasal group L, 5% GM Amygdala_superficial group L, 4% GM Hippocampus subiculum L, 2% WM Optic radiation L, 2% GM Hippocampus cornu ammonis L
4a	-22.0	-1.0	-38.0	3.47		HOCSA: 69% Parahippocampal Gyrus, anterior division, 6% Temporal Fusiform Cortex, anterior division, 3% Temporal Fusiform Cortex, posterior division, 1% Temporal Pole
						HOSSA: 85% Left Cerebral Cortex, 2% Left Cerebral White Matter
						JHA:

						98% GM Hippocampus entorhinal cortex L
5	13.0	-21.0	-31.0	3.59	14	HOSSA: 99% Brain-Stem

Supplementary Table 2: Significant clusters of the value difference effect

Clusters surviving family-wise error correction in whole-brain analysis ($p_{FWE} < .05$). Table lists MNI coordinates (X, Y, Z), statistical T values and atlas labels of peak voxels of the clusters. Atlas labels are based on the Harvard-Oxford Cortical Structural Atlas (HOCSA), Harvard-Oxford Subcortical Structural Atlas (HOSSA) and Juelich Histological Atlas (JHA) provided by FSL. If no label was found for a given atlas, the atlas is not listed. Subclusters are denoted by letters after the cluster ID.

Cluster ID	**X**	**Y**	**Z**	**T value**	**Cluster Size (voxels)**	**Atlas label**
Positive effect						
1	-34.0	2.0	14.0	10.95	32064	HOCSA: 37% Central Opercular Cortex, 15% Insular Cortex, 1% Frontal Operculum Cortex
						HOSSA: 73% Left Cerebral Cortex, 27% Left Cerebral White Matter
1a	-4.0	-3.0	46.0	10.39		HOCSA: 46% Cingulate Gyrus, anterior division, 40% Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)
						HOSSA: 87% Left Cerebral Cortex, 12% Left Cerebral White Matter
						JHA: 42% GM Premotor cortex BA6 L
1b	-49.0	-6.0	12.0	9.78		HOCSA: 49% Central Opercular Cortex, 2% Postcentral Gyrus, 1% Precentral Gyrus
						HOSSA: 51% Left Cerebral Cortex, 49% Left Cerebral White Matter
						JHA: 35% GM Secondary somatosensory cortex / Parietal operculum OP4 L, 9% GM Secondary somatosensory cortex / Parietal operculum OP1 L, 7% GM Broca's area BA44 L, 4% GM

						Primary auditory cortex TE1.0 L, 3% GM Secondary somatosensory cortex / Parietal operculum OP3 L
1c	50.0	-1.0	9.0	9.35		HOCSA: 59% Central Opercular Cortex, 2% Planum Polare, 1% Precentral Gyrus, 1% Inferior Frontal Gyrus, pars opercularis
						HOSSA: 63% Right Cerebral Cortex, 36% Right Cerebral White Matter
						JHA: 28% GM Secondary somatosensory cortex / Parietal operculum OP4 R, 20% GM Secondary somatosensory cortex / Parietal operculum OP3 R, 4% GM Inferior parietal lobule PFop R, 2% GM Primary auditory cortex TE1.2 R
2	23.0	-53.0	-56.0	7.21	83	No label
3	-22.0	-55.0	-58.0	6.21	31	No label
4	3.0	42.0	-8.0	5.99	1394	HOCSA: 63% Paracingulate Gyrus, 17% Frontal Medial Cortex, 16% Cingulate Gyrus, anterior division
						HOSSA: 91% Right Cerebral Cortex, 7% Left Cerebral Cortex, 1% Right Cerebral White Matter
4a	-4.0	32.0	-11.0	5.67		HOCSA: 33% Paracingulate Gyrus, 32% Subcallosal Cortex, 13% Cingulate Gyrus, anterior division, 9% Frontal Medial Cortex
						HOSSA: 90% Left Cerebral Cortex, 10% Left Cerebral White Matter, 0% Right Cerebral Cortex
4b	-7.0	27.0	-8.0	5.51		HOCSA: 63% Subcallosal Cortex, 3% Paracingulate Gyrus, 1% Cingulate Gyrus, anterior division
						HOSSA:

						69% Left Cerebral Cortex, 31% Left Cerebral White Matter
						JHA: 4% WM Cingulum L, 1% WM Callosal body
4c	-29.0	32.0	-14.0	5.40		HOCSA: 46% Frontal Orbital Cortex, 19% Frontal Pole
						HOSSA: 65% Left Cerebral Cortex, 35% Left Cerebral White Matter
5	33.0	34.0	-14.0	4.19	35	HOCSA: 48% Frontal Pole, 38% Frontal Orbital Cortex
						HOSSA: 87% Right Cerebral Cortex, 13% Right Cerebral White Matter
5a	18.0	22.0	-14.0	2.79		HOCSA: 12% Frontal Orbital Cortex
						HOSSA: 85% Right Cerebral White Matter, 15% Right Cerebral Cortex
6	-37.0	-11.0	-48.0	4.19	3	HOCSA: 11% Inferior Temporal Gyrus, anterior division, 8% Temporal Fusiform Cortex, posterior division, 8% Inferior Temporal Gyrus, posterior division, 6% Temporal Fusiform Cortex, anterior division
						HOSSA: 38% Left Cerebral Cortex, 0% Left Cerebral White Matter
7	1.0	14.0	14.0	4.14	9	HOSSA: 51% Right Lateral Ventricle, 36% Right Cerebral White Matter, 9% Left Cerebral White Matter, 4% Left Lateral Ventricle, 0% Left Cerebral Cortex
						JHA: 40% WM Callosal body
8	15.0	32.0	6.0	3.23	8	HOSSA:

						97% Right Cerebral White Matter, 3% Right Lateral Ventricle
						JHA: 97% WM Callosal body, 5% WM Cingulum R
Negative effect						
1	6.0	24.0	52.0	-10.28	11448	HOCSA: 52% Superior Frontal Gyrus, 10% Paracingulate Gyrus
						HOSSA: 79% Right Cerebral Cortex, 21% Right Cerebral White Matter
						JHA: 20% GM Premotor cortex BA6 R
1a	30.0	7.0	59.0	-9.56		HOCSA: 34% Middle Frontal Gyrus, 22% Superior Frontal Gyrus, 1% Precentral Gyrus
						HOSSA: 80% Right Cerebral Cortex, 10% Right Cerebral White Matter
1b	-2.0	29.0	39.0	-9.52		HOCSA: 61% Paracingulate Gyrus, 15% Superior Frontal Gyrus
						HOSSA: 86% Left Cerebral Cortex, 1% Right Cerebral Cortex, 0% Left Cerebral White Matter
1c	3.0	37.0	46.0	-9.49		HOCSA: 55% Superior Frontal Gyrus, 1% Paracingulate Gyrus
						HOSSA: 69% Right Cerebral Cortex, 9% Left Cerebral Cortex
						JHA: 1% GM Premotor cortex BA6 R
2	-37.0	-65.0	-28.0	-9.96	2796	No label
2a	-32.0	-65.0	-28.0	-9.93		No label
2b	30.0	-63.0	-28.0	-8.86		No label
2c	8.0	-78.0	-24.0	-8.52		HOCSA:

						3% Occipital Fusiform Gyrus, 2% Lingual Gyrus
						HOSSA: 5% Right Cerebral Cortex
3	43.0	-38.0	42.0	-9.09	3463	HOCSA: 41% Supramarginal Gyrus, posterior division, 10% Superior Parietal Lobule, 5% Angular Gyrus, 5% Postcentral Gyrus, 4% Supramarginal Gyrus, anterior division
						HOSSA: 68% Right Cerebral Cortex, 32% Right Cerebral White Matter
						JHA: 28% GM Anterior intra-parietal sulcus hIP3 R, 17% GM Anterior intra-parietal sulcus hIP2 R, 11% GM Superior parietal lobule 7PC R, 3% GM Anterior intra-parietal sulcus hIP1 R
3a	-47.0	-50.0	52.0	-8.25		HOCSA: 24% Supramarginal Gyrus, posterior division, 22% Angular Gyrus, 13% Superior Parietal Lobule, 2% Supramarginal Gyrus, anterior division, 1% Lateral Occipital Cortex, superior division
						HOSSA: 65% Left Cerebral Cortex, 25% Left Cerebral White Matter
						JHA: 40% GM Inferior parietal lobule PFm L, 32% GM Inferior parietal lobule PF L, 15% GM Anterior intra-parietal sulcus hIP2 L, 12% GM Inferior parietal lobule Pga L, 11% GM Anterior intra-parietal sulcus hIP1 L, 10% GM Superior parietal lobule 7PC L, 9% GM Anterior intra-parietal sulcus hIP3 L, 3% GM Superior parietal lobule 5L L
3b	48.0	-45.0	52.0	-8.04		HOCSA: 38% Supramarginal Gyrus, posterior division, 33% Angular Gyrus, 9% Superior Parietal Lobule

						HOSSA: 82% Right Cerebral Cortex, 17% Right Cerebral White Matter
						JHA: 77% GM Inferior parietal lobule PFm R, 13% GM Anterior intra-parietal sulcus hIP1 R, 12% GM Inferior parietal lobule Pga R, 12% GM Anterior intra-parietal sulcus hIP2 R, 2% GM Anterior intra-parietal sulcus hIP3 R, 1% GM Superior parietal lobule 7PC R
3c	38.0	-48.0	42.0	-7.39		HOCSA: 23% Superior Parietal Lobule, 18% Supramarginal Gyrus, posterior division, 15% Angular Gyrus
						HOSSA: 60% Right Cerebral Cortex, 39% Right Cerebral White Matter
						JHA: 45% GM Anterior intra-parietal sulcus hIP1 R, 29% GM Anterior intra-parietal sulcus hIP3 R, 20% GM Anterior intra-parietal sulcus hIP2 R
4	-12.0	2.0	-1.0	-8.09	1717	HOSSA: 60% Left Pallidum, 40% Left Cerebral White Matter
4a	8.0	7.0	9.0	-7.82		HOSSA: 57% Right Caudate, 41% Right Lateral Ventricle, 2% Right Cerebral White Matter
4b	10.0	2.0	14.0	-7.39		HOSSA: 68% Right Caudate, 30% Right Lateral Ventricle, 1% Right Cerebral White Matter, 1% Right Thalamus
4c	10.0	4.0	-1.0	-7.29		HOSSA: 90% Right Cerebral White Matter, 5% Right Caudate, 4% Right Pallidum, 0% Right Cerebral Cortex, 0% Right Thalamus, 0% Right Lateral Ventricle, 0% Right Accumbens
5	-19.0	42.0	-21.0	-6.0	19	HOCSA: 72% Frontal Pole, 3% Frontal Orbital Cortex

						HOSSA: 75% Left Cerebral Cortex, 20% Left Cerebral White Matter
6	-9.0	-58.0	-51.0	-5.73	5	No label

Supplementary Table 3: Significant clusters of the prospective value difference effect

Clusters surviving family-wise error correction in whole-brain analysis ($p_{FWE} < .05$). Table lists MNI coordinates (X, Y, Z), statistical T values and atlas labels of peak voxels of the clusters. Atlas labels are based on the Harvard-Oxford Cortical Structural Atlas (HOCSA), Harvard-Oxford Subcortical Structural Atlas (HOSSA) and Juelich Histological Atlas (JHA) provided by FSL. If no label was found for a given atlas, the atlas is not listed. Subclusters are denoted by letters after the cluster ID. In total, 96 clusters including subclusters were found. Table lists 10 clusters with the most extreme T values, separately for positive and negative effects.

Cluster ID	**X**	**Y**	**Z**	**T value**	**Cluster Size (voxels)**	**Atlas label**
Positive effect						
1	-34.0	-26.0	56.0	6.77	1029	HOCSA: 34% Postcentral Gyrus, 25% Precentral Gyrus
						HOSSA: 67% Left Cerebral Cortex, 32% Left Cerebral White Matter
						JHA: 44% WM Corticospinal tract L, 42% GM Primary motor cortex BA4p L, 28% GM Primary somatosensory cortex BA3b L, 28% GM Primary motor cortex BA4a L, 10% GM Primary somatosensory cortex BA1 L, 6% GM Primary somatosensory cortex BA2 L
1a	-32.0	-23.0	62.0	6.50		HOCSA: 34% Precentral Gyrus, 14% Postcentral Gyrus
						HOSSA: 55% Left Cerebral Cortex, 44% Left Cerebral White Matter
						JHA: 51% GM Premotor cortex BA6 L, 47% WM Corticospinal tract L, 32% GM Primary motor cortex BA4a L, 5% GM Primary somatosensory cortex BA3b L, 5% GM Primary somatosensory cortex BA1 L, 4% GM Primary motor cortex BA4p L

1b	-34.0	-33.0	66.0	6.39		HOCSA: 51% Postcentral Gyrus, 6% Superior Parietal Lobule, 3% Precentral Gyrus, 2% Supramarginal Gyrus, anterior division
						HOSSA: 66% Left Cerebral Cortex, 22% Left Cerebral White Matter
						JHA: 71% GM Primary somatosensory cortex BA1 L, 30% GM Primary motor cortex BA4a L, 29% GM Primary somatosensory cortex BA2 L, 27% GM Primary somatosensory cortex BA3b L, 12% GM Primary motor cortex BA4p L, 8% GM Superior parietal lobule 5L L, 5% GM Superior parietal lobule 7PC L, 1% WM Corticospinal tract L
1c	-54.0	-26.0	49.0	6.09		HOCSA: 61% Postcentral Gyrus, 17% Supramarginal Gyrus, anterior division
						HOSSA: 84% Left Cerebral Cortex, 11% Left Cerebral White Matter
						JHA: 58% GM Primary somatosensory cortex BA2 L, 56% GM Primary somatosensory cortex BA1 L, 31% GM Inferior parietal lobule PF L, 20% GM Inferior parietal lobule PFt L, 9% GM Anterior intra-parietal sulcus hIP2 L, 3% GM Primary somatosensory cortex BA3b L, 2% GM Inferior parietal lobule PFop L
2	45.0	-28.0	26.0	6.67	124	HOCSA: 31% Parietal Operculum Cortex, 10% Supramarginal Gyrus, anterior division, 1% Supramarginal Gyrus, posterior division
						HOSSA: 56% Right Cerebral White Matter, 44% Right Cerebral Cortex

						JHA: 40% GM Secondary somatosensory cortex / Parietal operculum OP1 R, 34% GM Inferior parietal lobule PFop R, 24% GM Inferior parietal lobule PFcm R, 4% GM Anterior intra-parietal sulcus hIP2 R
3	-49.0	-23.0	22.0	6.38	107	HOCSA: 31% Central Opercular Cortex, 15% Parietal Operculum Cortex, 4% Postcentral Gyrus, 1% Supramarginal Gyrus, anterior division
						HOSSA: 53% Left Cerebral Cortex, 47% Left Cerebral White Matter
						JHA: 72% GM Secondary somatosensory cortex / Parietal operculum OP1 L, 19% GM Inferior parietal lobule PFop L, 9% GM Secondary somatosensory cortex / Parietal operculum OP4 L, 7% GM Inferior parietal lobule PFcm L, 3% GM Primary auditory cortex TE1.0 L
3a	-44.0	-33.0	22.0	4.07		HOCSA: 60% Parietal Operculum Cortex, 3% Supramarginal Gyrus, posterior division, 1% Central Opercular Cortex, 1% Superior Temporal Gyrus, posterior division
						HOSSA: 67% Left Cerebral Cortex, 31% Left Cerebral White Matter
						JHA: 63% GM Inferior parietal lobule PFcm L, 30% GM Secondary somatosensory cortex / Parietal operculum OP1 L, 3% GM Inferior parietal lobule PFop L, 1% GM Inferior parietal lobule PF L
4	50.0	-63.0	-6.0	6.06	77	HOCSA: 44% Lateral Occipital Cortex, inferior division, 11% Inferior Temporal Gyrus, temporooccipital

						part, 8% Middle Temporal Gyrus, temporooccipital part, 2% Occipital Fusiform Gyrus
						HOSSA: 67% Right Cerebral Cortex, 33% Right Cerebral White Matter
						JHA: 8% GM Visual cortex V5 R
5	-39.0	-1.0	16.0	5.78	1093	HOCSA: 57% Central Opercular Cortex, 10% Insular Cortex
						HOSSA: 78% Left Cerebral Cortex, 22% Left Cerebral White Matter
5a	-29.0	7.0	-6.0	5.56		HOSSA: 73% Left Cerebral White Matter, 27% Left Putamen, 0% Left Cerebral Cortex
						JHA: 58% WM Inferior occipito-frontal fascicle L, 12% WM Uncinate fascicle L
5b	-32.0	-6.0	2.0	5.52		HOSSA: 61% Left Putamen, 39% Left Cerebral White Matter, 0% Left Cerebral Cortex
5c	-22.0	-13.0	-18.0	5.09		HOSSA: 91% Left Hippocampus, 2% Left Amygdala
						JHA: 80% GM Hippocampus cornu ammonis L, 42% GM Hippocampus subiculum L, 34% GM Hippocampus dentate gyrus L, 14% GM Amygdala_superficial group L, 13% GM Hippocampus hippocampal-amygdaloid transition area L, 13% GM Hippocampus entorhinal cortex L, 11% GM Amygdala_laterobasal group L, 4% GM Amygdala_centromedial group L
6	-54.0	-65.0	39.0	5.77	273	HOCSA:

						38% Lateral Occipital Cortex, superior division, 1% Angular Gyrus
						HOSSA: 48% Left Cerebral Cortex, 0% Left Cerebral White Matter
6a	-54.0	-65.0	34.0	5.56		HOCSA: 74% Lateral Occipital Cortex, superior division, 5% Angular Gyrus
						HOSSA: 85% Left Cerebral Cortex, 1% Left Cerebral White Matter
						JHA: 32% GM Inferior parietal lobule PGp L, 16% GM Inferior parietal lobule Pga L, 12% GM Inferior parietal lobule PFm L
6b	-52.0	-55.0	29.0	4.31		HOCSA: 49% Angular Gyrus, 14% Supramarginal Gyrus, posterior division, 5% Lateral Occipital Cortex, superior division
						HOSSA: 70% Left Cerebral Cortex, 29% Left Cerebral White Matter
						JHA: 47% GM Inferior parietal lobule Pga L, 31% GM Inferior parietal lobule PFm L, 20% GM Inferior parietal lobule PF L
7	-7.0	52.0	-8.0	5.75	1423	HOCSA: 53% Frontal Medial Cortex, 24% Paracingulate Gyrus, 13% Frontal Pole
						HOSSA: 91% Left Cerebral Cortex, 9% Left Cerebral White Matter
7a	-19.0	59.0	26.0	5.20		HOCSA: 75% Frontal Pole
						HOSSA: 79% Left Cerebral Cortex, 15% Left Cerebral White Matter

7b	-14.0	54.0	34.0	4.98		HOCSA: 79% Frontal Pole
						HOSSA: 82% Left Cerebral Cortex, 9% Left Cerebral White Matter
7c	-4.0	59.0	19.0	4.82		HOCSA: 53% Frontal Pole, 19% Superior Frontal Gyrus, 8% Paracingulate Gyrus
						HOSSA: 91% Left Cerebral Cortex, 4% Left Cerebral White Matter
8	28.0	-8.0	-18.0	5.16	539	HOSSA: 48% Right Amygdala, 34% Right Hippocampus, 2% Right Cerebral White Matter
						JHA: 79% GM Amygdala_laterobasal group R, 74% GM Hippocampus cornu ammonis R, 15% GM Hippocampus dentate gyrus R, 5% GM Amygdala_superficial group R, 3% GM Hippocampus subiculum R, 1% GM Amygdala_centromedial group R
8a	18.0	-11.0	-14.0	4.79		HOSSA: 78% Right Amygdala, 12% Right Hippocampus, 5% Right Cerebral White Matter, 0% Right Cerebral Cortex
						JHA: 35% GM Amygdala_superficial group R, 23% GM Hippocampus subiculum R, 22% GM Hippocampus hippocampal-amygdaloid transition area R, 14% WM Corticospinal tract R, 14% GM Amygdala_laterobasal group R, 12% GM Hippocampus cornu ammonis R, 8% GM Amygdala_centromedial group R, 6% GM Hippocampus entorhinal cortex R, 2% WM Acoustic radiation R
8b	25.0	12.0	-1.0	4.73		HOSSA:

						96% Right Putamen, 4% Right Cerebral White Matter
8c	40.0	-30.0	-21.0	4.22		HOCSA: 53% Temporal Fusiform Cortex, posterior division, 11% Inferior Temporal Gyrus, posterior division, 3% Inferior Temporal Gyrus, temporooccipital part, 2% Temporal Occipital Fusiform Cortex
						HOSSA: 69% Right Cerebral Cortex, 30% Right Cerebral White Matter
						JHA: 2% GM Hippocampus cornu ammonis R, 1% WM Optic radiation R
9	-2.0	-48.0	32.0	4.90	57	HOCSA: 84% Cingulate Gyrus, posterior division, 10% Precuneous Cortex
						HOSSA: 96% Left Cerebral Cortex, 1% Left Cerebral White Matter, 1% Right Cerebral Cortex
9a	-9.0	-53.0	29.0	4.59		HOCSA: 32% Cingulate Gyrus, posterior division, 11% Precuneous Cortex
						HOSSA: 57% Left Cerebral White Matter, 43% Left Cerebral Cortex
						JHA: 3% WM Callosal body
10	-7.0	-23.0	49.0	4.85	327	HOCSA: 54% Precentral Gyrus, 15% Cingulate Gyrus, posterior division, 3% Cingulate Gyrus, anterior division, 3% Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)
						HOSSA: 81% Left Cerebral Cortex, 19% Left Cerebral White Matter
						JHA: 40% GM Premotor cortex BA6 L, 34% GM Primary motor cortex

						BA4a L, 4% WM Corticospinal tract L, 3% GM Superior parietal lobule 5M L
10a	-4.0	-6.0	56.0	4.77		HOCSA: 78% Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex)
						HOSSA: 81% Left Cerebral Cortex, 17% Left Cerebral White Matter
						JHA: 78% GM Premotor cortex BA6 L
10b	-7.0	-11.0	59.0	4.70		HOCSA: 43% Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex), 6% Precentral Gyrus
						HOSSA: 52% Left Cerebral Cortex, 48% Left Cerebral White Matter
						JHA: 96% GM Premotor cortex BA6 L, 3% WM Corticospinal tract L
10c	8.0	-3.0	56.0	4.15		HOCSA: 44% Juxtapositional Lobule Cortex (formerly Supplementary Motor Cortex), 2% Precentral Gyrus, 1% Cingulate Gyrus, anterior division
						HOSSA: 50% Right Cerebral White Matter, 50% Right Cerebral Cortex
						JHA: 61% GM Premotor cortex BA6 R
Negative effect						
1	30.0	29.0	4.0	-6.77	23	HOCSA: 11% Frontal Orbital Cortex, 7% Insular Cortex, 6% Inferior Frontal Gyrus, pars triangularis, 3% Frontal Operculum Cortex
						HOSSA: 64% Right Cerebral White Matter, 36% Right Cerebral Cortex

2	6.0	-65.0	52.0	-6.51	1337	HOCSA: 66% Precuneous Cortex, 2% Lateral Occipital Cortex, superior division
						HOSSA: 70% Right Cerebral Cortex, 30% Right Cerebral White Matter
						JHA: 38% GM Superior parietal lobule 7P R, 32% GM Superior parietal lobule 7A R, 5% GM Superior parietal lobule 7M R
2a	-2.0	-63.0	54.0	-5.35		HOCSA: 83% Precuneous Cortex
						HOSSA: 85% Left Cerebral Cortex, 6% Left Cerebral White Matter, 1% Right Cerebral Cortex
						JHA: 42% GM Superior parietal lobule 7A L, 33% GM Superior parietal lobule 7P L
2b	-24.0	-53.0	42.0	-5.27		HOCSA: 24% Superior Parietal Lobule, 5% Lateral Occipital Cortex, superior division, 4% Supramarginal Gyrus, posterior division, 2% Angular Gyrus
						HOSSA: 64% Left Cerebral White Matter, 36% Left Cerebral Cortex
						JHA: 17% GM Anterior intra-parietal sulcus hIP3 L, 10% GM Anterior intra-parietal sulcus hIP1 L, 1% GM Superior parietal lobule 7A L
2c	10.0	-43.0	46.0	-5.25		HOCSA: 56% Precuneous Cortex, 16% Cingulate Gyrus, posterior division, 2% Postcentral Gyrus
						HOSSA: 76% Right Cerebral Cortex, 24% Right Cerebral White Matter

						JHA: 20% GM Superior parietal lobule 5Ci R, 5% GM Superior parietal lobule 5M R
3	1.0	17.0	49.0	-6.41	351	HOCSA: 54% Paracingulate Gyrus, 7% Superior Frontal Gyrus, 2% Cingulate Gyrus, anterior division
						HOSSA: 69% Right Cerebral Cortex, 7% Left Cerebral Cortex, 0% Right Cerebral White Matter
						JHA: 29% GM Premotor cortex BA6 R
3a	3.0	29.0	46.0	-4.15		HOCSA: 31% Superior Frontal Gyrus, 27% Paracingulate Gyrus
						HOSSA: 67% Right Cerebral Cortex, 9% Left Cerebral Cortex
						JHA: 3% GM Premotor cortex BA6 R
3b	8.0	32.0	44.0	-3.79		HOCSA: 31% Superior Frontal Gyrus, 11% Paracingulate Gyrus
						HOSSA: 57% Right Cerebral Cortex, 43% Right Cerebral White Matter
						JHA: 8% GM Premotor cortex BA6 R
4	1.0	-88.0	-11.0	-6.33	732	HOCSA: 45% Lingual Gyrus, 6% Occipital Pole, 3% Intracalcarine Cortex, 1% Occipital Fusiform Gyrus
						HOSSA: 32% Right Cerebral Cortex, 26% Left Cerebral Cortex, 4% Right Cerebral White Matter, 0% Left Cerebral White Matter
						JHA: 4% GM Visual cortex V1 BA17 R

4a	8.0	-80.0	-8.0	-5.38		HOCSA: 71% Lingual Gyrus, 6% Occipital Fusiform Gyrus, 1% Intracalcarine Cortex
						HOSSA: 78% Right Cerebral Cortex, 18% Right Cerebral White Matter, 0% Left Cerebral Cortex
						JHA: 56% GM Visual cortex V1 BA17 R, 52% GM Visual cortex V2 BA18 R, 12% WM Optic radiation R, 6% GM Visual cortex V3V R
4b	-12.0	-73.0	16.0	-4.24		HOCSA: 24% Intracalcarine Cortex, 9% Cuneal Cortex, 3% Supracalcarine Cortex, 2% Precuneous Cortex
						HOSSA: 59% Left Cerebral White Matter, 41% Left Cerebral Cortex
						JHA: 40% GM Visual cortex V1 BA17 L, 18% WM Optic radiation L, 10% GM Visual cortex V2 BA18 L
4c	13.0	-65.0	9.0	-4.18		HOCSA: 61% Intracalcarine Cortex, 4% Lingual Gyrus, 2% Supracalcarine Cortex
						HOSSA: 67% Right Cerebral Cortex, 33% Right Cerebral White Matter
						JHA: 74% GM Visual cortex V1 BA17 R, 33% GM Visual cortex V2 BA18 R, 32% WM Optic radiation R, 10% GM Visual cortex V3V R, 1% GM Visual cortex V4 R
5	30.0	-60.0	52.0	-5.93	517	HOCSA: 44% Lateral Occipital Cortex, superior division, 11% Superior Parietal Lobule, 7% Angular Gyrus
						HOSSA: 66% Right Cerebral Cortex, 29% Right Cerebral White Matter

						JHA: 37% GM Anterior intra-parietal sulcus hIP3 R, 28% GM Superior parietal lobule 7A R, 10% GM Anterior intra-parietal sulcus hIP1 R, 2% GM Superior parietal lobule 7PC R, 1% GM Superior parietal lobule 7P R
5a	33.0	-53.0	44.0	-5.0		HOCSA: 25% Superior Parietal Lobule, 21% Angular Gyrus, 8% Lateral Occipital Cortex, superior division, 2% Supramarginal Gyrus, posterior division
						HOSSA: 59% Right Cerebral Cortex, 41% Right Cerebral White Matter
						JHA: 25% GM Anterior intra-parietal sulcus hIP3 R, 11% GM Anterior intra-parietal sulcus hIP1 R, 6% GM Anterior intra-parietal sulcus hIP2 R, 2% GM Superior parietal lobule 7A R
5b	45.0	-43.0	49.0	-4.95		HOCSA: 41% Supramarginal Gyrus, posterior division, 18% Angular Gyrus, 8% Superior Parietal Lobule, 2% Postcentral Gyrus, 1% Supramarginal Gyrus, anterior division
						HOSSA: 71% Right Cerebral Cortex, 27% Right Cerebral White Matter
						JHA: 42% GM Inferior parietal lobule PFm R, 15% GM Anterior intra-parietal sulcus hIP2 R, 15% GM Anterior intra-parietal sulcus hIP1 R, 14% GM Anterior intra-parietal sulcus hIP3 R, 5% GM Inferior parietal lobule Pga R, 4% GM Superior parietal lobule 7PC R, 2% GM Inferior parietal lobule PF R
5c	50.0	-35.0	52.0	-4.88		HOCSA: 30% Supramarginal Gyrus, anterior division, 28% Supramarginal

						Gyrus, posterior division, 13% Postcentral Gyrus, 2% Superior Parietal Lobule, 1% Angular Gyrus
						HOSSA: 81% Right Cerebral Cortex, 15% Right Cerebral White Matter
						JHA: 39% GM Inferior parietal lobule PFt R, 28% GM Inferior parietal lobule PFm R, 25% GM Primary somatosensory cortex BA2 R, 19% GM Anterior intra-parietal sulcus hIP2 R, 10% GM Superior parietal lobule 7PC R, 2% GM Primary somatosensory cortex BA1 R
6	30.0	12.0	56.0	-5.38	72	HOCSA: 35% Middle Frontal Gyrus, 11% Superior Frontal Gyrus
						HOSSA: 66% Right Cerebral Cortex, 30% Right Cerebral White Matter
6a	30.0	9.0	64.0	-4.81		HOCSA: 22% Middle Frontal Gyrus, 18% Superior Frontal Gyrus
						HOSSA: 62% Right Cerebral Cortex, 2% Right Cerebral White Matter
7	50.0	37.0	26.0	-5.31	6	HOCSA: 41% Frontal Pole, 16% Middle Frontal Gyrus
						HOSSA: 67% Right Cerebral Cortex, 1% Right Cerebral White Matter
						JHA: 12% GM Broca's area BA45 R
8	53.0	34.0	26.0	-4.88	2	HOCSA: 15% Middle Frontal Gyrus, 10% Frontal Pole
						HOSSA: 38% Right Cerebral Cortex, 1% Right Cerebral White Matter
						JHA:

						4% GM Broca's area BA45 R
9	-42.0	9.0	24.0	-4.77	344	HOCSA: 31% Inferior Frontal Gyrus, pars opercularis, 12% Precentral Gyrus, 5% Middle Frontal Gyrus
						HOSSA: 56% Left Cerebral Cortex, 44% Left Cerebral White Matter
						JHA: 32% GM Broca's area BA44 L
9a	-44.0	-3.0	42.0	-4.2		HOCSA: 35% Precentral Gyrus, 7% Middle Frontal Gyrus
						HOSSA: 51% Left Cerebral Cortex, 49% Left Cerebral White Matter
						JHA: 19% GM Premotor cortex BA6 L, 10% WM Corticospinal tract L, 3% GM Primary motor cortex BA4a L
9b	-34.0	-1.0	49.0	-4.16		HOCSA: 28% Precentral Gyrus, 17% Middle Frontal Gyrus, 1% Superior Frontal Gyrus
						HOSSA: 57% Left Cerebral Cortex, 43% Left Cerebral White Matter
						JHA: 18% GM Premotor cortex BA6 L, 8% GM Primary motor cortex BA4a L, 4% WM Corticospinal tract L
9c	-47.0	-1.0	52.0	-4.11		HOCSA: 52% Precentral Gyrus, 18% Middle Frontal Gyrus
						HOSSA: 81% Left Cerebral Cortex, 15% Left Cerebral White Matter
						JHA: 81% GM Premotor cortex BA6 L, 1% GM Primary motor cortex BA4a L

10	-22.0	-26.0	24.0	-4.31	47	HOSSA: 95% Left Cerebral White Matter, 4% Left Lateral Ventricle, 1% Left Caudate
						JHA: 28% WM Corticospinal tract L, 19% WM Superior occipito-frontal fascicle L
10a	-9.0	-23.0	26.0	-3.88		HOSSA: 69% Left Cerebral White Matter, 31% Left Lateral Ventricle, 0% Left Cerebral Cortex
						JHA: 49% WM Callosal body, 2% WM Cingulum L

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*. https://doi.org/10.3389/fninf.2014.00014

Ahn, W.-Y., Haines, N., & Zhang, L. (2017). Revealing Neurocomputational Mechanisms of Reinforcement Learning and Decision-Making With the hBayesDM Package. *Computational Psychiatry*, *1*(0), Article 0. https://doi.org/10.1162/CPSY_a_00002

Avants, B., Epstein, C., Grossman, M., & Gee, J. (2008). Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis*, *12*(1), 26–41. https://doi.org/10.1016/j.media.2007.06.004

Bao, X., Gjorgieva, E., Shanahan, L. K., Howard, J. D., Kahnt, T., & Gottfried, J. A. (2019). Grid-like Neural Representations Support Olfactory Navigation of a Two-Dimensional Odor Space. *Neuron*, *102*(5), 1066-1075.e5. https://doi.org/10.1016/j.neuron.2019.03.034

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*, *76*, 412–427. https://doi.org/10.1016/j.neuroimage.2013.02.063

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., & Kurth-Nelson, Z. (2018). What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. *Neuron*, *100*(2), 490–509. https://doi.org/10.1016/j.neuron.2018.10.002

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, *37*(1), 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042

Bellmund, J. L. S., Deuker, L., Navarro Schröder, T., & Doeller, C. F. (2016). Grid-cell representations in mental simulation. *ELife*, *5*, e17089. https://doi.org/10.7554/eLife.17089

Bellmund, J. L. S., Gärdenfors, P., Moser, E. I., & Doeller, C. F. (2018). Navigating cognition: Spatial codes for human thinking. *Science*, *362*(6415), eaat6766. https://doi.org/10.1126/science.aat6766

Boccara, C. N., Nardin, M., Stella, F., O'Neill, J., & Csicsvari, J. (2019). The entorhinal cognitive map is attracted to goals. *Science*, *363*(6434), 1443–1447. https://doi.org/10.1126/science.aav4837

Bongioanni, A., Folloni, D., Verhagen, L., Sallet, J., Klein-Flügge, M. C., & Rushworth, M. F. S. (2021). Activation and disruption of a neural mechanism for novel choice in monkeys. *Nature*, *591*(7849), Article 7849. https://doi.org/10.1038/s41586-020-03115-5

Boorman, E. D., Behrens, T. E. J., Woolrich, M. W., & Rushworth, M. F. S. (2009). How Green Is the Grass on the Other Side? Frontopolar Cortex and the Evidence in Favor of Alternative Courses of Action. *Neuron*, *62*(5), 733–743. https://doi.org/10.1016/j.neuron.2009.05.014

Bornstein, A. M., & Daw, N. D. (2013). Cortical and Hippocampal Correlates of Deliberation during Model-Based Decisions for Rewards in Humans. *PLoS Computational Biology*, *9*(12), e1003387. https://doi.org/10.1371/journal.pcbi.1003387

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, *105*(38), 14325–14329. https://doi.org/10.1073/pnas.0803390105

Brooks, J. L. (2012). Counterbalancing for serial order carryover effects in experimental condition orders. *Psychological Methods*, *17*, 600–614. https://doi.org/10.1037/a0029310

Brown, T. I., Carr, V. A., LaRocque, K. F., Favila, S. E., Gordon, A. M., Bowles, B., Bailenson, J. N., & Wagner, A. D. (2016). Prospective representation of navigational goals in the human hippocampus. *Science*, *352*(6291), 1323–1326. https://doi.org/10.1126/science.aaf0784

Bush, D., Barry, C., Manson, D., & Burgess, N. (2015). Using Grid Cells for Navigation. *Neuron*, *87*(3), 507–520. https://doi.org/10.1016/j.neuron.2015.07.006

Butler, W. N., Hardcastle, K., & Giocomo, L. M. (2019). Remembered reward locations restructure entorhinal spatial maps. *Science*, *363*(6434), 1447–1452. https://doi.org/10.1126/science.aav5297

Cichy, R. M., Pantazis, D., & Oliva, A. (2016). Similarity-Based Fusion of MEG and fMRI Reveals Spatio-Temporal Dynamics in Human Cortex During Visual Object Recognition. *Cerebral Cortex*, *26*(8), 3563–3579. https://doi.org/10.1093/cercor/bhw135

Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. J. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, *352*(6292), 1464–1468. https://doi.org/10.1126/science.aaf0941

Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine*, *10*(4–5), 171–178. https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), Article 12. https://doi.org/10.1038/nn1560

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience*, *16*(1), Article 1. https://doi.org/10.1038/nn.3279

Doeller, C. F., Barry, C., & Burgess, N. (2010). Evidence for grid cells in a human memory network. *Nature*, *463*(7281), 657–661. https://doi.org/10.1038/nature08704

Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D., & Daw, N. D. (2015). Model-based choices involve prospective neural activity. *Nature Neuroscience*, *18*(5), 767–772. https://doi.org/10.1038/nn.3981

Epstein, R. A., Patai, E. Z., Julian, J. B., & Spiers, H. J. (2017). The cognitive map in humans: Spatial navigation and beyond. *Nature Neuroscience*, *20*(11), 1504–1513. https://doi.org/10.1038/nn.4656

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2018). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Esteban, Oscar, Markiewicz, Christopher J., Burns, Christopher, Goncalves, Mathias, Jarecka, Dorota, Ziegler, Erik, Berleant, Shoshana, Ellis, David Gage, Pinsard, Basile, Madison, Cindee, Waskom, Michael, Notter, Michael Philipp, Clark, Daniel, Manhães-Savio, Alexandre, Clark, Dav, Jordan, Kesshi, Dayan, Michael, Halchenko, Yaroslav O., Loney, Fred, … Ghosh, Satrajit. (2022). *nipy/nipype: 1.8.3* (1.8.3) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.596855

Esteban, Oscar, Markiewicz, Christopher J., Goncalves, Mathias, Provins, Céline, Kent, James D., DuPre, Elizabeth, Salo, Taylor, Ciric, Rastko, Pinsard, Basile, Blair, Ross W., Poldrack, Russell A., & Gorgolewski, Krzysztof J. (2022). *fMRIPrep: A robust preprocessing pipeline for functional MRI* (22.1.0) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.852659

Evans, A. C., Janke, A. L., Collins, D. L., & Baillet, S. (2012). Brain templates and atlases. *NeuroImage*, *62*(2), 911–922. https://doi.org/10.1016/j.neuroimage.2012.01.024

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Glasser, M. F., Miller, K. L., Ugurbil, K., & Yacoub, E. (2010). Multiplexed Echo Planar Imaging for Sub-Second Whole Brain FMRI and Fast Diffusion Imaging. *PLOS ONE*, *5*(12), e15710. https://doi.org/10.1371/journal.pone.0015710

FitzGerald, T. H. B., Seymour, B., & Dolan, R. J. (2009). The Role of Human Orbitofrontal Cortex in Value Comparison for Incommensurable Objects. *Journal of Neuroscience*, *29*(26), 8388–8395. https://doi.org/10.1523/JNEUROSCI.0717-09.2009

Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, *47*, S102. https://doi.org/10.1016/S1053-8119(09)70884-5

Garvert, M. M., Dolan, R. J., & Behrens, T. E. (2017). A map of abstract relational knowledge in the human hippocampal–entorhinal cortex. *ELife*, *6*, e17086. https://doi.org/10.7554/eLife.17086

Garvert, M. M., Saanum, T., Schulz, E., Schuck, N. W., & Doeller, C. F. (2023). Hippocampal spatio-predictive cognitive maps adaptively guide reward generalization. *Nature Neuroscience*, *26*(4), Article 4. https://doi.org/10.1038/s41593-023-01283-x

Gershman, S. J., & Daw, N. D. (2017). Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology*, *68*(1), 101–128. https://doi.org/10.1146/annurev-psych-122414-033625

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Frontiers in Neuroinformatics*, *5*. https://doi.org/10.3389/fninf.2011.00013

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., … Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing

outputs of neuroimaging experiments. *Scientific Data*, *3*(1), Article 1. https://doi.org/10.1038/sdata.2016.44

Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, *48*(1), 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, *436*(7052), 801–806. https://doi.org/10.1038/nature03721

Hall-McMaster, S., Dayan, P., & Schuck, N. W. (2021). Control over patch encounters changes foraging behavior. *IScience*, *24*(9), 103005. https://doi.org/10.1016/j.isci.2021.103005

Hayden, B. Y., Pearson, J. M., & Platt, M. L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience*, *14*(7), Article 7. https://doi.org/10.1038/nn.2856

Hegarty, M. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence*, *30*(5), 425–447. https://doi.org/10.1016/S0160-2896(02)00116-2

Horner, A. J., Bisby, J. A., Zotow, E., Bush, D., & Burgess, N. (2016). Grid-like Processing of Imagined Navigation. *Current Biology*, *26*(6), 842–847. https://doi.org/10.1016/j.cub.2016.01.042

Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F. S., & Behrens, T. E. J. (2012). Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience*, *15*(3), Article 3. https://doi.org/10.1038/nn.3017

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, *17*(2), 825–841. https://doi.org/10.1006/nimg.2002.1132

Juechems, K., Balaguer, J., Spitzer, B., & Summerfield, C. (2021). Optimal utility and probability functions for agents with finite computational precision. *Proceedings of the National Academy of Sciences*, *118*(2), e2002232118. https://doi.org/10.1073/pnas.2002232118

Julian, J. B., & Doeller, C. F. (2021). Remapping and realignment in the human hippocampal formation predict context-dependent spatial behavior. *Nature Neuroscience*, *24*(6), 863–872. https://doi.org/10.1038/s41593-021-00835-3

Julian, J. B., Keinath, A. T., Frazzetta, G., & Epstein, R. A. (2018). Human entorhinal cortex represents visual space using a boundary-anchored grid. *Nature Neuroscience*, *21*(2), Article 2. https://doi.org/10.1038/s41593-017-0049-1

Kaplan, R., Schuck, N. W., & Doeller, C. F. (2017). The Role of Mental Maps in Decision-Making. *Trends in Neurosciences*, *40*(5), 256–259. https://doi.org/10.1016/j.tins.2017.03.002

Kiani, R., Esteky, H., Mirpour, K., & Tanaka, K. (2007). Object Category Structure in Response Patterns of Neuronal Population in Monkey Inferior Temporal Cortex. *Journal of Neurophysiology*, *97*(6), 4296–4309. https://doi.org/10.1152/jn.00024.2007

Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., Lee, N., Rossa, B., Reuter, M., Chaibub Neto, E., & Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLOS Computational Biology*, *13*(2), e1005350. https://doi.org/10.1371/journal.pcbi.1005350

Knudsen, E. B., & Wallis, J. D. (2021). Hippocampal neurons construct a map of an abstract value space. *Cell*, *184*(18), 4640-4650.e10. https://doi.org/10.1016/j.cell.2021.07.010

Knutson, B., Taylor, J., Kaufman, M., Peterson, R., & Glover, G. (2005). Distributed Neural Representation of Expected Value. *Journal of Neuroscience*, *25*(19), 4806–4812. https://doi.org/10.1523/JNEUROSCI.0642-05.2005

Kolling, N., Behrens, T. E. J., Mars, R. B., & Rushworth, M. F. S. (2012). Neural Mechanisms of Foraging. *Science*, *336*(6077), 95–98. https://doi.org/10.1126/science.1216930

Kolling, N., Wittmann, M. K., Behrens, T. E. J., Boorman, E. D., Mars, R. B., & Rushworth, M. F. S. (2016). Value, search, persistence and model updating in anterior cingulate cortex. *Nature Neuroscience*, *19*(10), 1280–1285. https://doi.org/10.1038/nn.4382

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene Memory Is More Detailed Than You Think: The Role of Categories in Visual Long-Term Memory. *Psychological Science*, *21*(11), 1551–1556. https://doi.org/10.1177/0956797610385359

Kriegeskorte, N., Mur, M., Ruff, D. A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., & Bandettini, P. A. (2008). Matching Categorical Object Representations in Inferior Temporal Cortex of Man and Monkey. *Neuron*, *60*(6), 1126–1141. https://doi.org/10.1016/j.neuron.2008.10.043

Lanczos, C. (1964). Evaluation of Noisy Data. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, *1*(1), 76–85. https://doi.org/10.1137/0701007

Lee, S. W., Shimojo, S., & O'Doherty, J. P. (2014). Neural Computations Underlying Arbitration between Model-Based and Model-free Learning. *Neuron*, *81*(3), 687–699. https://doi.org/10.1016/j.neuron.2013.11.028

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology*, *22*(6), 1027–1038. https://doi.org/10.1016/j.conb.2012.06.001

Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Uğurbil, K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magnetic Resonance in Medicine*, *63*(5), 1144–1153. https://doi.org/10.1002/mrm.22361

Moneta, N., Garvert, M. M., Heekeren, H. R., & Schuck, N. W. (2023). Task state representations in vmPFC mediate relevant and irrelevant value signals and their behavioral influence. *Nature Communications*, *14*(1), Article 1. https://doi.org/10.1038/s41467-023-38709-w

Moser, E. I., Moser, M.-B., & McNaughton, B. L. (2017). Spatial representation in the hippocampal formation: A history. *Nature Neuroscience*, *20*(11), Article 11. https://doi.org/10.1038/nn.4653

Nau, M., Schröder, T. N., Bellmund, J. L. S., & Doeller, C. F. (2018). Hexadirectional coding of visual space in human entorhinal cortex. *Nature Neuroscience*, *21*(2), 188–190. https://doi.org/10.1038/s41593-017-0050-8

Navarro Schröder, T. N., Towse, B. W., Nau, M., Burgess, N., Barry, C., & Doeller, C. F. (2020). Environmental anchoring of grid-like representations minimizes spatial uncertainty during navigation. *BioRxiv*, 166306. https://doi.org/10.1101/166306

Nicolle, A., Klein-Flügge, M. C., Hunt, L. T., Vlaev, I., Dolan, R. J., & Behrens, T. E. J. (2012). An Agent Independent Axis for Executed and Modeled Choice in Medial Prefrontal Cortex. *Neuron*, *75*(6), 1114–1121. https://doi.org/10.1016/j.neuron.2012.07.023

Nyberg, N., Duvelle, É., Barry, C., & Spiers, H. J. (2022). Spatial goal coding in the hippocampal formation. *Neuron*, *110*(3), 394–422. https://doi.org/10.1016/j.neuron.2021.12.012

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, *4*(1), Article 1. https://doi.org/10.1038/82959

O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, *34*(1), 171–175. https://doi.org/10.1016/0006-8993(71)90358-1

O'Keefe, J., & Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press. https://repository.arizona.edu/handle/10150/620894

Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*(7090), Article 7090. https://doi.org/10.1038/nature04676

Palombo, D. J., Hayes, S. M., Reid, A. G., & Verfaellie, M. (2019). Hippocampal contributions to value-based learning: Converging evidence from fMRI and amnesia. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(3), 523–536. https://doi.org/10.3758/s13415-018-00687-8

Park, S. A., Miller, D. S., & Boorman, E. D. (2021). Inferences on a multidimensional social hierarchy use a grid-like code. *Nature Neuroscience*, *24*(9), 1292–1301. https://doi.org/10.1038/s41593-021-00916-3

Park, S. A., Miller, D. S., Nili, H., Ranganath, C., & Boorman, E. D. (2020). Map Making: Constructing, Combining, and Inferring on Abstract Cognitive Maps. *Neuron*, *107*(6), 1226-1238.e8. https://doi.org/10.1016/j.neuron.2020.06.030

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pelletier, G., & Fellows, L. K. (2019). A Critical Role for Human Ventromedial Frontal Lobe in Value Comparison of Complex Objects Based on Attribute Configuration. *The Journal of Neuroscience*, *39*(21), 4124–4132. https://doi.org/10.1523/JNEUROSCI.2969-18.2019

Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, *497*(7447), 74–79. https://doi.org/10.1038/nature12112

Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal Cortex Encodes Willingness to Pay in Everyday Economic Transactions. *Journal of Neuroscience*, *27*(37), 9984–9988. https://doi.org/10.1523/JNEUROSCI.2131-07.2007

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, *84*, 320–341. https://doi.org/10.1016/j.neuroimage.2013.08.048

Pruim, R. H. R., Mennes, M., van Rooij, D., Llera, A., Buitelaar, J. K., & Beckmann, C. F. (2015). ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage*, *112*, 267–277. https://doi.org/10.1016/j.neuroimage.2015.02.064

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II: Current Research and Theory: Vol. Vol. 2*.

Righi, G., Peissig, J. J., & Tarr, M. J. (2012). Recognizing disguised faces. *Visual Cognition*, *20*(2), 143–169. https://doi.org/10.1080/13506285.2012.654624

Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., Eickhoff, S. B., Hakonarson, H., Gur, R. C., Gur, R. E., & Wolf, D. H. (2013). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage*, *64*, 240–256. https://doi.org/10.1016/j.neuroimage.2012.08.052

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience*, *8*(9), Article 9. https://doi.org/10.1038/nrn2213

Schiller, D., Eichenbaum, H., Buffalo, E. A., Davachi, L., Foster, D. J., Leutgeb, S., & Ranganath, C. (2015). Memory and Space: Towards an Understanding of the Cognitive Map. *Journal of Neuroscience*, *35*(41), 13904–13911. https://doi.org/10.1523/JNEUROSCI.2618-15.2015

Schuck, N. W., Cai, M. B., Wilson, R. C., & Niv, Y. (2016). Human Orbitofrontal Cortex Represents a Cognitive Map of State Space. *Neuron*, *91*(6), 1402–1412. https://doi.org/10.1016/j.neuron.2016.08.019

Schuck, N. W., & Niv, Y. (2019). Sequential replay of nonspatial task states in the human hippocampus. *Science*, *364*(6447). https://doi.org/10.1126/science.aaw5181

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, *20*(11), 1643–1653. https://doi.org/10.1038/nn.4650

Stangl, M., Shine, J., & Wolbers, T. (2017). The GridCAT: A Toolbox for Automated Analysis of Human Grid Cell Codes in fMRI. *Frontiers in Neuroinformatics*, *11*. https://doi.org/10.3389/fninf.2017.00047

Stensola, T., Stensola, H., Moser, M.-B., & Moser, E. I. (2015). Shearing-induced asymmetry in entorhinal grid cells. *Nature*, *518*(7538), Article 7538. https://doi.org/10.1038/nature14151

Tavares, R. M., Mendelsohn, A., Grossman, Y., Williams, C. H., Shapiro, M., Trope, Y., & Schiller, D. (2015). A Map for Social Navigation in the Human Brain. *Neuron*, *87*(1), 231–243. https://doi.org/10.1016/j.neuron.2015.06.011

The Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., Earl, N., Starkman, N., Bradley, L., Shupe, D. L., Patil, A. A., Corrales, L., Brasseur, C. E., Nöthe, M., Donath, A., Tollerud, E., Morris, B. M., Ginsburg, A., Vaher, E., Weaver, B. A., Tocknell, J., Jamieson, W., … Zonca, A. (2022). *The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package*. https://doi.org/10.3847/1538-4357/ac7c74

Theves, S., Fernandez, G., & Doeller, C. F. (2019). The Hippocampus Encodes Distances in Multidimensional Feature Space. *Current Biology*, *29*(7), 1226-1231.e3. https://doi.org/10.1016/j.cub.2019.02.035

Theves, S., Fernández, G., & Doeller, C. F. (2020). The Hippocampus Maps Concept Space, Not Feature Space. *Journal of Neuroscience*, *40*(38), 7318–7325. https://doi.org/10.1523/JNEUROSCI.0494-20.2020

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208. https://doi.org/10.1037/h0061626

Tustison, N. J., Avants, B. B., Cook, P. A., Yuanjie Zheng, Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging*, *29*(6), 1310–1320. https://doi.org/10.1109/TMI.2010.2046908

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. https://doi.org/10.1007/BF00122574

Viganò, S., Rubino, V., Soccio, A. D., Buiatti, M., & Piazza, M. (2021). Grid-like and distance codes for representing word meaning in the human brain. *NeuroImage*, *232*, 117876. https://doi.org/10.1016/j.neuroimage.2021.117876

Vikbladh, O. M., Meager, M. R., King, J., Blackmon, K., Devinsky, O., Shohamy, D., Burgess, N., & Daw, N. D. (2019). Hippocampal Contributions to Model-Based Planning and

Spatial Memory. *Neuron*, *102*(3), 683-693.e4. https://doi.org/10.1016/j.neuron.2019.02.014

Wagner, I. C., Graichen, L. P., Todorova, B., Lüttig, A., Omer, D. B., Stangl, M., & Lamm, C. (2023). Entorhinal grid-like codes and time-locked network dynamics track others navigating through space. *Nature Communications*, *14*(1), Article 1. https://doi.org/10.1038/s41467-023-35819-3

Wikenheiser, A. M., & Schoenbaum, G. (2016). Over the river, through the woods: Cognitive maps in the hippocampus and orbitofrontal cortex. *Nature Reviews Neuroscience*, *17*(8), 513–523. https://doi.org/10.1038/nrn.2016.56

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., & Niv, Y. (2014). Orbitofrontal Cortex as a Cognitive Map of Task Space. *Neuron*, *81*(2), 267–279. https://doi.org/10.1016/j.neuron.2013.11.005

Wimmer, G. E., & Shohamy, D. (2012). Preference by Association: How Memory Mechanisms in the Hippocampus Bias Decisions. *Science*, *338*(6104), 270–273. https://doi.org/10.1126/science.1223252

Wittmann, M. K., Kolling, N., Akaishi, R., Chau, B. K. H., Brown, J. W., Nelissen, N., & Rushworth, M. F. S. (2016). Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nature Communications*, *7*(1), Article 1. https://doi.org/10.1038/ncomms12327

Zhang, Y., Brady, M., & Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, *20*(1), 45–57. https://doi.org/10.1109/42.906424

Acknowledgements

We would like to thank Theo Schaefer and Felix Deilmann for joint efforts in setting up MRI-compatibility of Psychopy tasks and co-development of scripts implementing MRI data analyses in Python. We would also like to thank Ulrike Horn for providing a Python script to interact with the MRI button box. We would like to thank Kerstin Schumer, Max Schulz, Anke Kummer, Simone Wipper, Sylvie Neubert, Mandy Jochemko, Nicole Pampus and Manuela Hofmann for help with data collection. Furthermore, we would like to thank Arthur Levasseur for help with piloting a previous task version. We would also like to thank Naomi de Haas for helpful input and discussions based on previous work investigating a value space. We thank the University of Minnesota Center for Magnetic Resonance Research for the provision of the multiband EPI sequence software and Toralf Mildner, Joeran Lepsien and colleagues of the Doellerlab as well as the MPI research group Adaptive Memory for joint efforts in further MRI sequence piloting. We would also like to thank all colleagues of the Doellerlab for helpful discussions of the study.

This work was supported by the Max Planck Society. NWS is funded by an Independent Max Planck Research Group grant awarded by the Max Planck Society (M.TN.A.BILD0004) and a Starting Grant from the European Union (ERC-2019-StG REPLAY-852669) and the Federal Ministry of Education and Research (BMBF) and the Free and Hanseatic City of Hamburg under the Excellence Strategy of the Federal Government and the Länder. CFD is supported by the Max Planck Society, the European Research Council (ERC-CoG GEOCOG 724836), the Kavli Foundation, the Jebsen Foundation, Helse Midt Norge and The Research Council of Norway (223262/F50, 197467/F50).

Data and code availability

Data to reproduce the figures, group-level statistical brain maps and analysis code will be made openly available.

Author contributions

A.N., M.M.G., N.W.S and C.F.D. conceived the experiment and developed the tasks. A.N. acquired the data. All authors planned the analyses. A.N. analyzed the data and all authors discussed the results. A.N. wrote the manuscript with input from all authors.