Neural systems for memory-based value judgment and decision-making

Avinash R. Vaidya[1]
David Badre[1, 2]

[1]Department of Cognitive, Linguistic, and Psychological Sciences
[2]Carney Institute for Brain Science
Brown University

RUNNING TITLE: MEMORY-BASED VALUE JUDGMENT

Keywords: decision-making; semantic memory; episodic memory; schemas; functional magnetic resonance imaging; value

Address correspondence to:
Avinash R. Vaidya
Brown University
Providence, RI 02912-1978
tel. 401-863-3634, fax 401-863-2255
email avinash_vaidya@brown.edu

Acknowledgements: This work was supported by a MURI award from the Office of Naval Research (N00014-16-1-2832), an award from the James S. McDonnell Foundation, and a Ruth L. Kirschstein National Research Service Award to A.V. (NIMH F32 MH116592-01A1). We are grateful to Michael Frank for helpful comments and suggestions on the manuscript, as well as Apoorva Bhandari and Andra Geana for constructive discussion on methods and analysis. Thanks also to Adriane Spiro for help with data collection, and Henry Jones for assistance with behavioral piloting and analysis of normative data.

Abstract

Real life choices often require that we draw inferences about the value of options based on structured, schematic knowledge about their utility for our current goals. Other times, value information may be retrieved directly from a specific prior experience with an option. In an fMRI experiment, we investigated the neural systems involved in retrieving and assessing information from different memory sources to support value-based choice. Participants completed a task in which items could be conferred positive or negative value based on schematic associations (i.e. schema value), or learned directly from experience with deterministic feedback (i.e. episodic value). We found that ventromedial prefrontal cortex (vmPFC) activity correlated with the influence of both episodic- and schema-based values on participants' decisions. Connectivity between vmPFC and middle temporal cortex also tracked the inferred value of items based on schematic associations on the first presentation of ingredients, prior to any feedback. In contrast, the striatum responded to participants' willingness to bet on ingredients as a function of the unsigned strength of their episodic memory for those ingredients' values. These results argue that striatum and vmPFC play distinct roles in memory-based value judgment and decision-making. Specifically, the vmPFC assesses the value of options based on information inferred from schematic knowledge and retrieved from episodic memory, while the striatum controls a decision to act on options based on the strength of retrieved episodic information.

Introduction

Adaptive decision-making depends on making choices that are in line with our goals. For example, when shopping for a tool to help build a table, a hammer would be a more appropriate choice than a ball of yarn. However, the opposite would be true when knitting a scarf. While it is unlikely you have tried to use yarn to build a table, you can probably infer its value for this task based on your knowledge of yarn and table construction. Structured knowledge about the world that can support this kind of inference is often referred to as a schema (Bartlett, 1932; Ghosh & Gilboa, 2014; Piaget, 1952). Building a table depends on knowledge about the materials (e.g. wood, nails) and processes (e.g. cutting, hammering) required, and their relative importance in this task. Judging the value of hammers and yarn for accomplishing these goals requires an inferential judgment about their utility based on this kind of schema-level knowledge rather than either feedback-based learning or mere retrieval of the associations one has with a table or yarn. Thus, the ability to judge the value of options that have not been directly experienced but that can be inferred based on knowledge of the situation is essential to prospective and adaptive behavior.

Importantly, value need not be inferred in all situations. Decision-making can also depend on the direct retrieval of memory for a specific event with arbitrary and selective significance. In the same shopping example, seeing a certain brand of varnish on display might remind you that your friend had recommended that varnish for your table last week — conferring positive expected value to this item. This kind of information does not depend on any kind of schematic knowledge about varnishes and tables, but instead relies on the retrieval of some positive association with a specific item from a spatiotemporally specific episode; in this case, your conversation with your friend (Tulving, 1972). Value in this

example is a memorized feature of one or a few previous episodes, which may be brought to mind through mechanisms of episodic retrieval.

Recent evidence suggests that distinct neurobiological systems are required for value judgments that rely on different memory sources. Lesion and temporary inactivation studies in rodents, non-human primates and human patients have demonstrated that ventromedial prefrontal cortex (vmPFC) and orbitofrontal cortex (OFC) are necessary for decision-making when option values must be inferred from knowledge of task structure, but not when values can be retrieved from directly learned associations (Bradfield, Dezfouli, van Holstein, Chieng, & Balleine, 2015; Izquierdo, Suda, & Murray, 2004; Jones et al., 2012; Reber et al., 2017; Rudebeck, Saunders, Prescott, Chau, & Murray, 2013). These findings have been taken as support for the theory that these regions encode schematic representations of task environments (Chan, Niv, & Norman, 2016; Schuck, Cai, Wilson, & Niv, 2016; Wilson, Takahashi, Schoenbaum, & Niv, 2014), and are generally in line with the hypothesis that this region has a necessary role in schema memory (Ghosh, Moscovitch, Melo Colella, & Gilboa, 2014; Kan, Larocque, Lafleche, Coslett, & Verfaellie, 2010; Moscovitch & Melo, 1997; Schnider, 2013). These schemas provide a knowledge structure through which value can be inferred to inform a choice when it is not known from direct experience.

A parallel line of research has begun to explore the role of the basal ganglia, and particularly the striatum, at the intersection of goal-directed retrieval and value-based decision-making. Recent work has argued that the striatum is involved in assessing the value of information stored in episodic memory and using this information to bias cortical representations of choices with higher value (Shadlen & Shohamy, 2016). A mutually compatible alternative to valuation of memory is that striatum integrates memory signals

into its basic decision making functions, including those related to memory itself (Han, Huettel, Raposo, Adcock, & Dobbins, 2010; Scimeca & Badre, 2012), though other work has suggested that these may be separate signals from those that assess the value of outcomes (Elward, Vilberg, & Rugg, 2015).

Value judgments based on different memory sources may also depend on pathways between regions involved in the storage and retrieval of relevant information and those that utilize this information for value judgment. vmPFC has reciprocal connections with regions of middle temporal cortex (MTC) and anterior ventrolateral prefrontal cortex (aVLPFC) (Price, 2007; Saleem, Kondo, & Price, 2008). These latter regions likely play an important role in semantic memory storage and retrieval, respectively (Badre, Poldrack, Pare-Blagoev, Insler, & Wagner, 2005b; Badre & Wagner, 2007; Barredo, Oztekin, & Badre, 2015; Barredo, Verstynen, & Badre, 2016; Binder & Desai, 2011; Jefferies & Lambon Ralph, 2006; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). Both vmPFC and striatum also receive input from the hippocampus (HPC), a region known to be critical for detailed episodic memory (Ergorul & Eichenbaum, 2004; Milner, Corkin, & Teuber, 1968; Moscovitch, Nadel, Winocur, Gilboa, & Rosenbaum, 2006). These structural and functional interactions are well placed to serve value judgments based on episodic and schematic memory. However, how these regions interact during adaptive retrieval and evaluation of information from memory is not clear.

In the current experiment, we investigated the neural systems involved in making value judgments based on different memory sources. In a functional magnetic resonance imaging (fMRI) experiment, we asked participants to complete a task during which they judged the value of several distinct items based on their associations with a positive and

negative schema. Participants would then learn the value of each item through deterministic feedback. After this initial experience, they could subsequently remember this outcome to guide their decision in two repeated presentations of that item. Thus, this task demanded that participants use both schema-level knowledge about real-world associations and information "cached" in episodic memory about item values acquired from task feedback.

By manipulating inferred versus cached value in this task, we were able to test three hypotheses about memory and value-based decision making. (1) vmPFC is preferentially engaged when value judgments are based on structured knowledge about the world (i.e. schema memory). (2) Striatum tracks the value of information cached in episodic memory. (3) vmPFC and striatum interact with different memory stores in accord with their respective roles in inferred (vmPFC) versus cached (striatum) value.

Results

Participants were told that they were going to take the role of a restaurant chef choosing whether or not to serve ingredients to a customer. Participants were instructed that if they fed the customer an ingredient that the customer liked, they would receive 50 points, however if they fed the customer an ingredient that the customer did not like, they would lose 25 points. If participants chose not to feed an ingredient to the customer, they would receive 0 points. Thus, participants were asked to make a bet on each trial regarding their belief that the customer would like or dislike the displayed ingredient. Importantly, the customer's preference for a given ingredient was always signaled by a cue above the correct response during feedback, regardless of whether the participant chose to feed the customer or not, and all customer preferences were deterministic. Thus, participants received complete information on each trial regardless of their response. The feedback provided a

source for learning a cached value about an individual ingredient in episodic memory. If the participant did not respond in time, no information about the customer's preferences was given and participants were shown the word 'slow' in place of feedback about points earned.

In addition to the information about ingredient preferences provided through feedback, participants were instructed at the start of each block that they were serving a new customer who liked one recipe (positive recipe) and disliked another (negative recipe). In this way, the participant could also infer the value of an ingredient on the basis schematic knowledge about these recipes, without having received feedback. For example, if the customer liked cake and disliked soup, onions might be conferred a negative value.

Participants were also instructed before the start of the task that each customer liked and disliked other ingredients that were unrelated to these recipes (i.e. 'independent ingredients'). The value related to these ingredients could not be anticipated in advance based on their schematic associations, but rather would have to be acquired through experience.

Each ingredient was repeated three times throughout a block whether it was a recipe item or independent item. Thus, the experiment had three factors by design: ingredient clustering (recipe or independent), value (positive or negative) and presentation number (1, 2, 3). For the purposes of behavioral and fMRI analysis, any missed responses were not considered in the count of ingredient presentations, as no value information was communicated to the participant on these trials (e.g. if a participant missed the response deadline on the first presentation, the second presentation would be considered the first, and the third presentation would be considered the second in all analyses).

In each trial, participants saw an ingredient and were asked to respond by moving a slider to either side of a bar to indicate their decision to serve an ingredient and their confidence in that decision with a single response, similar to (Scimeca, Katzman, & Badre, 2016) (Figure 1). Moving the slider past the midway point to either side was considered a binary selection of that side's response, while the distance of the slider from the midway point provided a continuous index of confidence. This slider position could also be interpreted as a measure of participants' belief about the value of each ingredient by taking the distance of the slider from either end of bar (i.e. participant-P(yes)).

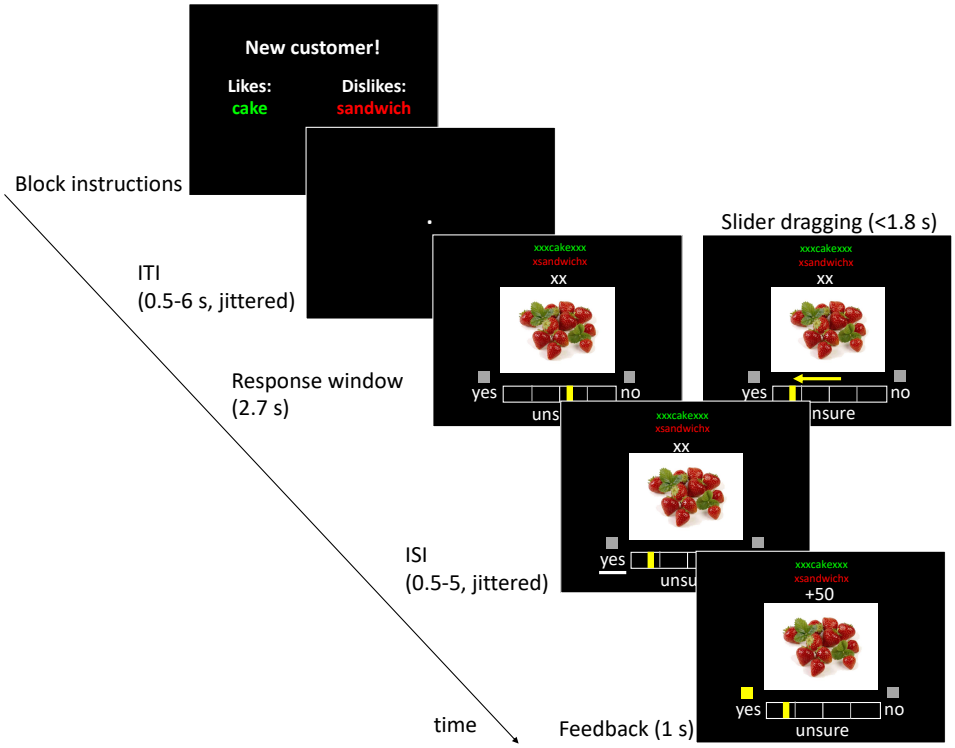


Figure 1. Task schematic showing instructions at the start of a block and organization of a single example trial. At the start of the block, participants were informed of the recipes that customer liked and disliked. These recipe names were presented in green and red above the ingredient on all trials. In each trial, participants indicated their decision to feed the displayed ingredient to the current customer, and their confidence in this decision, using a slider. Participants then saw whether or not the customer liked the ingredient, signaled by a box above either response cue, and the points earned or lost in each trial. ITI (inter-trial interval), ISI (inter-stimulus interval).

Behavioral data

We first examined whether participants inferred which ingredients were valuable based on customers' recipe preferences (Figure 2a). Indeed, on the first presentation of ingredients, before any feedback, participants accepted ingredients associated with the positive value recipe more often than ingredients associated with the negative value recipe and more often than positive value independent ingredients. This was evident in a significant three-way interaction between value, clustering and presentation number on participants' choices (RM-ANOVA: $F_{2,44} = 27.20$, $P < 0.0001$, $\eta_p^2 = 0.55$), and direct contrasts between the ingredient types (Bonferroni corrected t-tests: $t(22) \geq 7.79$, P's < 0.0001, d's ≥ 1.62).

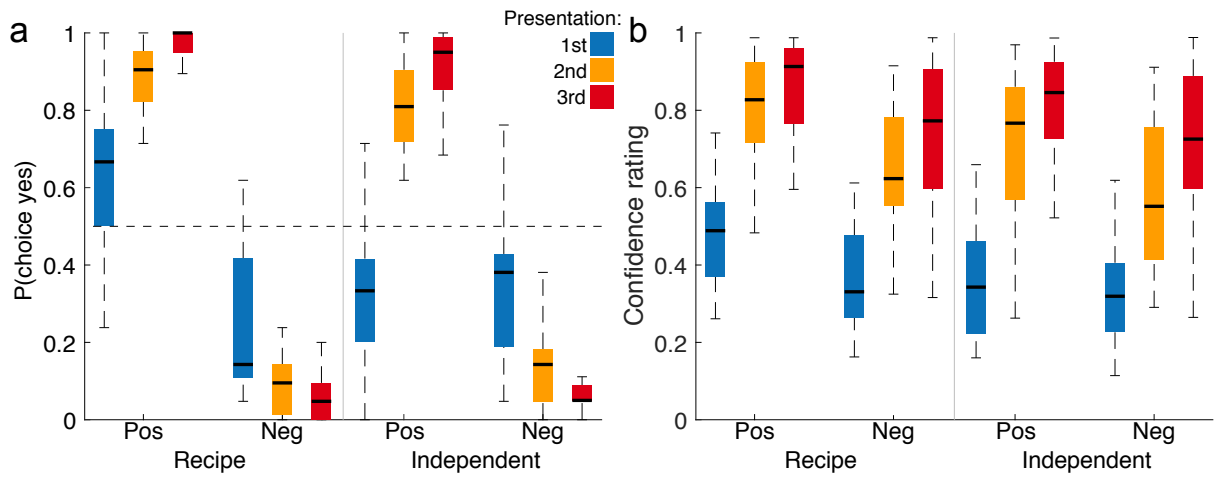


Figure 2. Choices and confidence ratings for positive (pos) and negative (neg) ingredients in the recipe and independent conditions across presentations. **a.** Frequency of 'yes' choices, **b.** Confidence ratings for correct responses. Boxes represent 25th, 50th and 75th quantiles of the data. Whiskers are approximately ±2.7 SD.

In contrast, there was no significant difference between acceptance rates for positive and negatively valued independent ingredients on the first presentation. This was as expected, as the ingredients in this condition did not have inferable value based on an association with a positive or negative recipe. The interaction between value and clustering on participants' choices was confirmed with a separate RM-ANOVA on the first presentation

only ($F_{1,22} = 61.31$, $P < 0.0001$, $\eta_p^2 = 0.73$). Participants accepted ingredients associated with the negative value recipe numerically less often than negative independent ingredients, but this difference was not significant after correction for multiple comparisons. Overall, the choice results indicate that participants successfully inferred value information from recipe associations, though these effects were stronger for the positively valued recipe.

There was clear evidence that participants learned ingredient values from feedback. Participants made choices consistent with the values of ingredients more often in the second and third presentations than in the first in all conditions, demonstrating retention of feedback from previous trials (Bonferroni corrected t-tests: t's(22) ≥ 4.54, P's < 0.01, d's ≥ 0.95). Similarly, participants' rate of value consistent choices increased significantly from the second presentation to the third in all conditions, except for independent positive ingredients (Bonferroni corrected t-tests: t's(22) ≥ 4.69, P's < 0.01, d's ≥ 0.98).

Participants also tended to reject both positive and negative independent ingredients in the first presentation significantly more often than chance (One sample, t-tests: t's(22) ≥ 8.69, P's < 0.0001, d's ≥ 1.81). Thus, participants appeared to strategically avoid endorsing independent ingredients before learning their values through feedback.

We next examined how participants' reported confidence changed on correct responses throughout the task (Figure 2b). Confidence increased with repeated presentations, and was also greater overall for recipe and positive value ingredients (RM-ANOVA: F's ≥ 27.88, $P \leq 0.0001$, $\eta_p^2 \geq 0.58$). Confidence for positive recipe ingredients was significantly higher than positive independent ingredients on the first presentation (Bonferroni corrected t-test: $t(20) = 4.57$, $P < 0.01$, $d = 0.99$) and third presentation (Bonferroni corrected t-test: $t(20) = 4.96$, $P < 0.01$, $d = 1.08$). However, there was no

difference in confidence for rejections of negative recipe and independent ingredients, resulting in a significant interaction between clustering and value (RM-ANOVA: $F_{1, 20} = 9.28$, $P = 0.006$, $\eta_p^2 = 0.32$),

Confidence also increased significantly from the first presentation to the second and third presentations in all conditions (Bonferroni corrected t-tests: t's(20) ≥ 4.57, P's ≤ 0.01, d's ≥ 1.29), and from the second to the third presentation in all conditions, except for negative recipe ingredients (Bonferroni corrected t-tests: t's(20) ≥ 3.98, P's ≤ 0.05, d's ≥ 0.87), leading to a modest interaction of value and presentation number ($F_{2, 40} = 3.34$, $P = 0.04$, $\eta_p^2 = 0.14$).

Supplementary analysis showed that both RT and confidence depended on response accuracy. Error trials were marked by reduced confidence for recipe ingredients on the first presentation, consistent with a failure to retrieve relevant information as a basis for these decisions (see Supplementary Results). Following the first presentation (on presentations two and three), errors were associated with reduced confidence and longer RT compared to correct responses. This pattern is consistent with longer RT on errors reflecting retrieval failures (Ratcliff, 1978). Participants' RT on correct trials showed a similar pattern to confidence ratings, speeding with presentation number, with fastest responses for positive recipe ingredients (Supplementary Figure 2).

In summary, behavioral results established that participants were able to infer the value of ingredients based on their recipe association and used these inferred values to make choices. Participants also showed a generally risk averse attitude on this first presentation by tending to reject independent ingredients. On subsequent presentations of an ingredient, participants accurately retrieved cached memory of deterministic feedback to make faster, more confident decisions. This was the case across both recipe and independent items.

244 However, participants continued to benefit from recipe associations in these later
245 presentations, with generally faster reaction times and greater confidence for correct
246 responses in this condition compared to independent ingredients. Finally, they tended to
247 asymmetrically utilize knowledge about the positive recipe more than the negative recipe
248 throughout all trials.

249 *Main effects of task conditions in whole-brain analyses*

250 In assessing fMRI effects at a whole-brain level, we first tested how ingredient
251 clustering, value, presentation number and accuracy affected activity using simple contrasts
252 from our condition-based general linear model (GLM; Supplementary Table 1). These
253 analyses were filtered to focus on correct responses, excluding trials where participants may
254 have made errors due to failure to appropriately recall information relevant to ingredient
255 values.

256 Recipe ingredients elicited greater activity than independent ingredients in medial
257 OFC (Figure 3a), but not elsewhere. Ingredients associated with positive value elicited
258 greater activity in OFC and medial PFC, as well as the ventral striatum (Figure 3b). Compared
259 with later ingredient presentations, the first encounter with an ingredient was associated
260 with widespread activity in lateral PFC, including left VLPFC, inclusive of the aVLPFC, and
261 the caudate nucleus (Figure 3c). Activity in left MTC, bilateral HPC and putamen was higher
262 in correct responses compared to errors across all presentations (Figure 3d).

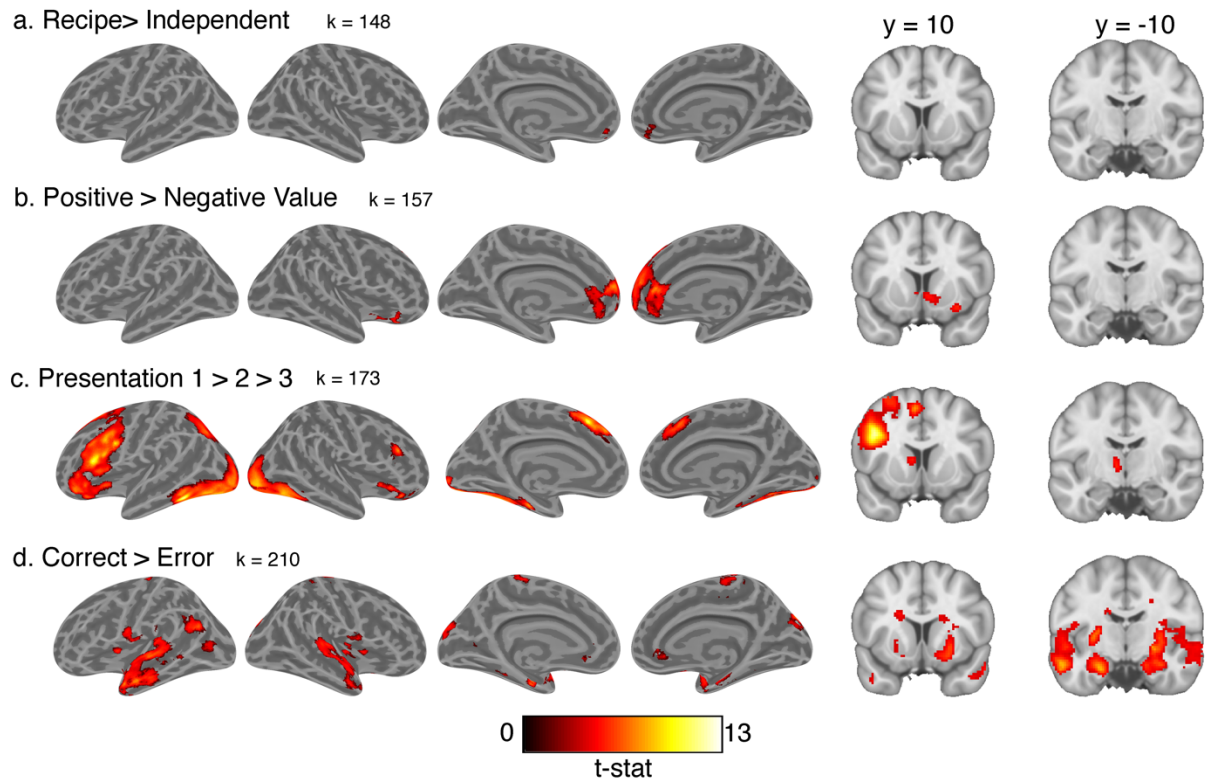


Figure 3. Whole brain contrasts projected onto inflated cortical surfaces and coronal slices for main effects of **a**. ingredient-recipe clustering, **b**. ingredient value, **c**. presentation number and **d**. response accuracy. All images are thresholded at $P < 0.05$, FWE cluster corrected. All main effects, except for accuracy, are filtered to only include correct responses. Cluster extent threshold for each contrast is given by value of k.

Region of interest analyses

To follow-up the whole-brain voxel-wise contrasts, we tested for more specific task-dependent changes in activity within a set of *a priori* regions of interest (ROIs) defined in Table 1 using a finite impulse response (FIR) model of the hemodynamic response. We were primarily interested in how the clustering status of ingredients influenced value-related activity across ingredient repetitions on correct responses. Here we focused on vmPFC and the striatum, since we had *a priori* expectations that these regions would differentially process value related to episodic and schematic memory. As the medial frontal wall is functionally heterogeneous (Braga & Buckner, 2017; Roy, Shohamy, & Wager, 2012), we

included two ROIs from this region, one that was defined by a meta-analysis of studies of expected-value (Bartra, McGuire, & Kable, 2013), and a more dorsal ROI (mPFC) that responded to shared associations between components of a recipe schema in another experiment (Barron, Dolan, & Behrens, 2013). This dorsal mPFC ROI responded to ingredient value, but not any other task manipulations and is described in the Supplementary Results.

Table 1. Description of ROIs

Region of interest	Definition	Center of mass (MNI)
vmPFC	Positive>Negative effects of subjective value (Bartra et al., 2013), masked using AAL definition of vmPFC from same reference	-1, 44, -7
mPFC	10 mm sphere around peak activation in a preference construction task (Barron et al., 2013)	-6, 53, 13
HPC	Left and right hippocampal AAL masks (Tzourio-Mazoyer et al., 2002)	
Striatum	Left and right putamen and caudate AAL masks	
aVLPFC	10 mm sphere around peak activation in semantic retrieval task (Badre et al., 2005b)	-54, 27, -9
MTC	10 mm sphere around peak activation in semantic classification task (Thompson-Schill et al., 1997)	-52, -55, 2

vmPFC was more active for positive than negative expected value, evident in a main effect of value (Figure 4a; RM-ANOVA: $F_{1,20} = 9.13$, $P = 0.007$, $\eta_p^2 = 0.31$), consistent with the whole-brain analysis. Activity in this region was greater for positive compared to negative recipe ingredients in the third presentation (Bonferroni corrected t-test: $t(20) = 4.90$, $P < 0.01$, $d = 1.07$), but there was no comparable value difference for independent ingredients (Bonferroni corrected t-test: $t(20) = 1.13$, $P > 0.05$, $d = 0.25$). However, there was also no significant difference between vmPFC activity in response to positive and negative recipe ingredients on the first presentation, leading to a three-way interaction between value,

presentation number and clustering ($F_{1, 20} = 5.06$, $P = 0.01$, $\eta_p^2 = 0.20$). This null result of value on the first presentation was not merely due to the ROI selection; an exploratory whole-brain contrast of positive and negative recipe conditions on the first presentation in the condition-based GLM was similarly null. vmPFC activity was generally also higher in later presentations, reflected in a marginal main effect of presentation number ($F_{2, 20} = 3.29$, $P = 0.05$, $\eta_p^2 = 0.14$), but there was no significant main effect of clustering ($F_{1, 20} = 1.42$, $P = 0.2$, $\eta_p^2 = 0.06$).

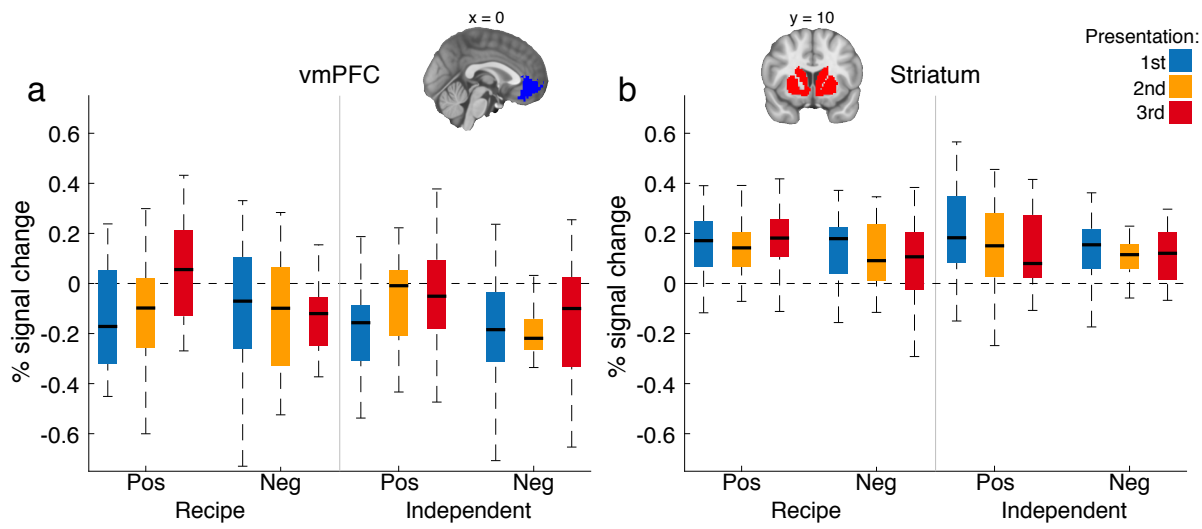


Figure 4. Region-of-interest analyses of **a.** ventromedial prefrontal cortex (vmPFC) and **b.** the bilateral striatum showing percent signal change for correct responses in each condition.

Within the striatum, there was greater activation for positive than negative value, evident in a strong main effect of value (Figure 4b; $F_{1, 20} = 9.17$, $P = 0.007$, $\eta_p^2 = 0.31$), consistent with the observation from the whole brain voxel-wise analysis. However, there were no other significant main effects or interactions.

We next tested secondary questions regarding how activity in MTC, HPC and aVLPFC responded to demands on episodic and schema memory, and how these responses were modulated by the accuracy of participants' choices (see Supplementary Results for details).

301 In brief, HPC activity was greater on correct responses than errors, unexpectedly more so for
302 recipe ingredients than independent ingredients. MTC activity appeared to respond to a
303 reduced need for semantic retrieval in later presentations of independent ingredients. In
304 contrast, aVLPFC responded principally to the novelty of these ingredients with no
305 modulatory influence of ingredient clustering or response accuracy.

306 In summary, ROI-based analyses confirmed whole-brain contrasts demonstrating
307 that the vmPFC and striatum carried information about ingredient values. Unlike striatum,
308 value-related activity in vmPFC was greater for recipe ingredients compared to independent
309 ingredients in later presentations, but not in the first presentation where the demand for
310 inferring ingredient values from recipe associations was the greatest.

311 *Computational modelling of behavior*

312 Behavioral evidence demonstrated that participants inferred the values of recipe
313 ingredients on the first presentation before receiving feedback, however whole-brain and
314 ROI-focused contrasts did not reveal any activity differences reflecting this process. We had
315 hypothesized that vmPFC would be preferentially engaged by schema-based values, while
316 the striatum would be involved in evaluating the contents of episodically encoded value
317 information. While value-related differences in vmPFC activity appeared to grow larger in
318 later presentations of recipe ingredients compared to independent ingredients, this change
319 in activity could not be easily differentiated from value information learned from feedback.
320 Importantly, however, this condition-based analysis ignores the continuous nature of
321 associations between independent ingredients and recipes. For example, strawberries are
322 considered part of the 'cake' cluster, but also have some association with 'pie.'

To better separate these components of expected value for testing our first two hypotheses regarding the neural systems supporting valuation of information from different memory sources, we developed a computational model of behavior to estimate the contribution of episodic and schema information to participants' choices in each trial. We then used these values in a model-based parametric GLM to determine the correlates of this information during value judgment.

Our computational model of behavior made a few basic assumptions about learning in this task. Within this model, each ingredient had an 'episodic value' (*EV*), which depended on feedback and decayed over time, and a 'schema value' (*SV*), which was derived from the differences in associative strength between an ingredient and the positive and negative recipes. The associative strength of an ingredient with each recipe was based on a latent semantic analysis of normative ingredient-recipe associations (Dumais, 2004; Landauer, Foltz, & Laham, 1998). Each of these value terms were weighted by a temperature parameter that determined their contribution to the model's output, a continuous estimate of the probability that it would endorse an ingredient, which we will refer to as model-P(yes).

Starting with these assumptions, we followed a step-wise process of adding parameters to a model with these features and comparing the Akiake Information Criterion (AIC) of the new model against the best simpler model. This allowed us to test if the new model better explained participants' behavior while penalizing for the addition of a new parameter. The details of comparisons between these alternate models are reported in the Supplementary Methods.

The final best fitting model allowed for a more sophisticated parameterization of SV where the number of dimensions in the latent semantic analysis of recipe-ingredient

associations was allowed to vary. This parameter essentially controlled the depth of recipe information considered in computing *SV* (from simple distinctions like savory-sweet to more fine-grained information about particular recipes; Supplementary Figure 1). The model also allowed participants' perception of the similarity of an ingredient's associations with positive and negative recipes to act as a weight on the respective contributions of the difference and sum of these associations in the calculation of *SV*. This enabled the model to determine whether ingredients, like the strawberry example above, would be attributed positive, negative, or null value if it had a strong association with both the positive and negative recipes (e.g. cake and pie).

This best fitting model also allowed *unfamiliarity* with an ingredient to influence decisions about betting or passing. Unfamiliarity was operationalized as the opposite of the unsigned episodic value of items. In other words, unfamiliarity was highest when ingredients were seen on the first presentation, or when ingredients had not been seen for many trials and their episodic value had decayed. This parameter thus effectively accounted for variance in participants' willingness to bet on ingredients when information about their episodic values were weaker. Lastly, an additional free parameter acted as an intercept term in the decision function, which effectively acted as a fixed bias on the model's tendency to bet on ingredients.

We first examined the covariance between parameters recovered from fitting this model to participants' behavioral data to test their independence (Table 2). The fixed bias parameter was negatively correlated with *SV* temperature, the schema similarity temperature, and the memory bias parameter. There was also a positive correlation between the recipe dimensions and *SV* temperature parameters, likely due to the cosine similarities

between ingredients growing smaller on average with higher recipe dimensions (Supplementary Figure 1).

Table 2. Mean parameter values and Pearson correlations between parameters in the best-fitting behavioral model

	Mean (SD)	**Decay** 0.02	**EV temp.** 5.05	**SV temp.** 2.07	**Schema similarity temp.**	**UF temp.**	**Fixed bias**
Decay	**0.02 (0.01)**	-					
EV temp.	**5.05 (0.32)**	0.35	-				
SV temp.	**2.07 (1.42)**	0.05	0.02	-			
Schema similarity temp.	**0.33 (1.75)**	-0.27	-0.01	0.25	-		
UF temp.	**-0.81 (0.99)**	0.09	0.31	0.31	0.06	-	
Fixed bias	**0.55 (0.83)**	0.15	-0.17	-0.45*	-0.47*	-0.85***	-
Recipe dimensions (log-space)	**2.56 (1.04)**	-0.04	0.07	0.60**	0.21	-0.04	-0.11

Episodic value (EV), schema value (SV), unfamiliarity (UF), *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$

We next tested if there were general tendencies for the two bias and schema similarity temperature parameters towards positive or negative values by conducting two-tailed, one-sample t-tests against zero. While the unfamiliarity bias parameter was significantly lower than zero ($t(22) = 3.91$, $P = 0.0008$, $d = 0.81$), the opposite was true of the fixed bias parameter ($t(22) = 3.18$, $P = 0.004$, $d = 0.66$). The schema similarity temperature parameter was not significantly different from zero ($t(22) = 0.90$, $P = 0.4$, $d = 0.19$). Thus, the unfamiliarity temperature parameter appeared to capture participants' aversion to betting on ingredients on their first presentation (i.e., before they had not learned the value of the item), while the fixed bias parameter exerted an inverse tendency across all trials,

biasing the model towards endorsing ingredients in opposition to the unfamiliarity parameter.

To confirm that all of these parameters could be uniquely recovered in our model fitting procedure, we generated 100 simulated participants with randomly assigned parameters using the best fitting model. We also tested if we could recover the same parameters that were found for participants' actual responses. In both cases we were able to recover these parameters from simulated data with very high fidelity (Supplementary Results).

Finally, we examined the simulated behavior of the model to test if it could recapitulate key features of participants' behavior across task conditions. To directly compare the model and participants' behavior, we focused on participants' continuous estimate of whether or not they should endorse an ingredient, based on position of the slider between the 'yes' and 'no' ends of the bar (i.e. participant-P(yes)). We compared these ratings to the model-P(yes) measure, and transformed model-P(yes) to confidence based on the distance from the point of maximum uncertainty (i.e. 0.5). Participant-P(yes) ratings followed a similar pattern to their averaged binary choices (Figure 5a; Supplementary Figure 4), including key features on the first presentation like higher P(yes) for positive recipe ingredients compared to negative recipe, and positive independent, ingredients, and a bias against endorsing independent ingredients.

We next examined the model-P(yes) and confidence ratings in the same way as participants' responses (Figure 5b). The model's behavior captured many of the same features observed in participants, including sensitivity to value information based on schematic associations on the first presentation in absence of feedback, and an aversion to

betting on independent ingredients on the first presentation. In subsequent presentations, the model also followed the values of options learned through feedback, with an advantage for recipe ingredients in model-P(yes) and confidence. However, the model utilized negative recipe associations more than participants, and did not show the same asymmetric greater confidence for positive value ingredients. An alternative model that explicitly gave different weights to positive and negative recipe associations did not pass our model fitting procedure, but marginally improved the model fitness, likely by accounting for this asymmetry.

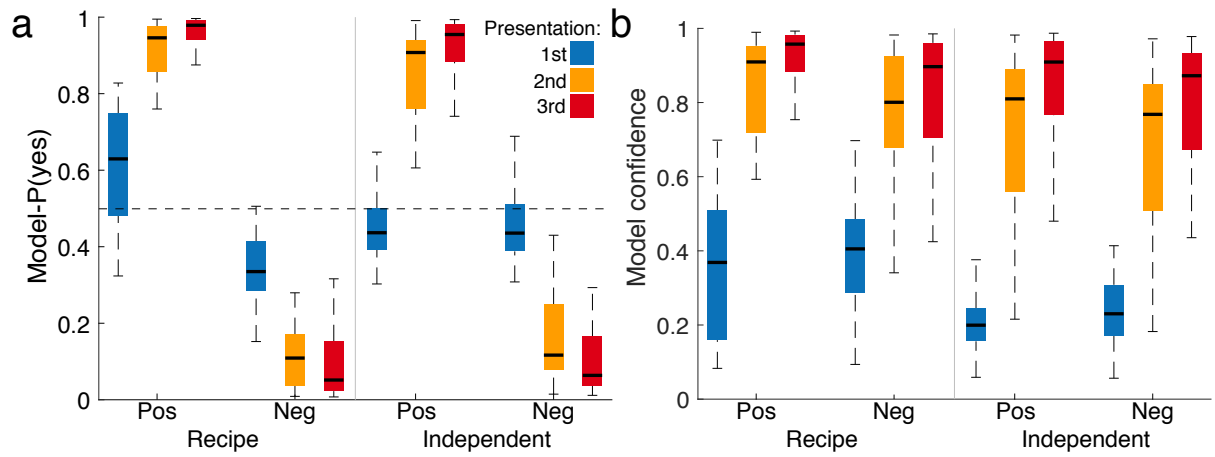


Figure 5. Model behavior across task conditions. **a.** Model-P(yes), a continuous indicator of the model's estimate of the value of these options. **B.** Model confidence on correct responses only.

In summary, our model captured core aspects of participants' behavior in this task, including utilization of recipe-based values and a bias against endorsing independent ingredients on the first presentation. The model was also able to recreate the basic pattern of participants' reported confidence rising with repeated presentations, with benefits for positive value and recipe associations. However, the model also gave greater weight to the negative value recipe than participants.

Model-based fMRI

As noted above, these model-derived analyses can provide greater sensitivity for detecting effects in fMRI than the condition-based analysis, and can tie these effects to interpretable psychological constructs. However, before testing for correlates of different model-derived aspects of value in the BOLD response, we first assessed whether the fMRI correlates of the model's predicted behavior were similar to those of participants' real continuous ratings in this task. These validation analyses found that model-derived terms for P(yes) and confidence, as well as feedback associated activity, were associated with a broadly very similar pattern of activity as recovered from terms based on participants' own ratings (Supplementary Figures 5, 6 and 7). Thus, the internal representations used in this model were able to reproduce behavior that correlated with a similar pattern of activity as participants' own ratings.

Having validated our model-based approach, we tested for correlates of the three terms that influenced value judgment in this behavioral model: schema value, episodic value and unfamiliarity. These terms were not always in alignment with participants' value judgments, particularly when the model and participants' behavior diverged (e.g. when participants failed to recall learned values). To capture activity that corresponded to the modulation of participants' real value judgments by these model-based terms, we also included interactions between these three terms and participant-P(yes). This model also included parametric modulators for participants' confidence ratings and ingredient subjective value ratings on the response regressor.

Whole-brain cluster corrected analyses found no significant main effects of participant-P(yes), episodic or schema value, or participants' subjective value ratings. There

were significant correlates for the main effect of unfamiliarity (Figure 6a), with increased activity in lateral PFC and visual cortex, similar to the main effect of presentation number in condition-based analysis above.

The interaction of schema value and participant-P(yes) was associated with activity in the pregenual ACC and medial OFC, as well as the right VLPFC and bilateral supramarginal gyri (Figure 6b). The interaction of episodic value and participant P(yes) was also associated with activity in medial OFC and a large area of the medial superior frontal gyrus and left angular gyrus (Figure 6c). There were no clusters that passed the whole-brain FWE cluster-corrected threshold for the interaction of unfamiliarity and participant-P(yes). However, given the challenge of detecting cluster corrected effects in deep brain structures, we applied a small volume correction using the striatum ROI. This analysis found that two clusters in the bilateral caudate nucleus responded to the interaction of unfamiliarity with participant-P(yes) (Figure 6d; FWE-corrected P's ≤ 0.03, MNI: 18, 20, -4; -8, 22, 0). Whole brain peak activations for this GLM are described in Supplementary Table 4.

We next specifically compared how activity related to participant-P(yes) in vmPFC and striatum, the two regions we identified *a priori* as being associated with valuation, was modulated by episodic value, schema value and unfamiliarity. We extracted the mean beta coefficients for the interaction between participant-P(yes) and these three terms in the vmPFC and striatum ROIs. The interaction of participant-P(yes) with episodic and schema value was significant in vmPFC (One-tailed, one sample t-tests against zero: t's(22) ≥ 2.99, P ≤ 0.003, d's ≥ 0.62), but not the striatum (t(22) ≤ 1.25, P ≥ 0.1, d's ≤ 0.26), whereas the interaction of participant-P(yes) with unfamiliarity was significant in the striatum (t(22) = 2.65, P = 0.007, d = 0.55), but not vmPFC (t(22) = 1.30, P = 0.1, d = 0.27). Thus, vmPFC

468 responded to greater agreement between both episodic and schema values with participant-
469 P(yes), while striatum activity increased in response to participant-P(yes) when
470 unfamiliarity was high.

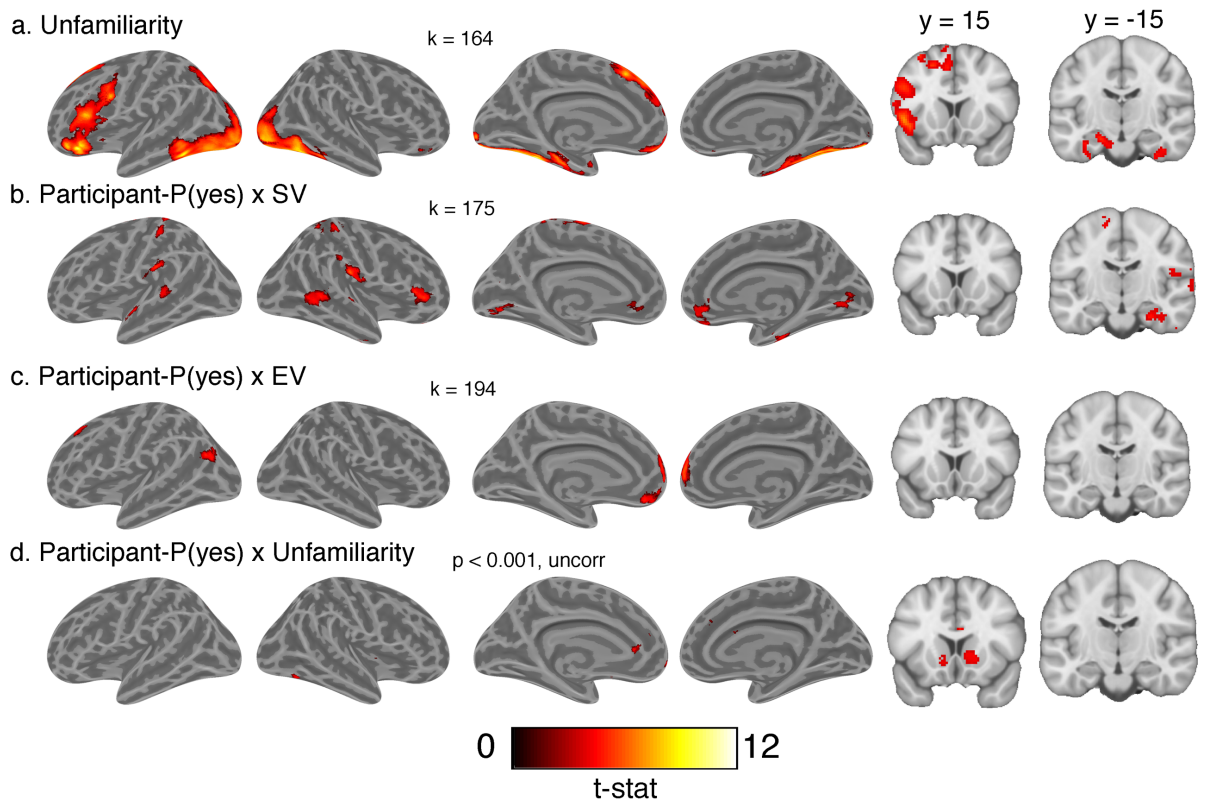


Figure 6. Whole brain t-statistic maps for model-based value terms and their interaction with partcipant-P(yes). **a.** Unfamiliarity, **b**. Interaction of participant-P(yes) and schema value (SV), **c.** Interaction of participant-P(yes) and episodic value (EV), **d**. Interaction of participant-P(yes) and unfamiliarity. All images are thresholded at $P < 0.05$, FWE cluster corrected, except for those shown in **d**. Cluster extent threshold for each map is given by value of k.

471

472 To examine whether ROIs differed in their response to these value-related terms, we
473 then tested an ROI (striatum, vmPFC) by modulator (interactions of participant-P(yes) with
474 episodic value, schema value and unfamiliarity) interaction on these average beta
475 coefficients (Figure 7). This interaction was highly significant ($F_{2,44} = 12.01$, $P < 0.0001$, $\eta_p^2 =$
476 0.35), as was a main effect of ROI ($F_{1,22} = 14.09$, $P < 0.01$, $\eta_p^2 = 0.39$), but not modulator ($F_{2,44}$

= 0.25, P = 0.8, η_p^2 = 0.01). To test for the source of this interaction, we compared the difference of the betas in the vmPFC and striatum for each modulator. This difference was significantly greater for the interaction of participant-P(yes) with episodic value compared to interactions with schema value or unfamiliarity (Bonferroni corrected t-tests t's(22) ≥ 3.75, P's < 0.01, d ≥ 0.78), but there was no comparable significant difference between schema value and unfamiliarity.

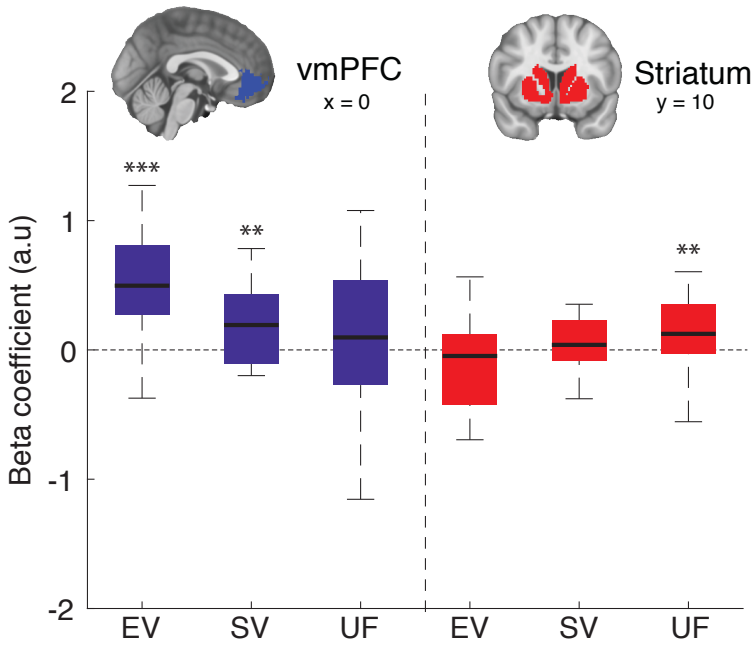


Figure 7. Boxplots showing mean beta coefficients from voxels in ventromedial prefrontal cortex (vmPFC) and striatum ROIs for interactions of participant-P(yes) with episodic value (EV), schema value (SV) and unfamiliarity (UF). *** P < 0.001, ** P < 0.01, one-tailed one-sample t-test against zero.

To summarize, vmPFC activity reflected the influence of episodic and schema values on participants' value judgments, and preferentially reflected the influence of episodic values on participants' value judgments compared to the striatum. Striatal activity reflected the effect of unfamiliarity on participants' value judgments, but not value information cached in episodic memory or inferred from schematic associations.

488 *Generalized psychophysical interactions*

489 We hypothesized that differential demand on memory systems during decision-
490 making would also lead to distinct patterns of functional connectivity, reflecting the
491 contributions of different memory stores to the assessment of options. Specifically, we were
492 interested in whether vmPFC and MTC connectivity increased with greater demands on
493 schema-based decision-making, and whether striatum and HPC connectivity increased
494 during the assessment of episodically encoded values. To address this hypothesis, we carried
495 out generalized psychophysical interaction (gPPI) analyses that contrasted how connectivity
496 between pairs of ROIs listed in Table 1 changed between task conditions, as well as more
497 exploratory seed-to-voxel whole-brain gPPI analyses focused on vmPFC, MTC, HPC and
498 striatum.

499 We first examined whether connectivity with vmPFC and MTC was modulated by
500 demands on schematic knowledge about ingredients and their value on correct responses in
501 the first presentation. Within an MTC seed, there was increased connectivity with vmPFC on
502 the first presentation of positive recipe ingredients compared to negative recipe ingredients
503 (Bonferroni corrected t-test: $t(21) = 4.00$, $P < 0.01$, $d = 0.85$), but not any other ROI ($P \geq 0.2$,
504 Bonferroni corrected). Similarly, using vmPFC as a seed ROI revealed increased connectivity
505 with the MTC ROI in the same contrast (Bonferroni corrected t-test: $t(21) = 3.68$, $P < 0.05$, d
506 $= 0.78$), but not any other ROI. An exploratory whole-brain seed-to-voxel analysis on this
507 contrast with the MTC ROI as a seed found that this region had significantly greater
508 connectivity with a cluster on the medial frontal wall overlapping with our vmPFC ROI
509 (Figure 8, MNI: -2, 46, -12), and a second cluster in lateral occipital cortex (MNI: -52, -68, 12).

The same whole-brain ROI-to-voxel contrast with the vmPFC ROI as a seed did not reveal any significant clusters at an FWE cluster-corrected threshold.

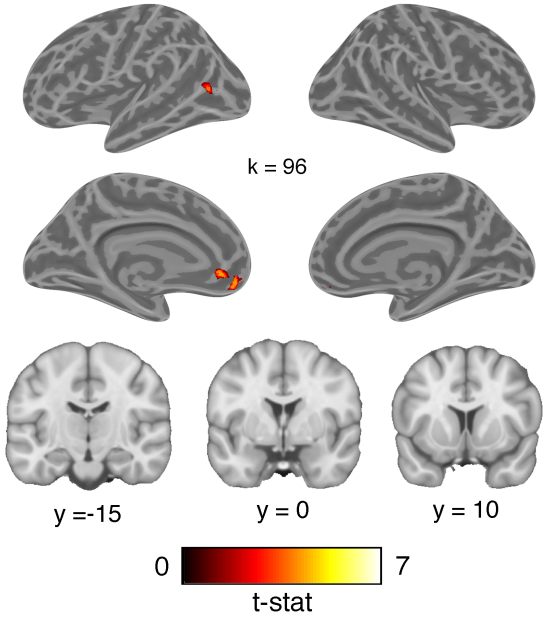


Figure 8. Whole brain t-statistic maps for seed-to-voxel generalized psychophysical interaction analyses contrasting the first presentation of positive recipe ingredients than negative recipe ingredients. Voxels passing the P <0.05 FWE cluster corrected threshold for greater connectivity with middle temporal cortex (MTC) seed are shown. Cluster extent threshold is given by k.

These data provide evidence that vmPFC and MTC connectivity changed depending on the value of recipe ingredients on the first presentation. To examine whether this effect might depend on successful retrieval of relevant schema-value information, we tested the interaction of value and accuracy on the first presentation of recipe ingredients. This interaction was significant when either vmPFC and MTC was used as a seed (Bonferroni corrected t-tests: t's(20) ≥ 2.95, P < 0.05, d ≥ 0.64), but not between these seeds and any other ROIs. We further examined whether this value related increase in vmPFC-MTC connectivity was higher on the first presentation than subsequent presentations by testing

521 the interaction of presentation number and value for recipe ingredients, but this interaction
522 did not reach significance after correction for multiple comparisons in either seed.

523 We also tested whether connectivity between MTC and vmPFC and any other ROIs
524 depended on clustering status by contrasting the recipe and independent conditions on the
525 first presentation, collapsing across value. There was no significant change in connectivity
526 between MTC, vmPFC and any other ROIs in this contrast. We also contrasted connectivity
527 for correct and erroneous responses for recipe ingredients on the first presentation and
528 similarly did not any difference in connectivity between ROIs. Exploratory seed-to-voxel
529 analyses for both of these contrasts were similarly null.

530 Next, we investigated if connectivity between with the HPC and striatum depended
531 on demands on episodic retrieval in the second and third presentations. Contrasting positive
532 and negative independent ingredients in these trials, we did not find any significant increase
533 in connectivity between HPC and striatum with each other or any other region in ROI-to-ROI,
534 or seed-to-voxel analyses. We further examined if connectivity with these ROIs was higher
535 for positive than negative ingredients in the independent condition relative to the recipe
536 condition by testing the interaction of value and clustering. This analysis was similarly null
537 for both seeds.

538 Finally, we tested if HPC and striatum connectivity was influenced by greater demand
539 on episodic memory. Collapsing across value conditions, we contrasted independent and
540 recipe ingredients on the second and third presentations where ingredient values could be
541 retrieved from episodic memory, with the expectation that demand on episodic memory
542 would be greater in the independent condition. There was no significant difference in
543 connectivity between the striatum or HPC with any other ROIs in this contrast. However, a

seed-to-voxel whole-brain analysis with the HPC seed showed significantly greater connectivity with a cluster in the right putamen and insula (Figure 9a; MNI: 26, 10, 0). The reverse seed-to-voxel analysis with the striatum ROI as a seed did not reveal significant clusters within a small-volume corrected HPC ROI.

To test if this increase in HPC and hippocampus connectivity was related to retrieval success we also contrasted correct and error responses on the second and third presentation of independent ingredients. Using the HPC as a seed in a seed-to-voxel analysis, we found increased connectivity with this ROI and the right putamen (MNI: 28, 2, 4) after small volume correction within the striatum ROI. An exploratory whole-brain analysis also found that HPC significantly increased connectivity with clusters in bilateral anterior rostrolateral PFC and right insula (Figure 9b). However, there were no significant differences in connectivity between these conditions in ROI-to-ROI contrasts with either HPC or striatum as seeds. The reverse ROI-to-voxel analysis with the striatum ROI as a seed did not reveal significant increase in connectivity in a small-volume corrected HPC ROI. All peaks from seed-to-voxel analyses are reported in Supplementary Table 5.

In summary, we found that vmPFC-MTC connectivity depended on recipe value associations on the first presentation where this information was key to guiding choice. Meanwhile, HPC appeared to increase connectivity with the right putamen in response to greater demand on episodic memory compared to conditions where schema memory could still be used, and in trials where information was successfully retrieved.

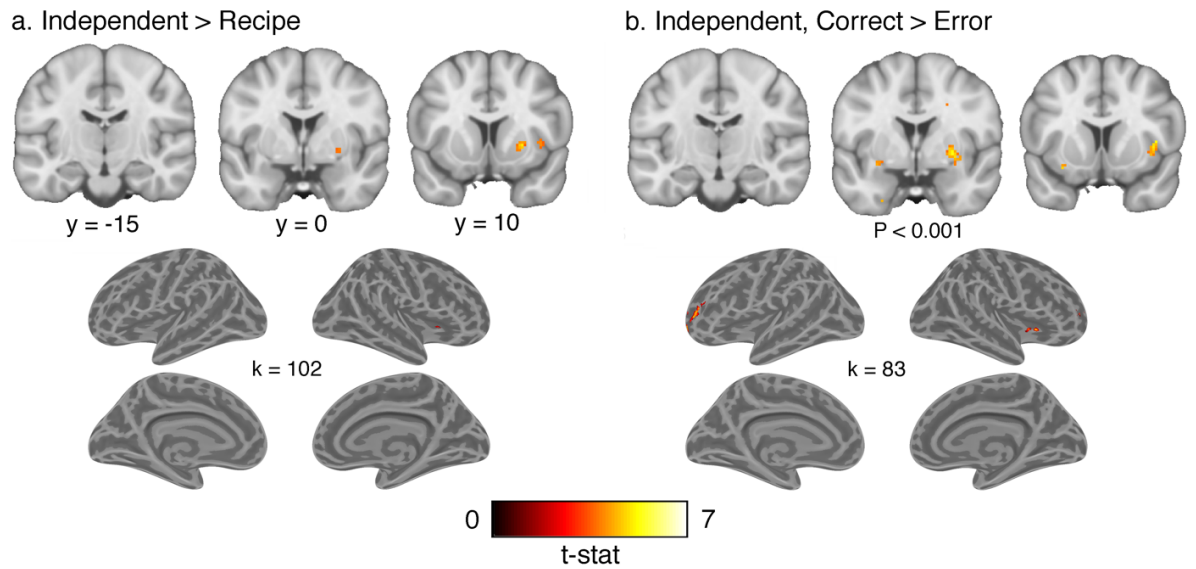


Figure 9. Whole brain t-statistic maps for generalized psychophysical interaction analyses showing voxels with increased connectivity with the hippocampus (HPC) during presentations two and three. **a.** Greater HPC connectivity for independent ingredients over recipe ingredients at the P <0.05 FWE cluster corrected threshold. **b.** Greater connectivity on correct than erroneous responses. Voxels in coronal slices are thresholded at P < 0.001, uncorrected. Projections onto cortical surface are thresholded at the P <0.05, cluster corrected threshold. Cluster extend threshold is given by k.

Discussion

Here we evaluated the neural circuitry involved in value judgments based on multiple memory sources. Functional neuroimaging results implicate the vmPFC in judging options based on both schema-level knowledge and episodic, cached memory, while striatal activity correlated with participants willingness to bet on ingredients modulated by the unsigned strength of their episodic memory for those ingredients' values. Thus, these regions have distinct roles in memory-based value judgment and decision-making.

Participants clearly used schematic knowledge about recipes to complete this task. They were able to infer the value of ingredients based on schematic knowledge without feedback, and continued showing a benefit from recipe information in subsequent presentations. Our computational model was able to capture this behavior using normative

information about ingredient-recipe associations. Modeling also revealed interesting variation in how participants used these associations. Notably, freely fitting the number of dimensions used in reconstructing these recipe-ingredient associations substantially improved the fit of the model, implying that participants varied in the depth of semantic knowledge they employed in this task. Participants also differed in how they strategically evaluated ingredients based on the similarity of an ingredient's associations with both positive and negative recipes. Some participants attributed associative strength to one recipe or the other, while others treated these ingredients as having little net schema value. The positive recipe also generally had a larger influence on behavior than the negative recipe, indicating that participants prioritized searching associations with this recipe first — possibly because they adopted a default response to pass on ingredients where they could not retrieve information that would confer positive value.

Notably, memory strength, expressed as unfamiliarity, had a substantial impact on participants' willingness to endorse ingredients. Participants demonstrated a general tendency against betting on independent ingredients on the first presentation. Given that there was an equal proportion of positive and negative value ingredients, and there was a 2:1 ratio of points gained to lost on a bet, this behavior reflects a deviation from a rational maximization of expected value, or loss aversion (Kahneman & Tversky, 1979). Whereas uncertainty is frequently parameterized as descriptive or experienced probabilities in behavioral economics, here uncertainty depended on the strength of participants' episodic memory. These results are also consistent with previous work that has suggested that memory of item identity can make an option a more appealing candidate option (Gluth, Sommer, Rieskamp, & Buchel, 2015). In our behavioral model, unfamiliarity modulated

602 option values, as items with weaker memory strength were riskier bets. In real terms, people
603 appear to apply a meta-cognitive strategy where the strength of evidence from memory
604 affects option value, regardless of its content. We would thus predict that a strong memory
605 for a weakly positive or negative value would be more likely to drive a response than a weak
606 memory for a strongly positive or negative value.

607 While we predicted that vmPFC would be preferentially engaged by value judgments
608 based on schema-level knowledge, we found that this region tracked both episodic and
609 schematic dimensions of option values, as revealed in model-based and model-free analyses.
610 Past studies have demonstrated that the value function represented in vmPFC can flexibly
611 bend to accommodate attentional priority toward different option attributes (Hare,
612 Malmaud, & Rangel, 2011; Hunt, Dolan, & Behrens, 2014; Leong, Radulescu, Daniel,
613 DeWoskin, & Niv, 2017; Rudorf & Hare, 2014). Our data extend these findings by
614 demonstrating that activity within this region accommodates value information retrieved
615 from different types of memory stores within the same task.

616 The ubiquity of vmPFC value responses to different stimuli and tasks has been taken
617 as evidence for this region encoding a common currency value representation (Chib, Rangel,
618 Shimojo, & O'Doherty, 2009; Levy & Glimcher, 2012). However, lesion evidence in humans
619 has implied that this region is not necessary for value judgment *per se* (Reber et al., 2017;
620 Vaidya & Fellows, 2015; Yu, Dana, & Kable, 2018), and may have a more nuanced role in some
621 component of the valuation process (Vaidya, Sefranek, & Fellows, 2017; Xia, Stolle, Gidengil,
622 & Fellows, 2015). Our initial prediction about the role of vmPFC was informed by this work
623 and evidence that vmPFC or OFC damage and inactivation in animal models does not affect
624 decision-making or value judgment in situations where value information can be retrieved

625 from prior direct experience, but does when this information must be inferred from
626 structured knowledge about a task (Bradfield et al., 2015; Izquierdo et al., 2004; Jones et al.,
627 2012; Rudebeck et al., 2013).

628 There may be a few factors contributing to this discrepancy. While 'cached' values
629 were learned over many repeated exposures in these lesion experiments, here cached values
630 depended on just one or two trials. Encoding values through limited, spaced experience may
631 depend on more flexible systems available in the PFC, while values learned through lots of
632 continuous experience might become less dependent on these systems and rely instead on
633 the basal ganglia (Balleine & O'Doherty, 2010; Graybiel, 2008). Recent work has shown that
634 spaced, brief training over a period of days in a reinforcement learning task increases fMRI
635 value signals in vmPFC compared to an equivalent number of trials within a single session
636 (Wimmer, Li, Gorgolewski, & Poldrack, 2018). These findings imply that this region may be
637 engaged more when value judgments rely on episodic memory rather than massed
638 experience in a single session.

639 Importantly, our data cannot speak to the necessity of vmPFC in decisions based on
640 either episodic or schema value. fMRI can only provide evidence of the association of vmPFC
641 in this process without making causal claims about its specific contribution. Although these
642 data suggest the involvement of vmPFC in representing both episodic or schema-based
643 value, this information may be robustly represented by other regions that do not depend on
644 vmPFC to influence behavior. Studying the effects of vmPFC damage in this task could help
645 assess when this region is critical for memory-based value judgment.

646 Prior studies have also shown that value judgments based on semantic information
647 are associated with broader network interactions between vmPFC and regions of left middle

648 and superior temporal cortex, in fMRI functional connectivity (Lim, O'Doherty, & Rangel,
649 2013). Here, value information was associated with greater vmPFC connectivity with MTC
650 on the first presentation, and this connectivity was stronger for correct than error trials.
651 These results are consistent with our hypothesis that the interaction between these regions
652 has an important role in the judgment of option values based on their schema associations,
653 as schemas are particularly important on the first presentation, when no episodic
654 information is available.

655 This value-based difference in vmPFC-MTC connectivity could reflect the
656 accumulation of positive or negative value information through the interaction between
657 these regions. However it may also reflect the differential weight given to positive and
658 negative recipes, as there was clearer behavioral evidence that participants utilized positive
659 recipe associations on this first presentation. In either case, these data provide evidence that
660 engagement between vmPFC and semantic memory stores is involved in inferential value
661 judgment based on real-world structured knowledge. Connectivity between these regions
662 did not differ by clustering status (i.e. there was no difference in MTC-vmPFC connectivity
663 for recipe and independent ingredients on the first presentation). However, this contrast is
664 complicated by the likely possibility that both these conditions involve attempts at memory
665 retrieval, while retrieval success is more likely in the recipe condition.

666 Our prior work has shown that aVLPFC is engaged by controlled retrieval of both
667 episodic and semantic information via its connections with middle temporal cortex and the
668 hippocampal formation (Badre, Poldrack, Pare-Blagoev, Insler, & Wagner, 2005a; Badre &
669 Wagner, 2007; Barredo et al., 2015; Barredo et al., 2016). Here, activity in this region was
670 principally modulated by presentation number, not by clustering or value, and connectivity

between this region with MTC and HPC did not change based on memory demands. aVLPFC shares connections with vmPFC and MTC (Saleem et al., 2008), which may be key for controlled retrieval in memory-based value judgment. While we did not manipulate demands on controlled retrieval here, we might expect that connectivity between these three regions to increase when value judgments require retrieving weak semantic associates.

This experiment explicitly crossed two features commonly associated with the vmPFC, but rarely compared within the same task: value coding and schema-based memory and inference (Chan et al., 2016; Delgado et al., 2016; Ghosh et al., 2014; Moscovitch & Melo, 1997). While there was no main effect of ingredient clustering in an ROI-based analysis of vmPFC, there was an overall effect of the recipe clustering status of ingredients in an overlapping region of medial OFC in a whole-brain analysis. The vmPFC ROI used in this study was chosen from a meta-analysis of expected value effects in neuroimaging studies rather than schema memory (Bartra et al., 2013), which may have influenced this difference between the ROI and whole-brain results. Although there was substantial overlap between the ROI and voxels responding to clustering in the whole-brain analysis, the effect may have been washed out by averaging in the ROI volume. Our data are consistent both with schema-related knowledge playing a role in value construction in vmPFC, and with this region responding to schema-related knowledge. Recent rodent electrophysiological data also indicate that value and schematic representations of task co-exist in OFC (Zhou et al., 2019), indicating that this region may represent value and schematic knowledge with separable codes.

The role of striatum in memory tasks has been the focus of recent research, with several hypotheses being considered (Elward et al., 2015; King, Chastelaine, Elward, Wang,

694 & Rugg, 2018; Schwarze, Bingel, Badre, & Sommer, 2013; Scimeca & Badre, 2012). Here,
695 activity in the striatum appeared to integrate the influence of episodic strength on
696 participants' decisions to bet or pass on ingredients, as its activity correlated with the
697 interaction of participant-P(yes) and unfamiliarity. Relatedly, recent work from our group
698 has observed that the striatum tracked prediction errors for retrieval success during a
699 recognition memory task (Scimeca et al., 2016), analogous to its role in action selection and
700 reinforcement-learning (Garrison, Erdeniz, & Done, 2013; Haber & Knutson, 2010;
701 Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006). These prediction errors were
702 correlated with changes in the criterion by which evidence from memory was evaluated.
703 However, unlike more conventional reinforcement learning settings, in this memory task,
704 the striatum was sensitive to prediction errors that derived from memory strength rather
705 than the likely outcome of a response, positive or negative. We expand on those findings
706 here, demonstrating that the striatum responds to the influence of memory strength on
707 participants' willingness to make a bet in a value-based decision-making task, independent
708 of memory valence. Together, these results imply a critical role for the striatum in a cognitive
709 decision about when to act based on memory strength, equivalent to its role in controlling a
710 tendency to bet in risky decision-making (Knutson, Wimmer, Kuhnen, & Winkielman, 2008;
711 Nachev et al., 2015; Zalocusky et al., 2016), or in gating working memory during cognitive
712 control (Chatham & Badre, 2015).

713 In several cases, the striatum appeared to principally respond to unsigned
714 information about memory strength in this task (e.g. memory strength and surprise), rather
715 than the value of options themselves (i.e. episodic and schema value or RPE), as it has in
716 many other decision-making and learning experiments (Bartra et al., 2013; Garrison et al.,

717 2013). While the striatum has also been shown to respond to successful memory retrieval in
718 many studies (Spaniol et al., 2009), other work has demonstrated that these signals
719 dynamically depend on the current task, changing with participants' goals (Han et al., 2010),
720 or has indicated that striatal reward and retrieval signals are distinct (Elward et al., 2015).
721 Here we found evidence that unsigned episodic memory strength modulates striatal value
722 signals, but the signed episodic value of options did not. These data indicate that value and
723 retrieval signals in the striatum may follow a common dimension that corresponds to the
724 utility of retrieved information based on task demands — possibly biasing cortical
725 representations of associated actions based on the strength of information recovered from
726 retrieval.

727 We also observed that connectivity between HPC and the striatum increased in the
728 second and third presentations of independent ingredients compared to recipe ingredients,
729 indicating decisions with greater demand on episodic memory for value judgment increased
730 communication between these regions. This connectivity was also higher on correct choices
731 for independent ingredients than on error trials, suggesting that HPC-striatal
732 communication was involved in the successful retrieval of ingredient-value associations.
733 However, connectivity between these regions did not depend on the episodic value of these
734 ingredients. Instead, this increase in connectivity was observed while collapsing across value
735 conditions, arguing against this change in connectivity being specific to conferring positive
736 value to previously rewarded ingredients.

737 This increase in connectivity between the HPC and striatum was only evident in the
738 right putamen in a whole-brain analysis when the HPC was used as a seed ROI, not in ROI-
739 to-ROI analyses or with the striatal ROI as a seed. The apparent selectivity of this effect across

740 analyses may owe to our striatal ROI averaging over functionally distinct subregions. This
741 ROI was large *a priori*, experiments that have examined reward and retrieval success signals
742 together have found somewhat variable peak coordinates (Elward et al., 2015; Han et al.,
743 2010). At a liberal threshold (P <0.01, uncorrected), HPC connectivity was also increased in
744 the left putamen, but not the caudate.

745 In conclusion, we studied how people integrate evidence from episodic and semantic
746 memory during value based decision making. We observed that vmPFC is involved in judging
747 the value of options based on information retrieved from both schema-based inference and
748 episodic memory retrieval, while the striatum was implicated in controlling willingness to
749 take action based on the strength of episodic memory. These functions are supported
750 through interactions with systems implicated in the storage of both episodic and semantic
751 memory. Our data provide fresh perspectives into the dynamics of this circuitry, and a view
752 into the intersection of decision-making and memory that is the seat of much real-world
753 choice behavior. These results thus have clear implications for understanding the origins of
754 decision-making impairments in neurodegenerative diseases that affect semantic and
755 episodic memory such as Alzheimer's disease and frontotemporal dementia (Salmon &
756 Bondi, 2009), as well as understanding how disruption of dopaminergic circuits in
757 Parkinson's disease may affect memory (Edelstyn, Shepherd, Mayes, Sherman, & Ellis, 2010).
758 New insights into the interaction of these circuits may have implications for clinical
759 interventions or treatments for these disease symptoms.

760 **Materials and Methods**

761 *Participants*

762 Twenty-six participants were recruited from the Providence, RI area to participate in
763 the fMRI experiment. Two participants were excluded due to excessive motion during the
764 functional scan, and one was excluded for falling asleep during the experiment. Of the
765 remaining 23 participants, four were male, their mean age was 22.5 years (SD = 4.0), and
766 their mean years of education was 15.5 (SD = 2.3). All participants were right-handed, had
767 normal or corrected-to-normal vision, and were screened for the presence of psychiatric or
768 neurological conditions, psychoactive medication use, and contraindications for MRI. Two
769 participants rejected all ingredients on the first presentation of the positive independent
770 condition, prohibiting analysis of their behavior or ROI-level responses in a full 12-condition
771 ANOVA after filtering for correct responses. One participant did not commit any errors on
772 the second or third presentation of the positive recipe ingredients, prohibiting analysis of
773 their error trials in this condition.

774 Two hundred participants from the United States were recruited through Amazon
775 Mechanical Turk (AMT) to provide normative naming responses for images of food
776 ingredients. Four of these subjects were rejected for exceptionally low accuracy in this task
777 (> 50%). The mean age of the remaining participants was 35.4 years (SD = 10.5), and 107 of
778 them were male.

779 A second sample of 221 participants were recruited from the United States through
780 AMT to provide normative data for ingredient-recipe associations. Twenty-nine of these
781 participants were rejected for failing to follow task instructions (most often by simply
782 writing the names of the ingredients rather than associated recipes). The remaining
783 participants had a mean age of 36.0 years (SD = 11.1), and 71 of them were male.

All participants across experimental and norming studies gave informed consent as approved by the Human Research Protections Office at Brown University, and were compensated for their participation. Participants in the fMRI experiment also received a monetary bonus proportional to the points they earned in the main task.

Stimuli

Ingredient stimuli were obtained from the bank of standardized stimuli (BOSS; N = 65) (Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010; Brodeur, Guerard, & Bouras, 2014), food-pics database (N = 50) (Blechert, Meule, Busch, & Ohla, 2014), with a small set of supplemental food images collected by A.V. from Google Images (N = 4). These images were selected to include basic food ingredients — i.e. items that might be used as parts of a recipe, rather than completed meals. For example, an image of cheese would be included rather than an image of a pizza. All images were scaled to the same size and padded with white pixels or cropped to fit the same area. Ingredients from the BOSS database were excluded if more than 20% of the normative sample either responded that they did not know the name of the ingredient or did not recognize the ingredient. Ingredients were also rejected if the modal name for the ingredient given by respondents was incorrect (indicating low recognition accuracy), and if the H-value (a measure of naming disagreement across participants, similar to entropy) was above 3.00.

Because similar naming data were not available for ingredients from food-pics or Google Images, normative naming data were obtained through an AMT experiment. Participants were shown ingredient images and asked to write the name of the ingredient in a text box below, or write “I don’t know.” Responses that included common variants of ingredient names were considered correct, while overly general names were marked as

incorrect (e.g. 'nuts' for an image of pecans). Ingredients where more than 20% of participants did not know the name, or gave incorrect answers, were rejected. Naming agreement was also examined by calculating an H-value for each ingredient, however this value did not exceed 3.00 for any ingredient in this set.

Normative recipe associations

Associations between ingredients and recipes were obtained through a norming experiment conducted on AMT akin to single word generation paradigms used to derive word associations (Postman & Keppel, 1970). Participants were shown ingredient images one at a time and were asked to write the name of the first recipe that came to mind using the ingredient shown on screen. Unique recipe names were then extracted for each ingredient and two researchers (A.V. and H.J.) manually processed these data. Alternative names, semantically similar responses and misspellings were all collapsed into single recipes, and ingredient specific information was removed to facilitate detection of overlap in ingredient-recipe associations (e.g. responses like 'pumpkin soup' and 'mushroom soup' were both reclassified as 'soup'). The end result of this process was a matrix with endorsement counts for 102 ingredients in 1876 recipes. Six ingredients were then dropped from this set, as they were visually too similar to other ingredients in the experiment, leaving 96 ingredients total.

Of these recipes, many (915) were only named by a single AMT participant for a single ingredient. As our task design required that each recipe have a large number of associate ingredients (at least seven recipe ingredients, and some additional ingredients to be used as independent items), we cycled through different parameters for filtering out recipes. With each of these filtered lists, ingredients were assigned to hierarchically organized clusters

using the Matlab (Mathworks, Natick, MA) function 'linkage,' which sorted ingredients based on the mean cosine similarity of their naming frequencies. Through this process we selected six recipes where the clustering solution provided ingredient cluster sizes that were large enough for our task design (pasta (N=10), soup (N=13), cake (N=15), pie (N=21), salad (N=17) and sandwich (N=20)).

fMRI Experimental design

The main experimental task, and all other tasks not administered through AMT were programmed in Psychtoolbox-3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007). At the start of each block, participants were told that they were about to feed a new customer, and were given the names of the recipes that the customer liked and disliked. The two recipes in each block were randomly selected without replacement from a set of six recipes.

The screen was then left blank for 12 s until the start of the first trial. Trials were preceded by a jittered inter-trial interval (ITI) drawn from a lognormal distribution (0.5-6 s). In each trial, participants saw a food ingredient centrally displayed for up to 2.7 s, during which time they could choose whether or not to feed the ingredient to the customer ('yes'/'no'). The names of the recipes that each customer liked and disliked appeared at the top of the screen on each trial, color coded to indicate the customer's preference.

Participants indicated both their choice and their confidence using a combined response on a sliding bar. The slider would start in a random position on each trial and participants could then drag it toward the left or the right by holding down a button on a fiber optic response pad (Current Designs, Philadelphia, PA) using their right index or middle fingers. The slider would stop moving when it reached an end of the bar, or when participants let the button go. Participants could not change the slider position after letting

go of the button. The side of bar where participants placed the slider signaled their choice, and the distance of the from the middle of the bar indicated their confidence. After a response, there was a jittered inter-stimulus interval (ISI) drawn from a uniform distribution (0.5-5.5 s). The remaining time from the response window and slider movement period was also added to this interval. Participants then saw feedback about their decision for 1 s. A square above the correct answer would light up, and the amount of points added or subtracted from participants' total would be displayed in this period.

Before the fMRI experiment, participants completed a practice task in which they learned to use the slider while verifying various general knowledge statements (e.g., 'Black bears are omnivores'), and indicate their confidence in their response. Participants received further practice outside the scanner by completing a short practice block of the restaurant task where they fed a customer with recipes and ingredients that did not appear in the fMRI experiment.

Trial-order, ITI and ISI lengths were drawn from sequences optimized for efficient estimation of the hemodynamic response across conditions of interest. These sequences were optimized by generating 5000 random sequences, and using SPM12 to create event-based design matrices for stimuli and feedback events in each sequence. The average variance for each regressor explained by all other regressors in the matrix was calculated for each sequence and multiplied by the total duration of the sequence to approximate design efficiency. The three sequences with the lowest values in this measure (i.e. least covariance in regressors with the shortest experiment time) were used for the fMRI experiment and were randomly ordered for each subject.

Value rating task

At the start of the experimental session, participants completed a short rating task where they were asked to provide ratings for how much they would hypothetically want to have each ingredient at the end of the session. Participants provided responses on a seven point Likert scale from -3 to 3, with labels on -3 reading 'Not at all,' 0 reading 'Indifferent' and 3 reading 'Very much.' These subjective value ratings were included as a nuisance variable in each of the parametric GLM analyses. This task also served to visually familiarize participants with all ingredient images.

Computational modelling of behavior

Computational models of behavior were used to characterize the dynamics underlying participants' behavior in this task. In total, 16 alternative models were tested that included combinations of seven main variants that each built off of a base model by adding a new parameter. To briefly summarize, alternative models included different operationalizations of recipe associations, separate parameters for weighing positive and negative recipe associations, trade-offs in the use of episodic and schema value information and a model with a separate recollection process in addition to a graded episodic memory. The details of each of these variants are described in the Supplementary Methods. Here we will summarize the features of the best fitting model that combines some of these features and is the focus of model-based fMRI analyses.

Within the best fitting model, the value (*V)* of the ingredient (*i)* on each trial was calculated as:

$$V_i = \beta_0 + \beta_u(1 - |EV_i|) + \beta_e EV_i + \beta_s SV_i \text{ (Eq. 1)}$$

Where β_0 is an intercept term that captures overall, item-independent biases toward positive or negative value, $\beta_u$is a parameter that controls the influence of unfamiliarity, β_e

controls the influence of episodically encoded values (EV), and β_s controls the influence of schema values (SV) for the current trial ingredient. The model's willingness to bet on each trial (i.e. the model-P(yes)) was then calculated by placing V in a simple logistic function.

On each trial where participants responded, the episodic value for each ingredient was updated according to a simple Rescorla-Wagner learning rule (Rescorla & Wagner, 1972):

$$EV_i = EV_i + \alpha(outcome - EV_i) \text{ (Eq. 2)}$$

Where the value for *outcome* could be 1 or -1, depending on whether the feedback was positive or negative. While this rule is more commonly used in the context of reinforcement learning, it was adopted here as a simple means of operationalizing a graded form of episodic memory that is signed and becomes stronger with repetitions of consistent value associations. The parameter α was a fixed learning rate that was fit for each subject in an alternate version of the base model where β_e was held at different constant values (1, 2, 5 and 8). The model fit with a constant value of 5 for $\beta_e$had the lowest average negative log-likelihood of constants tested, and so α values from this fit were used in the base model and all further variations. This procedure to separately fit α was followed because β_e and α could not be stably estimated within the same model. Estimating these parameters separately allowed us to roughly capture individual differences in the encoding of EV, while separately estimating the contribution of EV to choice behavior through β_e. Allowing α to vary was also found to reduce the negative log-likelihood of the base model over a variant where α was fixed to 1.0.

To reflect forgetting, the values for all ingredients also decayed towards zero with each trial according to:

$$EV_{t+1} = EV_t - dEV_t \text{ (Eq. 3)}$$

Where d is a parameter that controls the rate of this decay.

Schema values were calculated through a weighted difference of the associative strength of ingredients with positive and negative recipes. The associative strength of each ingredient with these recipes was determined through a procedure similar to latent semantic analysis (LSA) (Dumais, 2004; Landauer et al., 1998). The endorsement frequencies for ingredients across all recipes in the normative AMT data were weighted by their entropy across recipes and a singular value decomposition (SVD) was carried out on this matrix. This matrix was then reconstituted with a smaller number of dimensions set by a free parameter c (see Supplementary Figure 1 for reconstituted matrices at different values of c). The association of an ingredient with each recipe was calculated as the mean cosine similarity between that ingredient and all ingredients in each recipe cluster within this reconstituted matrix.

The schema similarity model estimated participants' uncertainty about the recipe associations of each ingredient as the probability that an ingredient was associated with the same recipe ($P(SR)$).

$$P(SR)_i = \frac{1}{1+\exp\left(-\beta_{sim}\left(1-\frac{|PR_i - NR_i|}{RRD}\right)\right)} \text{ (Eq. 4)}$$

Where RRD corresponds to the range of recipe association differences (i.e. the difference between the maximum and minimum recipe associative strength for all ingredients and recipes). This probability was then used in calculating weights for the sum and difference of the positive and negative recipe associations, where the weight for the sum of these associations was calculated as

$$w_{same} = 2(\,P(SR)_i - 0.5) \text{ (Eq. 5)}$$

And the weight for the difference of these ingredient associations was calculated as

$$w_{diff} = 1 - 2\,|P(SR)_i - 0.5| \text{ (Eq. 6)}$$

SV was then calculated as the weighted combination of the difference and sum of each ingredient's associative strength with the positive and negative recipes:

$$SV_i = w_{same}\,(PR_i + NR_i) + w_{diff}\,(PR_i - NR_i) \text{ (Eq. 7)}$$

The parameter β_{sim} controlled the sensitivity of participants to the similarity of ingredients' recipe associations, and the relative contribution of the difference or sum of positive and recipe associations to SV. This parameter also controlled the valence of the additive strength of these associations: positive or negative values for β_{sim} would give the same valence to the additive strength of recipe associations, respectively. Alternatively, if the β_{sim} parameter was zero, participants would rely solely on the difference of the ingredient's associative strength with the positive and negative recipes.

Participants' responses consisted of two main features: a binary response indicating whether or not to feed an ingredient to a customer, and a continuous rating of their confidence in this binary response. This continuous confidence response could also be transformed into a continuous measure of participants' belief that a customer would like an ingredient (i.e. P(yes)) by measuring the position of the slider between the 'yes' and 'no' options rather than the distance of the slider from the middle, similar to (Scimeca et al., 2016). Models were fit to both participants' continuous and binary choice responses. The overall model likelihood was calculated as the product of the model likelihood for participants' binary responses, and the model likelihood of the linear fit between the logit transformed model-P(yes) and participant-P(yes). The negative log of this likelihood

function served as the objective function to be minimized in the fitting procedure carried out by Matlab function 'fmincon.'

Computational models were fit to participants' behavior in a step-wise process. At each step, a simpler preferred model was modified by the addition of new free parameters and the Akaike Information Criterion (AIC) of the simpler model and modified models were compared with a one-tailed t-test, as the addition of these new parameters was predicted to reduce AIC. The model with the best improvement in model fitness across subjects (i.e. the largest t-statistic in this comparison) was then considered the preferred model in the next step. Additional parameters that significantly reduced AIC were added to the preferred model and compared in the next step ($P < 0.05$). The search for the best fitting model was stopped once additional parameters no longer significantly improved model fitness, or when the parameters to search across were exhausted.

Each model was fit separately for each subject 20 times with randomized starting positions. The set of parameters with the lowest negative log-likelihood from each set of fits was saved and the other parameter fits were discarded. These 20 separate fits helped guard against the influence of idiosyncratic differences in starting point on the recovered parameters, and the possibility that the model would converge on local minima in the likelihood function.

To verify that parameters of the best fitting model could be recovered, we generated 100 simulated datasets with randomized parameters. The parameter d was generated from a uniform random distribution from 0 to 0.05. Parameters β_e and $\beta_s$were generated from a lognormal distribution with a mean of 5 and standard deviation of 0.5. Parameters β_0, $\beta_m$and β_{sim}were all generated from a standard normal distribution (mean 0, standard deviation of

990 1). The *c* parameter was generated from a random normal distribution ranging from 2 to 58.
991 Given that these parameters were randomly generated and did not share the covariance of
992 the real parameters fit to participants' behavior, this test could overestimate their
993 recoverability. Thus, we further verified the recoverability of these parameters by fitting the
994 model to data simulated using the parameters fit to participants' real behavior.

995 *fMRI data acquisition and preprocessing*

996 Whole-brain imaging was acquired using a Siemens 3T Magnetom Prisma system
997 with a 64-channel head coil. First, a high resolution T1 weighted MPRAGE image was
998 acquired for visualization (repetition time (TR), 1900 ms; echo time (TE), 3.02 ms; flip angle,
999 9°; 160 sagittal slices; 1 x 1 x 1 mm voxels). Functional volumes were acquired using a
1000 gradient-echo echo planar sequence (TR, 2000 ms; TE, 28 ms; flip angle 90°; 38 interleaved
1001 axial slices tilted approximately 15° from the AC-PC plane; 3 x 3 x 3 mm voxels). Functional
1002 data were acquired over three runs. Each run lasted 16.5 min on average (Mean: 494.8
1003 acquisitions). After the functional runs, a brief in-plane anatomical T1 image was collected,
1004 which was used to define a brain mask that respected the contours of the brain and the space
1005 of the functional runs (TR, 350 ms; TE 2.5 ms; flip angle 70°; 38 axial slices; 1.5 x 1.5 x 3 mm
1006 voxels). Soft padding was used to restrict head motion throughout the experiment. Stimuli
1007 were presented on a 32-inch monitor at the back of the bore of the magnet, and subjects
1008 viewed the screen through a mirror attached to the head coil.

1009 Functional data were preprocessed using SPM12. Quality assurance for the functional
1010 data of each subject was first assessed through visual inspection and TSdiffAna
1011 (https://sourceforge.net/projects/spmtools/) and ArtRepair
1012 (http://cibsr.stanford.edu/tools/human-brain-project/artrepair-software.html). Outlier

volumes (defined by standard deviations from the global average signal) were interpolated Slice timing correction was then carried by resample all slices to match the first slice. Next, motion during functional runs was corrected by registering volumes to the mean using rigid-body transformation. In any case where movement across runs was in excess of the voxel size in any direction (3 mm), realignment was recalculated run-by-run. Any participants where motion within a run was larger than a voxel were excluded from further analysis (N = 2). Functional volumes were then normalized to Montreal Neurological Institute (MNI) stereotaxic space using 4th order B-spline interpolation, and resampled to 2 x 2 x 2 mm voxels using trilinear interpolation. Finally, volumes were spatially smoothed using an 8 mm full-width half-maximum isotropic Gaussian kernel. The in-plane anatomical T1 image for each subject was also normalized to MNI space and a brain mask for functional analysis was created using the Brain Extraction Tool in FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/).

fMRI data analysis

Functional data were analyzed under the assumptions of the GLM using SPM12. All regressors and parametric modulators were convolved with the SPM canonical hemodynamic response function (HRF). Subject motion (six translational and rotational components) were also included as nuisance regressors. Beta coefficients for single subject effects were estimated using a fixed-effects model in a first-level analysis. For whole-brain contrasts, these estimates were then included in a second-level analysis using a one-sample t-test against 0 at each voxel. Whole brain t-statistic maps were thresholded at a cluster defining threshold of $P < 0.001$ uncorrected, and then evaluated for statistical significance at the cluster corrected FWE threshold of $P < 0.05$ to control for multiple comparisons. The cluster extent for this threshold varied by test and is reported with each analysis.

1036 *fMRI general linear models*

1037 Five different GLMs were fit to fMRI data:

1038 (1) Condition-based GLM

1039 First, a condition-based GLM was fit to the functional data. This GLM included
1040 separate regressors at the response and feedback phases for the following factors: clustering
1041 (recipe, independent), value (positive, negative), presentation number (1, 2, 3) and accuracy
1042 (correct, error). The duration of the boxcar for the response regressors was set to the median
1043 of subjects' combined RT and slider movement durations. Trial-by-trial deviation from this
1044 median duration was included as an additional nuisance regressor. This GLM also included
1045 a nuisance regressor for the time the stimulus display was on screen (i.e. the response
1046 window, ISI and feedback).

1047 (2) Finite impulse response GLM

1048 For ROI-based analyses on contrasts of task conditions, a separate finite impulse
1049 response (FIR) GLM was fit to each participant. This model had the identical conditions as
1050 the condition-based GLM described immediately above, but used an FIR model with eight
1051 time points (16 s window) as a basis function for the hemodynamic response function.

1052 (3) Participant-behavior GLM

1053 Participants' responses using the slider were used to estimate participants' beliefs
1054 during value judgment and processing of feedback. We tested for correlates of these putative
1055 processes with a parametric GLM. This model had three boxcar regressors: a response
1056 regressor with onsets at stimulus presentation and durations lasting to the point
1057 participants finished moving the slider, a feedback regressor with onsets and durations
1058 lasting the presentation of feedback, and a nuisance regressor for the stimulus display. Three

parametric modulators of the response regressor were included in this model: participants' reported P(yes), reported confidence and subjective value ratings for each ingredient from the original value rating task. The feedback regressor also had two parametric modulators: RPEs (i.e. the difference between participant-P(yes) and outcomes) and surprise (unsigned RPEs). Parametric modulators were not orthogonalized in this model, or in any of the following GLMs.

(4) Model-behavior GLM

A parametric GLM based on the parameters estimated from the computational model fits to participant behavior was estimated to compare with the parametric GLM based on participants' raw behavior. This GLM included the same three regressors as the first GLM, but with parametric modulators derived from the computational model's estimates of participants' behavior. There were five modulators on the response regressor: model-P(yes), model confidence, residuals for model-P(yes), residuals for model confidence and subjective value ratings. There were also four modulators on the feedback regressor: model RPE, model surprise and residuals on model RPE and surprise. All residuals were calculated by simply taking the difference of participants' behavior and model-derived estimates.

(5) Model-based value components GLM

The final GLM included parametric modulators based on the model-derived value signals, and their interaction with participants' responses. This GLM had the same three main regressors as GLMs 3 and 4. There were eight parametric modulators on the response regressor: participant-P(yes), participant confidence, episodic value, schema value, unfamiliarity, interactions between participant-P(yes) and episodic value, schema value and

1081 unfamiliarity, as well as subjective value ratings for ingredients. The feedback regressor had
1082 two parametric modulators: participants' RPEs and surprise.

1083 *Region of interest analyses*

1084 Regions of interest (ROIs) were defined for testing region-specific hypotheses using
1085 unbiased criteria. These ROIs were either defined using masks from the automated
1086 anatomical labelling atlas (AAL), or using 10 mm spheres around peak coordinates from
1087 independent studies. These spheres were then masked to only include voxels that were more
1088 than 50% likely to be gray matter in SPM's tissue probability map. The coordinates and
1089 references for these ROIs are listed in Table 1.

1090 In ROI analyses contrasting conditions, the time course of activity within each ROI
1091 was estimated using the MarsBar Toolbox (Brett, Anton, Valabregue, & Poline, 2002) from
1092 the FIR GLM. For each ROI, the peak activity was found across conditions and subjects, and
1093 the mean activity ± 1 bin around this time point was extracted and tested using RM-ANOVAs.
1094 In ROI analyses of parametric GLM models, the mean beta weights within ROI voxels were
1095 extracted instead.

1096 *Functional connectivity analysis*

1097 Functional connectivity analyses were carried out using the CONN toolbox in
1098 conjunction with SPM12 (Whitfield-Gabrieli & Nieto-Castanon, 2012). The same ROIs
1099 described in the ROI-based analyses in Table 1 were used as seeds for focused ROI-ROI
1100 connectivity analyses and exploratory whole-brain analyses. Task-related activity was
1101 regressed out using the condition-based GLM described above, including six regressors for
1102 translational and rotational motion voxels and trial-by-trial deviations from the median of
1103 RT and slider movement durations, as well as additional nuisance regressors including five

principal components for both cerebrospinal fluid voxels and white matter, and a regressor for linear trends in the signal. The residuals of this model were then high-pass filtered at 0.008 Hz.

Connectivity data were analyzed in a two-step process. In a first-level analysis, connectivity between *a priori* pairs of ROIs was measured as the bivariate correlation between the mean residuals of voxel activities. In a second level analysis, correlation coefficients for each ROI-ROI in each subject were Fisher transformed and subjected to a one-sample *t*-test against zero. Exploratory seed-to-voxel whole brain connectivity analyses were similarly carried out by calculating the bivariate correlation of mean residual activity within an ROI and all other voxels in each subject. These correlations were then Fisher-transformed and compared to zero with one-sample t-tests. As with above fMRI analyses, all whole-brain effects were first peak thresholded at $P < 0.001$ and further thresholded at FWE cluster corrected threshold of $P < 0.05$ to control for multiple comparisons. Cluster thresholds are reported separately for each analysis.

References

Badre, D., Poldrack, R. A., Pare-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005a). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47*(6), 907-918. doi:10.1016/j.neuron.2005.07.023

Badre, D., Poldrack, R. A., Pare-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005b). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron, 47*(6), 907-918. doi:10.1016/j.neuron.2005.07.023

Badre, D., & Wagner, A. D. (2007). Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia, 45*(13), 2883-2901. doi:10.1016/j.neuropsychologia.2007.06.015

Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology, 35*(1), 48-69. doi:10.1038/npp.2009.131

Barredo, J., Oztekin, I., & Badre, D. (2015). Ventral Fronto-Temporal Pathway Supporting Cognitive Control of Episodic Memory Retrieval. *Cerebral Cortex, 25*(4), 1004-1019. doi:10.1093/cercor/bht291

Barredo, J., Verstynen, T. D., & Badre, D. (2016). Organization of cortico-cortical pathways supporting memory retrieval across subregions of the left ventrolateral prefrontal cortex. *J Neurophysiol, 116*(3), 920-937. doi:10.1152/jn.00157.2016

Barron, H. C., Dolan, R. J., & Behrens, T. E. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. *Nat Neurosci, 16*(10), 1492-1498. doi:10.1038/nn.3515

Bartlett, F. C. (1932). *Remembering: A Study in Experimental and Social Psychology*: Cambridge University Press.

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. doi:10.1016/j.neuroimage.2013.02.063

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences, 15*(11), 527-536. doi:10.1016/j.tics.2011.10.001

Blechert, J., Meule, A., Busch, N. A., & Ohla, K. (2014). Food-pics: an image database for experimental research on eating and appetite. *Front Psychol, 5*, 617. doi:10.3389/fpsyg.2014.00617

Bradfield, L. A., Dezfouli, A., van Holstein, M., Chieng, B., & Balleine, B. W. (2015). Medial Orbitofrontal Cortex Mediates Outcome Retrieval in Partially Observable Task Situations. *Neuron, 88*(6), 1268-1280. doi:10.1016/j.neuron.2015.10.044

Braga, R. M., & Buckner, R. L. (2017). Parallel Interdigitated Distributed Networks within the Individual Estimated by Intrinsic Functional Connectivity. *Neuron, 95*(2), 457-471 e455. doi:10.1016/j.neuron.2017.06.038

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433-436. doi:Doi 10.1163/156856897x00357

Brett, M., Anton, J., Valabregue, R., & Poline, J. (2002). *Region of interest analysis using an SPM toolbox*. Paper presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.

Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS One, 5*(5), e10773. doi:10.1371/journal.pone.0010773

Brodeur, M. B., Guerard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PLoS One, 9*(9), e106953. doi:10.1371/journal.pone.0106953

Chan, S. C. Y., Niv, Y., & Norman, K. A. (2016). A Probability Distribution over Latent Causes, in the Orbitofrontal Cortex. *Journal of Neuroscience, 36*(30), 7817-7828. doi:10.1523/JNEUROSCI.0659-16.2016

Chatham, C. H., & Badre, D. (2015). Multiple gates on working memory. *Curr Opin Behav Sci, 1*, 23-31. doi:10.1016/j.cobeha.2014.08.001

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *J Neurosci, 29*(39), 12315-12320. doi:10.1523/JNEUROSCI.2575-09.2009

Delgado, M. R., Beer, J. S., Fellows, L. K., Huettel, S. A., Platt, M. L., Quirk, G. J., & Schiller, D. (2016). Viewpoints: Dialogues on the functional role of the ventromedial prefrontal cortex. *Nat Neurosci, 19*(12), 1545-1552. doi:10.1038/nn.4438

Dumais, S. T. (2004). Latent semantic analysis. *Annual Review of Information Science and Technology, 38*, 189-230.

Edelstyn, N. M., Shepherd, T. A., Mayes, A. R., Sherman, S. M., & Ellis, S. J. (2010). Effect of disease severity and dopaminergic medication on recollection and familiarity in patients with idiopathic nondementing Parkinson's. *Neuropsychologia, 48*(5), 1367-1375. doi:10.1016/j.neuropsychologia.2009.12.039

Elward, R. L., Vilberg, K. L., & Rugg, M. D. (2015). Motivated Memories: Effects of Reward and Recollection in the Core Recollection Network and Beyond. *Cereb Cortex, 25*(9), 3159-3166. doi:10.1093/cercor/bhu109

Ergorul, C., & Eichenbaum, H. (2004). The hippocampus and memory for "what," "where," and "when". *Learn Mem, 11*(4), 397-405. doi:10.1101/lm.73304

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: a meta-analysis of neuroimaging studies. *Neurosci Biobehav Rev, 37*(7), 1297-1310. doi:10.1016/j.neubiorev.2013.03.023

Ghosh, V. E., & Gilboa, A. (2014). What is a memory schema? A historical perspective on current neuroscience literature. *Neuropsychologia, 53*, 104-114. doi:10.1016/j.neuropsychologia.2013.11.010

Ghosh, V. E., Moscovitch, M., Melo Colella, B., & Gilboa, A. (2014). Schema representation in patients with ventromedial PFC lesions. *J Neurosci, 34*(36), 12057-12070. doi:10.1523/JNEUROSCI.0740-14.2014

Gluth, S., Sommer, T., Rieskamp, J., & Buchel, C. (2015). Effective Connectivity between Hippocampus and Ventromedial Prefrontal Cortex Controls Preferential Choices from Memory. *Neuron, 86*(4), 1078-1090. doi:10.1016/j.neuron.2015.04.023

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience, 31*, 359-387. doi:10.1146/annurev.neuro.29.051605.112851

Haber, S. N., & Knutson, B. (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology, 35*(1), 4-26. doi:10.1038/npp.2009.129

Han, S., Huettel, S. A., Raposo, A., Adcock, R. A., & Dobbins, I. G. (2010). Functional Significance of Striatal Responses during Episodic Decisions: Recovery or Goal Attainment? *Journal of Neuroscience, 30*(13), 4767-4775. doi:10.1523/Jneurosci.3077-09.2010

Hare, T. A., Malmaud, J., & Rangel, A. (2011). Focusing attention on the health aspects of foods changes value signals in vmPFC and improves dietary choice. *J Neurosci, 31*(30), 11077-11087. doi:10.1523/JNEUROSCI.6383-10.2011

Hunt, L. T., Dolan, R. J., & Behrens, T. E. (2014). Hierarchical competitions subserving multi-attribute choice. *Nat Neurosci, 17*(11), 1613-1622. doi:10.1038/nn.3836

Izquierdo, A., Suda, R. K., & Murray, E. A. (2004). Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci, 24*(34), 7540-7548. doi:10.1523/JNEUROSCI.1921-04.2004

Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. *Brain, 129*(Pt 8), 2132-2147. doi:10.1093/brain/awl153

Jones, J. L., Esber, G. R., McDannald, M. A., Gruber, A. J., Hernandez, A., Mirenzi, A., & Schoenbaum, G. (2012). Orbitofrontal Cortex Supports Behavior and Learning Using Inferred But Not Cached Values. *Science, 338*(6109), 953-956. doi:Doi 10.1126/Science.1227489

Kahneman, D., & Tversky, A. (1979). Prospect Theory - Analysis of Decision under Risk. *Econometrica, 47*(2), 263-291. doi:Doi 10.2307/1914185

Kan, I. P., Larocque, K. F., Lafleche, G., Coslett, H. B., & Verfaellie, M. (2010). Memory monitoring failure in confabulation: evidence from the semantic illusion paradigm. *J Int Neuropsychol Soc, 16*(6), 1006-1017. doi:10.1017/S1355617710000536

King, D. R., Chastelaine, M., Elward, R. L., Wang, T. H., & Rugg, M. D. (2018). Dissociation between the neural correlates of recollection and familiarity in the striatum and hippocampus: Across-study convergence. *Behav Brain Res, 354*, 1-7. doi:10.1016/j.bbr.2017.07.031

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception, 36*, 14-14.

Knutson, B., Wimmer, G. E., Kuhnen, C. M., & Winkielman, P. (2008). Nucleus accumbens activation mediates the influence of reward cues on financial risk taking. *Neuroreport, 19*(5), 509-513. doi:10.1097/WNR.0b013e3282f85c01

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*(2-3), 259-284. doi:Doi 10.1080/01638539809545028

Leong, Y. C., Radulescu, A., Daniel, R., DeWoskin, V., & Niv, Y. (2017). Dynamic Interaction between Reinforcement Learning and Attention in Multidimensional Environments. *Neuron, 93*(2), 451-463. doi:10.1016/j.neuron.2016.12.040

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Curr Opin Neurobiol, 22*(6), 1027-1038. doi:10.1016/j.conb.2012.06.001

Lim, S. L., O'Doherty, J. P., & Rangel, A. (2013). Stimulus value signals in ventromedial PFC reflect the integration of attribute value signals computed in fusiform gyrus and posterior superior temporal gyrus. *J Neurosci, 33*(20), 8729-8741. doi:10.1523/JNEUROSCI.4809-12.2013

Milner, B., Corkin, S., & Teuber, H. L. (1968). Further Analysis of Hippocampal Amnesic Syndrome - 14-Year Follow-up Study of Hm. *Neuropsychologia, 6*(3), 215-&. doi:Doi 10.1016/0028-3932(68)90021-3

Moscovitch, M., & Melo, B. (1997). Strategic retrieval and the frontal lobes: evidence from confabulation and amnesia. *Neuropsychologia, 35*(7), 1017-1034.

Moscovitch, M., Nadel, L., Winocur, G., Gilboa, A., & Rosenbaum, R. S. (2006). The cognitive neuroscience of remote episodic, semantic and spatial memory. *Curr Opin Neurobiol, 16*(2), 179-190. doi:10.1016/j.conb.2006.03.013

Nachev, P., Lopez-Sosa, F., Gonzalez-Rosa, J. J., Galarza, A., Avecillas, J., Pineda-Pardo, J. A., . . . Strange, B. (2015). Dynamic risk control by human nucleus accumbens. *Brain, 138*, 3496-3502. doi:10.1093/brain/awv285

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature, 442*(7106), 1042-1045. doi:10.1038/nature05051

Piaget, J. (1952). *The Origins of Intelligence in Children* (Vol. 8). International Universities Press.

Postman, L., & Keppel, G. (1970). *Norms of Word Asociation*. New York, New York: Academic Press.

Price, J. L. (2007). Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Ann N Y Acad Sci, 1121*, 54-71. doi:annals.1401.008 [pii]
10.1196/annals.1401.008

Ratcliff, R. (1978). Theory of Memory Retrieval. *Psychological Review, 85*(2), 59-108. doi:Doi 10.1037//0033-295x.85.2.59

Reber, J., Feinstein, J. S., O'Doherty, J. P., Liljeholm, M., Adolphs, R., & Tranel, D. (2017). Selective impairment of goal-directed decision-making following lesions to the human ventromedial prefrontal cortex. *Brain, 140*(6), 1743-1756. doi:10.1093/brain/awx105

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effective- ness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning. II. Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

Roy, M., Shohamy, D., & Wager, T. D. (2012). Ventromedial prefrontal-subcortical systems and the generation of affective meaning. *Trends Cogn Sci, 16*(3), 147-156. doi:10.1016/j.tics.2012.01.005

Rudebeck, P. H., Saunders, R. C., Prescott, A. T., Chau, L. S., & Murray, E. A. (2013). Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat Neurosci, 16*(8), 1140-1145. doi:10.1038/nn.3440

Rudorf, S., & Hare, T. A. (2014). Interactions between Dorsolateral and Ventromedial Prefrontal Cortex Underlie Context-Dependent Stimulus Valuation in Goal-Directed Choice. *Journal of Neuroscience, 34*(48), 15988-15996. doi:10.1523/JNEUROSCI.3192-14.2014

Saleem, K. S., Kondo, H., & Price, J. L. (2008). Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. *J Comp Neurol, 506*(4), 659-693. doi:10.1002/cne.21577

Salmon, D. P., & Bondi, M. W. (2009). Neuropsychological assessment of dementia. *Annu Rev Psychol, 60*, 257-282. doi:10.1146/annurev.psych.57.102904.190024

Schnider, A. (2013). Orbitofrontal reality filtering. *Front Behav Neurosci, 7*, 67. doi:10.3389/fnbeh.2013.00067

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. doi:10.1016/j.neuron.2016.08.019

Schwarze, U., Bingel, U., Badre, D., & Sommer, T. (2013). Ventral striatal activity correlates with memory confidence for old- and new-responses in a difficult recognition test. *PLoS One, 8*(3), e54324. doi:10.1371/journal.pone.0054324

Scimeca, J. M., & Badre, D. (2012). Striatal contributions to declarative memory retrieval. *Neuron, 75*(3), 380-392. doi:10.1016/j.neuron.2012.07.014

Scimeca, J. M., Katzman, P. L., & Badre, D. (2016). Striatal prediction errors support dynamic control of declarative memory decisions. *Nat Commun, 7*, 13061. doi:10.1038/ncomms13061

Shadlen, M. N., & Shohamy, D. (2016). Decision Making and Sequential Sampling from Memory. *Neuron, 90*(5), 927-939. doi:10.1016/j.neuron.2016.04.036

Spaniol, J., Davidson, P. S., Kim, A. S., Han, H., Moscovitch, M., & Grady, C. L. (2009). Event-related fMRI studies of episodic encoding and retrieval: meta-analyses using activation likelihood estimation. *Neuropsychologia, 47*(8-9), 1765-1779. doi:10.1016/j.neuropsychologia.2009.02.028

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc Natl Acad Sci U S A, 94*(26), 14792-14797.

Tulving, E. (1972). Episodic and Semantic Memory. In E. Tulving & W. Donaldson (Eds.), *Organization of Memory*. London, UK: Academic Press, Inc.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . . Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15*(1), 273-289. doi:10.1006/nimg.2001.0978

Vaidya, A. R., & Fellows, L. K. (2015). Testing necessary regional frontal contributions to value assessment and fixation-based updating. *Nat Commun, 6*, 10120. doi:10.1038/ncomms10120

Vaidya, A. R., Sefranek, M., & Fellows, L. K. (2017). Ventromedial Frontal Lobe Damage Alters how Specific Attributes are Weighed in Subjective Valuation. *Cereb Cortex*, 1-11. doi:10.1093/cercor/bhx246

Whitfield-Gabrieli, S., & Nieto-Castanon, A. (2012). Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect, 2*(3), 125-141. doi:10.1089/brain.2012.0073

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. doi:10.1016/j.neuron.2013.11.005

Wimmer, G. E., Li, J. M. K., Gorgolewski, K. J., & Poldrack, R. A. (2018). Reward Learning over Weeks Versus Minutes Increases the Neural Representation of Value in the Human Brain. *Journal of Neuroscience, 38*(35), 7649-7666. doi:10.1523/Jneurosci.0075-18.2018

Xia, C., Stolle, D., Gidengil, E., & Fellows, L. K. (2015). Lateral orbitofrontal cortex links social impressions to political choices. *J Neurosci, 35*(22), 8507-8514. doi:10.1523/JNEUROSCI.0526-15.2015

Yu, L. Q., Dana, J., & Kable, J. W. (2018). Individuals with ventromedial frontal damage have more unstable but still fundamentally transitive preferences. *bioRxiv, 384024*. doi:10.1101/384024

Zalocusky, K. A., Ramakrishnan, C., Lerner, T. N., Davidson, T. J., Knutson, B., & Deisseroth, K. (2016). Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature, 531*(7596), 642-646. doi:10.1038/nature17400

Zhou, J., Gardner, M. P. H., Stalnaker, T. A., Ramus, S. J., Wikenheiser, A. M., Niv, Y., & Schoenbaum, G. (2019). Rat Orbitofrontal Ensemble Activity Contains Multiplexed but Dissociable Representations of Value and Task Structure in an Odor Sequence Task. *Curr Biol, 29*(6), 897-907 e893. doi:10.1016/j.cub.2019.01.048

Neural systems for memory-based value judgment and decision-making

A.R. Vaidya & D. Badre

Supplementary methods

Alternate models

We developed a simple base model that made a set of assumptions about behavior in this task that were carried over to all alternative models. Each alternative model made a modification of this base model, or the next simplest model from the last step. Alternative models were compared based on their improvement in AIC scores from the last best fitting model from a previous step. The results of this stepwise fitting procedure are provided in Supplementary Tables 6-9. Detailed description of the base model and each of these modifications are below:

(1) Base model

In the base model, the expected value (V) of each ingredient (i) was the weighted sum of its episodic value (EV) and schema value (SV) (these are defined below) as in

$$V_i = \beta_e EV_i + \beta_s SV_i \text{ (Eq. 8)}$$

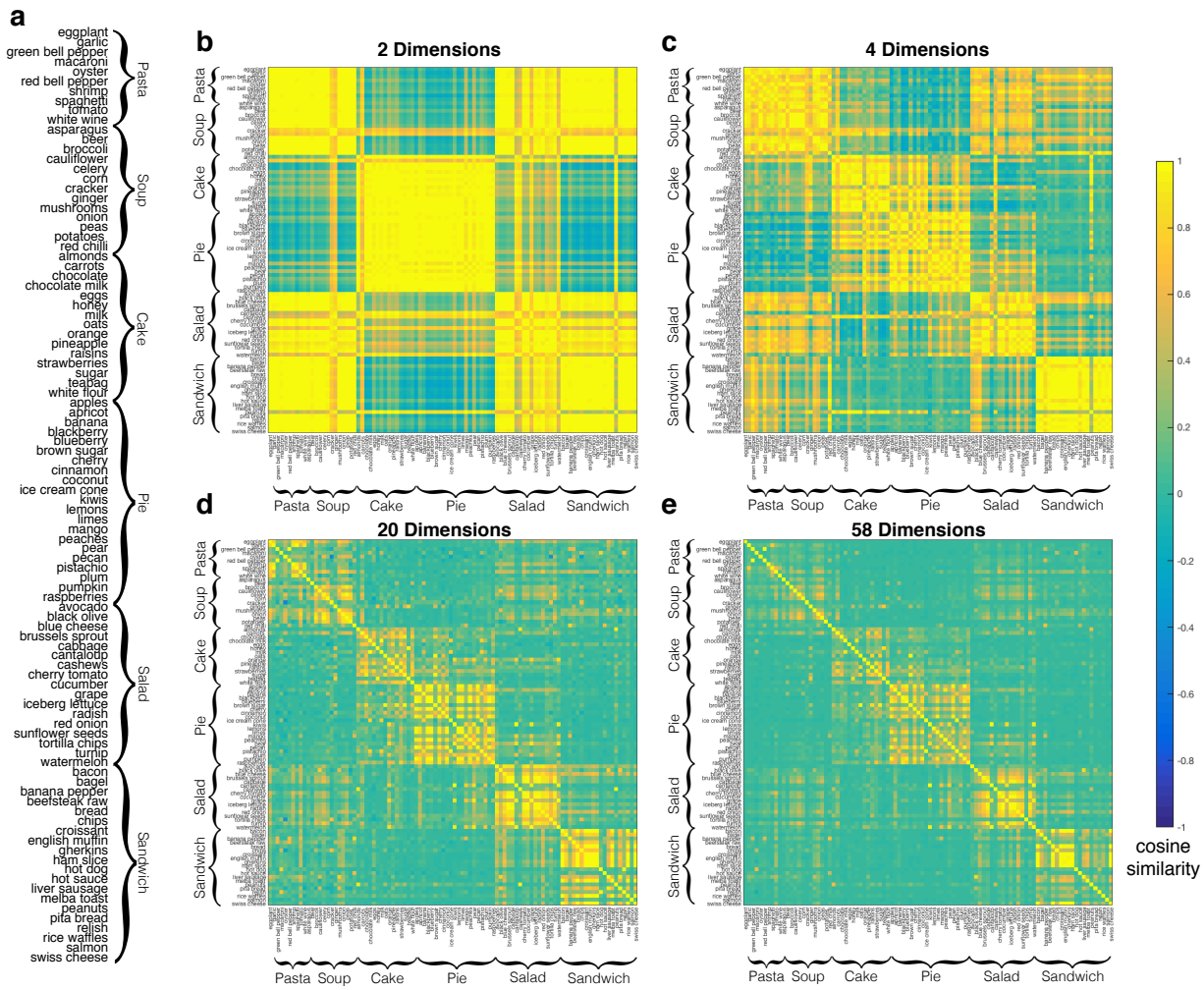
Where β_e and β_s are free temperature parameters that determine the weight of these different values to V_i. This ingredient value was then placed in a logistic function to calculate the model-P(yes) for each trial.

On each trial where participants responded, the episodic value for each ingredient was updated according to a simple Rescorla-Wagner learning rule, as in Eq. 2., described in the main methods section. On each trial, the EV of all ingredients would decay towards zero according to Eq. 3.

The *SV* of each ingredient was simply the difference of the associative strength of that ingredient with the positive recipe (*PR*) and negative recipe (*NR*) in the base model. These associations were determined using a procedure similar to Latent Semantic Analysis (Dumais, 2004; Landauer, Foltz, & Laham, 1998). The endorsement frequencies for ingredients across all recipes in the normative mTurk data were weighted by their entropy across recipes and a singular value decomposition (SVD) was carried out on this matrix. The association of an ingredient with each recipe was calculated as the mean cosine similarity between that ingredient and all ingredients in each recipe cluster within this reconstituted matrix. In the base model, the number of eigenvalues used to reconstitute the matrix was set to 58. This was found to be the minimum number of eigenvalues needed to almost perfectly reconstitute the ingredient-ingredient cosine similarity matrix using the raw recipe endorsement frequencies in the mTurk data (Pearson $r = 0.99$).

(2) <u>Recipe dimensions model</u>

The recipe dimensions model was identical to the base model, except the number of dimensions to retain before reconstituting the ingredient-by-recipe endorsement frequency matrix was allowed to vary according to a free parameter *c* (see Supplementary Figure 1). This parameter essentially controlled the dimensionality of recipe space that the model would consider, so that a lower dimensional model would separate ingredients on only simple features (i.e. savory or sweet), while a higher dimensional model would include increasingly granular information that separates ingredients along information about specific recipes.



Supplementary Figure 1. Example ingredient by ingredient cosine similarity matrices from reconstructions of ingredient-recipe associations with increasing dimensions. **a.** Ingredient names and clustering by recipe. **b.** Cosine similarity matrix with two dimensions **c.** with four dimensions, **d.** with 20 dimensions, and **e.** with 58 dimensions.

(3) <u>Asymmetric schemas model</u>

This model included separate temperature parameters for positive and negative schema values, so that the value of an ingredient was:

$$V_i = \beta_e \, EV_i + \beta_{sp} PR_i - \beta_{sn} NR_i \text{ (Eq. 9)}$$

Where β_{sp} and β_{sn} are temperature parameters that control the influence of ingredient associations with positive and negative recipes, respectively.

(4) Schema similarity weighting model

The recipes in a block could vary in their similarity, making some ingredients plausibly associated with both positive or negative recipes. The schema similarity model estimated participants' uncertainty about the recipe associations of each ingredient as the probability that an ingredient could be used in both recipes ($P(SR)$), calculated as in Eq. 4. This probability was calculated using an additional parameter β_{sim}, which controlled the sensitivity of participants to the similarity of ingredients' recipe associations, and the relative contribution of the difference or sum of positive and negative recipe associations to *SV.* This probability was then used in calculating weights for the sum and difference of positive and negative recipe associations. These weights were then applied in calculating *SV* as in Eq. 7.

(5) Unfamiliarity bias model

The unfamiliarity bias model allowed the willingness to endorse an ingredient vary according to the strength of the episodic memory for that ingredient's value (i.e. the absolute value of *EV*). In this model the value of each ingredient was calculated as:

$$V_i = \beta_u(1 - |EV_i|) + \beta_e\, EV_i + \beta_s SV_i \text{ (Eq. 10)}$$

Where β_u is a free temperature parameter that could be positive or negative and controls the model's tendency to endorse or reject an ingredient based on the strength of episodic memory.

(6) Fixed bias model

This model introduced a simple intercept term to the calculation of ingredient values, so that the value of an ingredient on each trial was:

$$V_i = \beta_0 + \beta_e EV_i + \beta_s SV_i \text{ (Eq. 11)}$$

Where β_0 is a free parameter that could be positive or negative and controls an overall tendency to endorse or reject ingredients across all trials.

(7) Schema discounting model

As participants learned the deterministic values of ingredients from feedback, it might be useful to use schema information less. This model assumed that when ingredients were associated with stronger episodic memories (i.e. the absolute value of *EV*), the schema value information would be exponentially discounted. This would suggest that participants trade off in retrieving information from different memory sources, especially when one is more reliable. For each ingredient, a discount factor was calculated as:

$$discount = \exp(-k|EV_i|) \text{ (Eq. 12)}$$

Where k was a free parameter that determined the steepness of this discounting. This discount factor was then included in the calculation of ingredient values:

$$V_i = \beta_e\, EV_i + discount\, \beta_s\, SV_i \text{ (Eq. 13)}$$

So that the contribution of schema values scaled with this exponential discounting factor.

(8) Dual Process Recollection Model

While the above models assume that episodic memory is a graded process, other models developed in memory judgment tasks suggest that episodic memory has a graded component (familiarity) and a discrete component (recollection) (Yonelinas, 1994). To test whether episodic values in this task may also rely on an additional non-graded recollection process, ingredient values were calculated as:

$$V_i = \beta_e((1 - P_r)EV_i + P_r RV_i) + \beta_s SV_i \text{ (Eq. 14)}$$

Where *RV* is the recollected value of an ingredient, i.e. the outcome from previous trials that does not decay or change over time (-1 or 1), and P_r is a fixed probability of recollection that

does not change over trials. Thus, in this model, values learned through feedback contributed to the decision as the sum of recalled and graded value information, weighted by the probability of recall.

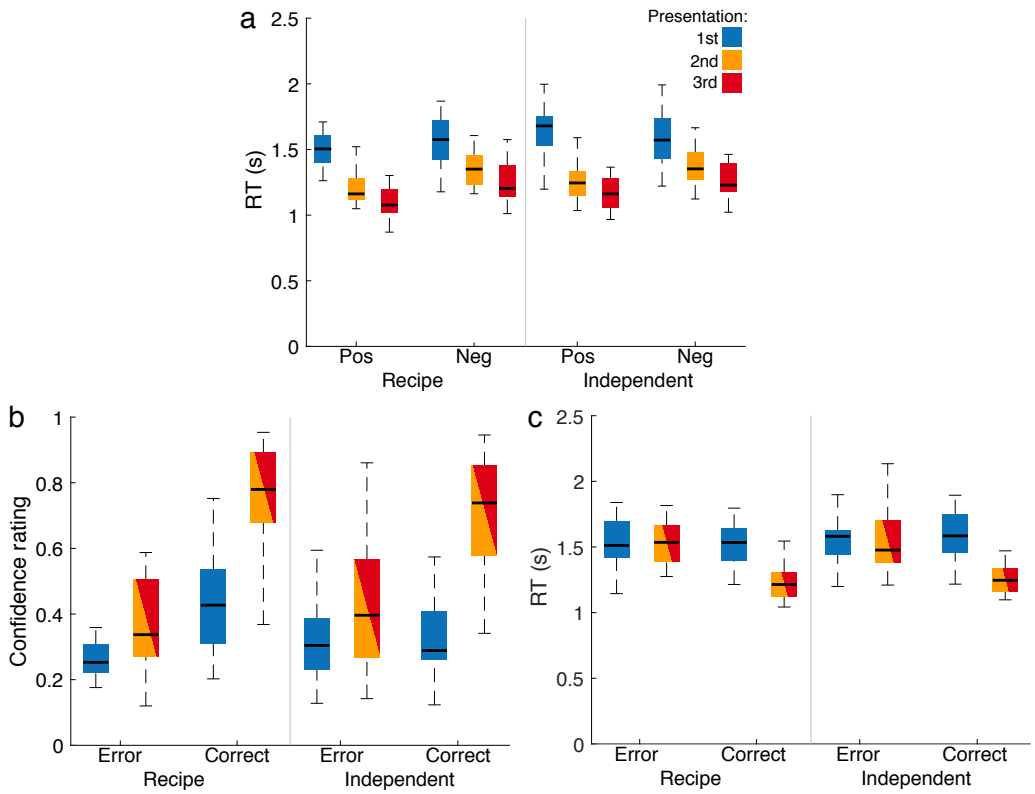
Supplementary Results

Confidence and reaction times

Similar to reported confidence, participants' reaction times were also strongly influenced by experience with ingredients with stronger decrements in subsequent presentations of positive recipe ingredients (Supplementary Figure 2a). There were significant interactions between value and clustering (RM-ANOVA: $F_{1,20} = 4.66$, $P = 0.04$, $\eta_p^2 = 0.19$), and value with presentation number ($F_{2,40} = 8.50$, $P = 0.0008$, $\eta_p^2 = 0.30$), but there was no significant three-way interaction. Responses in the first presentation did not significantly differ by cluster condition or value. In the third presentation, participants were significantly faster in accepting positive recipe ingredients than positive independent ingredients (Bonferroni corrected t-tests: $t(20) = 4.60$, $P < 0.01$, $d = 1.00$), though there were no other significant reaction time differences between recipe and independent conditions. Reaction times were also consistently faster in the second and third presentation compared to the first presentation across all conditions (Bonferroni corrected t-tests: t's$(20) \geq 4.60$, P's < 0.01, $d \geq 1.36$), and faster in the third presentation than the second (Bonferroni corrected t-tests: t's$(20) \geq 3.57$, P's < 0.05, d's ≥ 0.78).

We separately tested the effects of accuracy on confidence ratings and reaction times, collapsing across value levels and the second and third presentations to have enough error trials per participant for this analysis (Supplementary Figure 2cd). Confidence ratings were influenced by significant interactions between accuracy and clustering (RM-ANOVA: $F_{1,21} = 15.43$, $P = 0.0007$, $\eta_p^2 = 0.42$), and accuracy and presentation number ($F_{1,21} = 43.69$, $P < 0.0001$, $\eta_p^2 = 0.67$), but no three-way interaction. On the first presentation, participants were significantly more confident on correct responses for recipe ingredients than independent

ingredients (Bonferroni corrected t-tests: $t(21) = 3.77$, $P < 0.05$, $d = 0.80$), and error responses on recipe ingredients (Bonferroni corrected t-tests: $t(21) = 5.60$, $P < 0.01$, $d = 1.19$). There was no comparable significant difference in confidence for correct and error responses for independent ingredients in the first presentation. We confirmed that the effects of accuracy differed between clustering conditions on this first presentation by separately testing this interaction (RM-ANOVA: $F_{1, 21} = 14.34$, $P = 0.001$, $\eta_p^2 = 0.40$). Confidence was also higher on accurate responses compared to errors on the second and third presentations of both recipe and independent ingredients (Bonferroni corrected t-tests: t's$(21) \geq 6.64$, P's < 0.0001, $d = 1.42$).



Supplmentary Figure 2. Confidence and reaction times (RT) across task conditions. **a**. RT for correct responses in recipe and independent conditions, for positive (pos) and negative (neg) ingredients. **b.** Confidence ratings for errors and correct responses in recipe and independent conditions, collapsing across value. Second and third presentations are also collapsed. **c.** RT for correct and error responses in the same conditions as **b**.

134

135 Reaction times were influenced by an interaction of accuracy and presentation
136 number (RM-ANOVA: $F_{1,21} = 49.84$, $P < 0.0001$, $\eta_p^2 = 0.70$), but there was only a trend toward
137 an interaction between accuracy and clustering ($F_{1,21} = 3.81$, $P = 0.06$, $\eta_p^2 = 0.15$), and no
138 three-way interaction. Participants were significantly faster on correct responses than
139 errors on the second and third presentations of both recipe and independent ingredients
140 (Bonferroni corrected t-tests: t's(21) ≥ 5.03, P's < 0.01, d ≥ 1.07). In contrast, there were no
141 significant differences in reaction times for correct and error responses in the first
142 presentation for either clustering condition.

143 *Medial prefrontal cortex ROI*
144
145 As a supplement to our tests within the vmPFC ROI defined based on studies of
146 expected value, we also examined a more dorsal medial PFC (mPFC) ROI centered around a
147 region previously shown to be involved in inferring the common recipe associations of
148 ingredients in a separate experiment (Barron, Dolan, & Behrens, 2013) . This ROI also
149 responded strongly to value ($F_{1,20} = 21.19$, $P = 0.0002$, $\eta_p^2 = 0.51$), but not to any other factor
150 or their interaction.

151 We also included this mPFC ROI in the generalized psychophysical interaction (gPPI)
152 analysis to test if connectivity between this region and any other ROI differed on the first
153 presentation of positive and negative value recipe ingredients, but found no significant
154 difference.

155 *Region of interest analyses — accuracy*

156

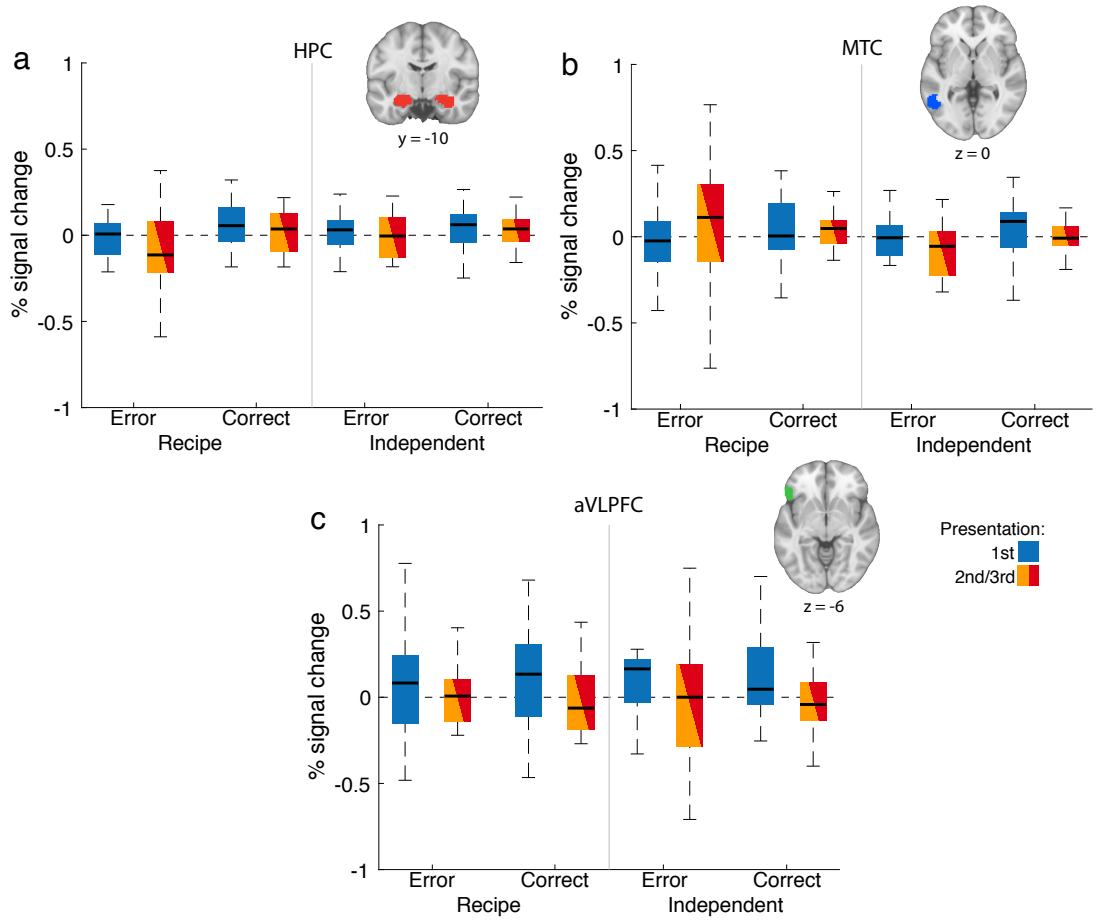
We next assessed how ROIs hypothesized to be involved in episodic and semantic memory were engaged by differential memory demands in this task and by retrieval success and failure. To do so, we evaluated the effects of clustering status, presentation number and accuracy in ROI-based analyses of the HPC, MTC and aVLPFC. We collapsed across value conditions and combined presentations two and three into a single level to have enough error trials for this analysis.

HPC activity was greater on accurate trials, confirming the whole-brain results (Supplementary Figure 3a; RM-ANOVA: $F_{1, 21} = 13.14$, $P = 0.002$, $\eta_p^2 = 0.38$). HPC activity was also generally lower on later presentations ($F_{1, 21} = 6.54$, $P = 0.02$, $\eta_p^2 = 0.24$). The effect of accuracy on HPC activity was modulated marginally by clustering ($F_{1, 21} = 4.65$, $P = 0.04$, $\eta_p^2 = 0.18$), with greater HPC activity on correct trials than error trials for recipe ingredients, collapsing across presentations (Bonferroni corrected t-test: $t(21) = 3.47$, $P < 0.01$, $d = 0.74$), but no such accuracy effect for independent ingredients. There was no main effect of clustering, or three-way interaction of clustering, accuracy and presentation.

MTC activity responded to an interaction of clustering and presentation number (Supplementary Figure 3b; $F_{1, 21} = 5.60$, $P = 0.03$, $\eta_p^2 = 0.21$). Posthoc tests collapsed across accuracy revealed a significantly greater reduction in MTC activity for independent ingredients in the second and third presentations compared to the first presentation (Bonferroni corrected t-test: $t(21) = 2.87$, $P < 0.05$, $d = 0.61$), but no significant difference for recipe ingredients.

aVLPFC activity was larger on the first presentation compared to the second and third, consistent with the whole-brain analysis, revealed in a main effect of presentation number (Supplementary Figure 3c; $F_{1, 21} = 10.76$, $P = 0.004$, $\eta_p^2 = 0.34$). The response of

aVLPFC did not change with accuracy or clustering status, or with any interaction between task variables.



Supplementary Figure 3. Region-of-interest analyses for correct and erroneous responses for recipe and independent ingredients, collapsing across value. Presentations two and three have been collapsed into one level. Percent signal change in **a.** hippocampus (HPC), **b.** middle temporal cortex (MTC) and **c.** anterior ventrolateral prefrontal cortex (aVLPFC).

Parameter recovery test

To test the recoverability of the parameters of the best fitting behavioral model, we generated 100 simulated datasets with randomized parameters. We then tested if these parameters could be recovered by fitting the same model to these simulated datasets in

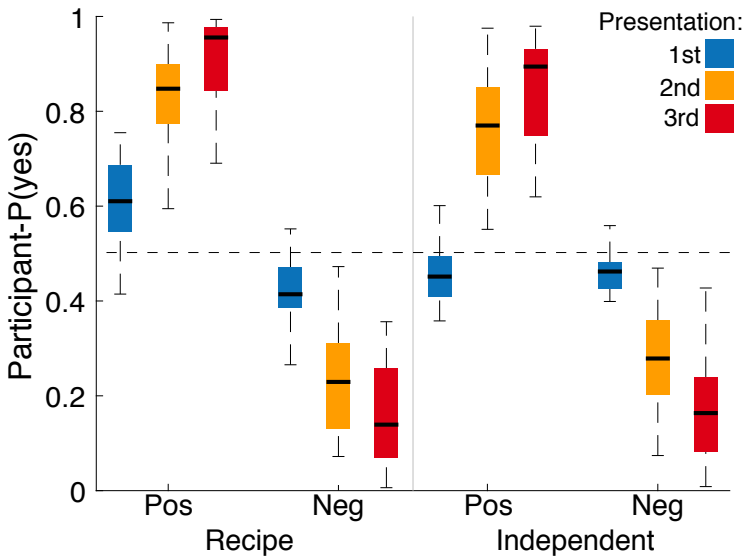
exactly the same way as real participant data. We were able to reliably and accurately recover all parameters in these data, with high correlations between simulated and recovered parameters (Pearson r's(98) ≥ 0.99, P's < 0.0001). However, as these data were simulated from randomly generated parameters, they did not share the covariance structure of the parameter estimates fit to participants' behavior described above.

To test if participants' own parameters could be reliably recovered, we fit the model to simulated data generated using parameter estimates from the model fit to participants' behavior. For all seven parameters, there was a perfect correlation between the parameter estimates from participants' behavior and the simulated data (Pearson r's(21) = 1.00, P's < 0.0001).

Participant-P(yes) ratings

We examined how the continuous participant-p(yes) ratings changed as a function of value (positive versus negative), clustering (recipe versus independent) and presentation number (one, two and three; Supplementary Figure 4). There was a three-way interaction between these three factors (RM-ANOVA: $F_{2, 44} = 8.82$, $P = 0.0006$, $\eta_p^2 = 0.29$). Critically, on the first presentation, participant-P(yes) was significantly higher for positive recipe ingredients than negative recipe ingredients, and positive independent ingredients (Bonferroni corrected t-test: $t(22) = 6.57$, P < 0.0001, $d = 1.37$), while there was no such difference between positive and negative independent ingredients. Participant-P(yes) was also lower than chance on the first presentation for positive and negative independent ingredients (One-sample t-tests against chance: t's(22) ≥ 3.15, P's ≤ 0.005, d's ≥ 0.66), indicating a bias to reject options prior to learning their values. These ratings tracked the learned value of ingredients after feedback in presentations two and three compared to one

(Bonferroni corrected t-tests: t's(22) ≥ 7.20, P's < 0.0001, d's ≥ 1.50), and presentation three compared to two (Bonferroni corrected t-tests: t's(22) ≥ 5.40, P's < 0.01, d's ≥ 1.13). Both positive and negative recipe ingredients showed an advantage over independent ingredients in the second presentation (Bonferroni corrected t-tests: t's(22) ≥ 3.59, P's < 0.05, d's ≥ 1.13), though only positive ingredients showed a significant advantage in the third presentation (Bonferroni corrected t-test: $t(22) = 4.83$, $P < 0.01$, $d = 1.00$).



Supplementary Figure 4. Participant-P(yes) ratings across task conditions.

Model behavior analysis

Like with participant-P(yes), there was a strong three-way interaction between value, clustering and presentation number (RM-ANOVA: $F_{2, 44} = 53.82$, $P < 0.0001$, $\eta_p^2 = 0.71$). The model showed evidence of using recipe information to make decisions on the first presentation, with higher model-P(yes) ratings for positive recipe ingredients over negative recipe ingredients, and over independent positive ingredients (Bonferroni corrected t-tests: t's(22) ≥ 7.38, P's < 0.0001, d's ≥ 1.54), and no significant difference in model-P(yes) for positive and negative independent ingredients on this first presentation.

Unlike participant-P(yes), the model gave significantly lower ratings to negative recipe ingredients than negative independent ingredients on the first presentation (Bonferroni corrected t-test: $t(22) = 6.64$, $P < 0.0001$, $d = 1.38$), suggesting that it used information about the negative recipe more than participants. The model-P(yes) for independent ingredients was also significantly lower than 0.5 on the first presentation, indicating a bias against betting on these ingredients before knowing their value (one-sample t-tests: $t(22) \geq 2.16$, $P \leq 0.04$, d's ≥ 0.45) . The model also showed clear evidence of learning ingredient values, with the model-P(yes) consistently hewing closer to ingredient values in presentation two and three compared to one (Bonferroni corrected t-tests: t's$(22) \geq 9.23$, P's < 0.0001, $d \geq 1.93$), and presentation three compared to two (Bonferroni corrected t-tests: t's$(22) \geq 5.72$, P's < 0.01, $d \geq 1.19$). Like with control subjects, positive recipe ingredients maintained an advantage in the second and third presentations over positive independent ingredients, but the same was not true in the contrast of negative recipe and independent ingredients.

Similarly, we examined whether the model's predicted confidence followed the same pattern as participants' responses across task conditions on trials where participants responded correctly. Model confidence rose predictably with presentation number (RM-ANOVA: $F_{2, 40} = 150.65$, $P < 0.0001$, $\eta_p^2 = 0.88$), and was higher overall for recipe ingredients than independent ingredients ($F_{1, 20} = 46.18$, $P < 0.0001$, $\eta_p^2 = 0.70$), however there was only a trend toward a main effect of value ($F_{1, 20} = 3.45$, $P = 0.08$, $\eta_p^2 = 0.15$). There was no significant three-way interaction between these factors. Like participants' ratings, there was also a significant two-way interaction of value and presentation number ($F_{2, 40} = 6.59$, $P = 0.003$, $\eta_p^2 = 0.25$). However, unlike participants, there was also a significant interaction of clustering and presentation number ($F_{2, 40} = 8.65$, $P = 0.0007$, $\eta_p^2 = 0.30$) and no significant

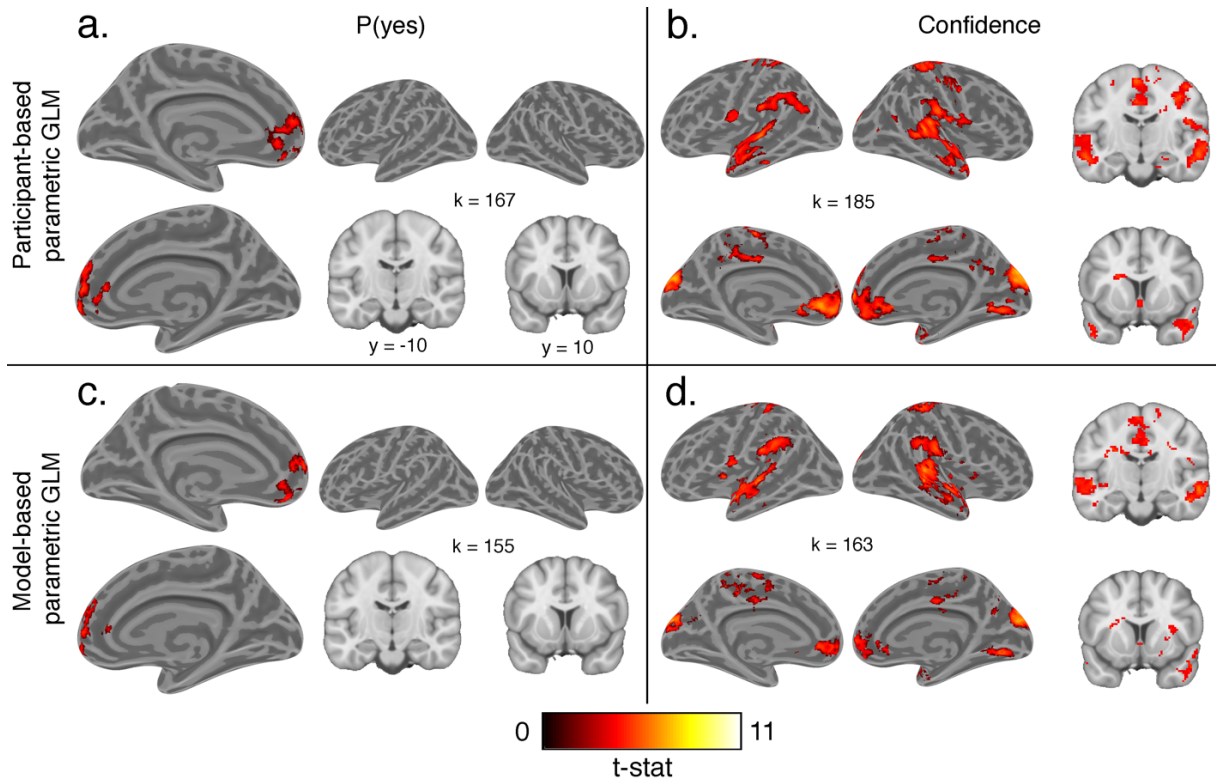
interaction of value and clustering. Model confidence was notably higher for both positive and negative recipe ingredients compared to independent ingredients on the first presentation (Bonferroni corrected t-tests: t's(20) ≥ 4.02, P's < 0.05, d's ≥ 0.88), rather than just positive recipe ingredients. While the model was able to capture most of the key features of participants' confidence responses, it underestimated the effect of value on confidence and did not reflect the asymmetric confidence for positive and negative recipe ingredients.

Validation analyses for model-based fMRI

Two parametric GLMs were run on these functional data for this purpose. The first included parametric modulators of the response regressor for participant-P(yes), confidence ratings and a nuisance term containing participants' own subjective value ratings for each ingredient from a separate task. This model also had modulators on the feedback regressor for the reward prediction error (RPE; based on the difference of the outcome and participant-P(yes)) and surprise (the absolute value of the RPE). The second GLM included the same parametric modulators, but were derived from the model-P(yes), as well as residual terms for model-P(yes), confidence, RPE and surprise, calculated from the difference of the model's prediction and participants' behavior in each trial. These residual terms were included to let us examine variance in activity associated with value judgment and processing of feedback that was not captured by our model.

Supplementary Figure 5 shows the whole-brain cluster-corrected statistical map for the P(yes) and confidence modulators of both GLMs. vmPFC responded to both participant- and model-P(yes), which was confirmed by an ROI analysis of the beta coefficients from these GLMs (one-tailed, one-sample t-tests: t's(22) ≥ 4.20, P ≤ 0.0002, d's = 0.88). The striatum ROI also showed a significant effect for participant-P(yes) (t(22) = 2.09, P = 0.02, d

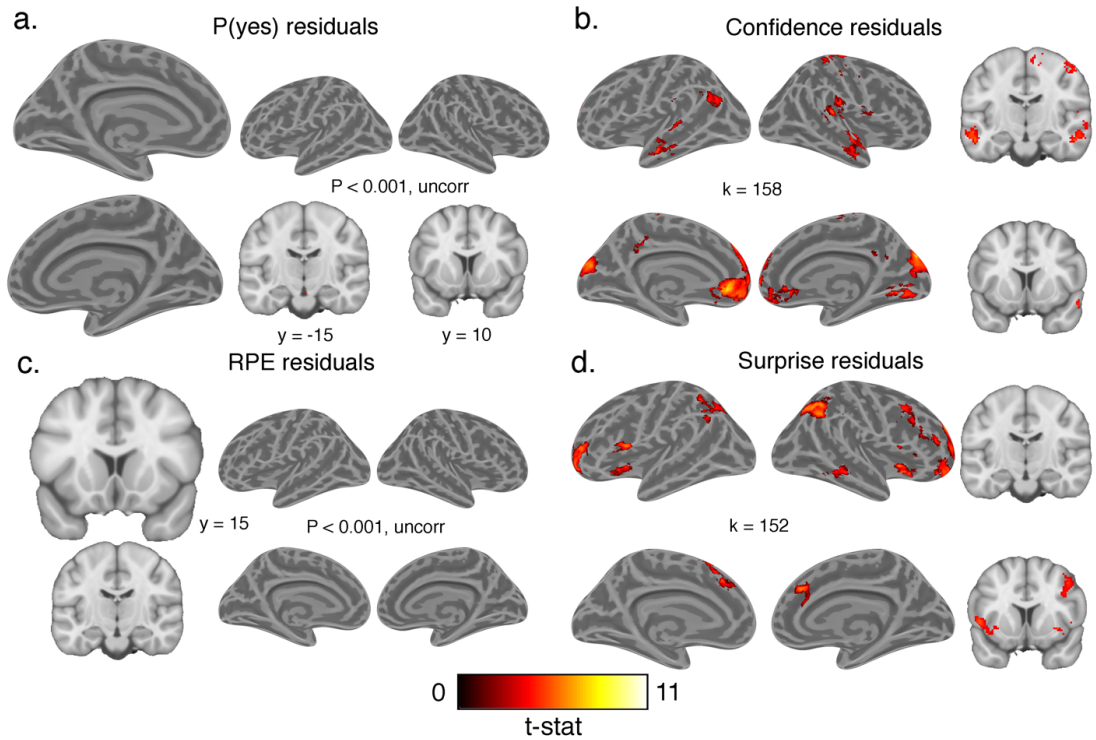
= 0.43), though this did not pass whole-brain cluster corrected threshold in the voxel-wise analysis. The effect of model-P(yes) in the striatum similarly did not reach significance at a cluster-corrected threshold, though there was a trend in an ROI-level analysis ($t(22) = 1.64$, $P = 0.06$, $d = 0.34$). There was no significant activity associated with P(yes) residuals in a whole-brain analysis, either as a cluster corrected threshold or an uncorrected peak threshold of $P = 0.001$ (Supplementary Figure 6).



Supplementary Figure 5. Whole brain t-statistic maps from parametric models based on participant and model behavior. Top row shows correlates for participant **a.** P(yes) and **b.** confidence. Bottom row shows correlates for model **c.** P(yes) and **d**. confidence. All images are thresholded at $P < 0.05$, FWE cluster corrected. Cluster extent threshold for each map is given by value of k.

vmPFC was also significantly associated with participant and model confidence at a whole-brain cluster corrected threshold and ROI level (one-tail, one-sample t-tests: t's (22) ≥ 3.55, $P \leq 0.0008$, $d = 0.74$), indicating that value and confidence signals were

simultaneously present within this region, as in previous work (Lebreton, Abitbol, Daunizeau, & Pessiglione, 2015). Participants' confidence was also significantly associated with activity in the HPC at an ROI-level ($t(22) = 2.70$, $P = 0.006$, $d = 0.56$), and marginally with MTC at an ROI-level ($t(22) = 1.68$, $P = 0.05$). The model confidence term was significantly associated with activity in HPC at an ROI-level ($t(22) = 1.96$, $P = 0.03$, $d = 0.41$), though no clusters in this region survived whole-brain cluster correction. Model confidence did not reach significance in the MTC ROI ($t(22) = 1.05$, $P = 0.15$, $d = 0.22$). There were significant effects for confidence residuals at a cluster corrected whole-brain level in several regions, including vmPFC (Supplementary Figure 6). This residual activity related to confidence was unsurprising given that the model was fit directly to participants' P(yes) ratings, and only captured confidence ratings incidentally.

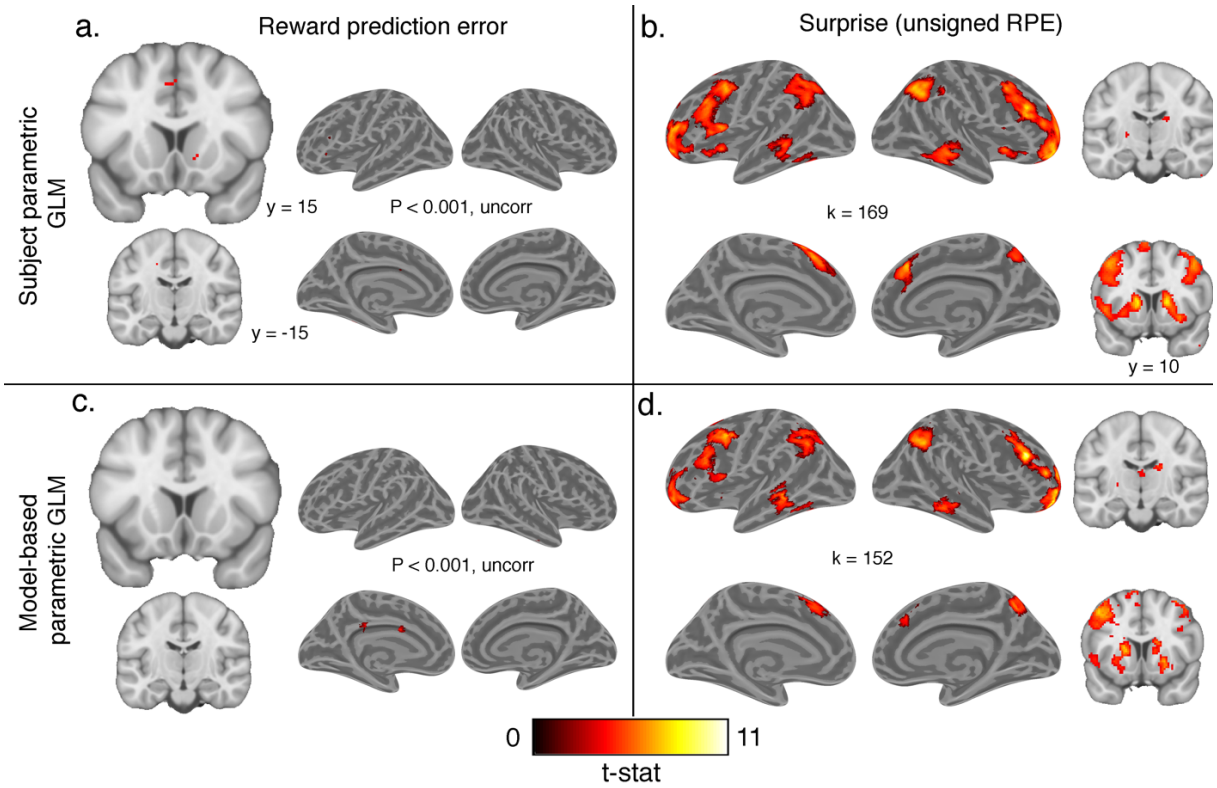


Supplementary Figure 6. Whole brain t-statistic maps showing activations associated with residuals of model-based regressors. **a.** P(yes), **b.** confidence, **c.** Reward prediction error (RPE) and **d**. surprise. P(yes) and RPE maps are thresholded at $P < 0.001$, uncorrected. Maps for surprise are thresholded at $P < 0.05$, FWE cluster corrected. Cluster extent threshold for each map is given by value of k.

In a secondary analysis of feedback-based modulators, we also found good correspondence between the model and participant RPE and surprise regressors (Supplementary Figure 7). There were no significant clusters associated with participants' RPE at a whole-brain level, nor was there any significant effect for the RPE in the striatum in an ROI analysis ($t(22) = 1.55$, $P = 0.07$, $d = 0.32$). Similarly, there were no significant effects of model RPE at a whole brain level or in the striatum ROI ($t(22) = 0.18$, $P = 0.4$, $d = 0.04$). However, at the whole-brain level, there was robust activation in response to both model and participant surprise in bilateral areas of the striatum, lateral PFC, supplementary motor area and angular gyrus. There were no significant effects for RPE residuals in a whole-brain

cluster corrected analysis, but some significant effects of surprise that generally followed the same pattern as participant-derived surprise signals (Supplementary Figure 5).

We also tested if there was any activity related to participants' subjective value for ingredients, distinct from task-related expected values (i.e. participant-P(yes)) in the participant-behavior GLM. There were also no clusters of activity associated with subjective value ratings at a whole-brain cluster corrected threshold, or in ROI analyses in vmPFC and the striatum (t's(22) ≤ 0.65, P ≥ 0.6, d's ≤ 0.13). Full information about activation peaks for the participant- and model-behavior parametric GLMs are available in Supplementary Tables 2 and 3, respectively.



Supplementary Figure 7. Whole brain t-statistic maps from parametric models based on participant and model behavior. Top row shows correlates for participant **a.** reward prediction error (RPE) and **b.** surprise. Bottom row shows correlates for model **c.** RPE and **d**. surprise. RPE maps are thresholded at $P < 0.001$, uncorrected. Maps for surprise are thresholded at $P < 0.05$, FWE cluster corrected. Cluster extent threshold for each map is given by value of k.

Supplementary Table 1. fMRI activations for main effects of task conditions shown in Figure 3

Region (AAL2)	MNI Coordinates			Number of voxels	Peak *t*-value
	x	y	z		
Recipe > Independent, correct (k = 148)					
Left superior frontal gyrus, medial orbital	-10	44	-12	240	4.31
Positive > Negative Value, correct (k = 157)					
Right anterior cingulate	12	40	8	4134	7.60
Left superior frontal gyrus, medial	-10	54	14		7.54
Right superior frontal gyrus, medial orbital	4	34	-14		5.84
Right posterior orbital gyrus	24	26	-14	282	5.78
Right caudate nucleus	6	12	-4	172	4.66
Presentation 1 > 2 > 3 (k = 173)					
Left inferior frontal gyrus, opercularis	-42	4	28	8816	12.39
Left supplementary motor area	-6	22	46		11.67
Left inferior frontal gyrus, triangularis	-46	24	16		11.30
Left inferior temporal gyrus	-48	-52	-16	4672	11.83
Left fusiform gyrus	-22	-86	-12		8.76
Left fusiform gyrus	-36	-70	-12		8.59
Right fusiform gyrus	36	-58	-12	3287	10.21
Right inferior occipital gyrus	34	-88	-6		8.51
Right fusiform gyrus	36	-28	-28		6.24
Right middle frontal gyrus	50	32	22	578	8.27
Left middle occipital gyrus	-24	-68	36		7.59
Left superior parietal gyrus	-30	-62	66		4.14
Right inferior frontal gyrus, orbital	32	30	-10	621	6.91
Left thalamus	-8	-8	4	370	5.36
Correct > error (k = 210)					
Right pallidum	26	-6	-6	4394	8.87
Right superior temporal gyrus	56	-26	4	4394	6.39
Right middle temporal gyrus	62	-2	-18	4394	6.13
Right paracentral lobule	10	-26	68	1065	5.29
Left precentral lobule	-22	-22	60	1065	4.61
Left middle temporal gyrus	-46	2	-28	6667	8.37
Left hippocampus	-24	-10	-22	6667	7.70
Left putamen	-30	-2	2	6667	7.68

Right superior occipital gyrus	24	-90	30	1195	7.21
Left cuneus	-8	-94	22	1195	5.31
Left anterior cingulate	-2	28	-8	241	4.54
Right superior frontal gyrus, medial	6	50	0	241	3.79

All reported clusters were significant at the $P < 0.05$, FWE-corrected threshold at the cluster level after peak thresholding at $P < 0.001$. The critical cluster extent threshold for each contrast is given by the value of k.

Supplementary Table 2. fMRI activations from participant behavior-based GLM, from Figure 7 and supplementary figure 4

Region (AAL2)	MNI Coordinates			Number of voxels	Peak *t*-value
	x	y	z		
Participant-P(yes) (k = 167)					
Right superior frontal gyrus, medial	8	54	8	2076	5.77
Right superior frontal gyrus, medial	6	64	26		4.93
Left anterior cingulate	-10	36	-4		4.78
Confidence (k = 185)					
Left cuneus	-2	-90	28	3194	8.04
Right cuneus	18	-74	24		4.34
Left superior frontal gyrus, medial orbital	-6	42	-8	3870	7.13
Right superior frontal gyrus, medial	8	60	6		6.25
Right superior temporal gyrus	60	-36	8	5251	6.79
Right superior temporal gyrus	62	2	-12		6.00
Right Rolandic operculum	50	-20	22		5.90
Left superior temporal gyrus	-68	-28	10	3864	6.67
Left temporal pole, middle temporal gyrus	-48	16	-34		5.80
Left middle temporal gyrus	-56	-10	-18		5.78
Left precuneus	-2	-44	38	2929	6.33
Left supplementary motor area	-4	-18	62		5.88
Right precentral gyrus	16	-28	68		5.76
Right lingual gyrus	14	-70	-4	435	5.74
Surprise (k = 169)					
Left middle frontal gyrus	-34	56	12	8393	10.69
Left supplementary motor area	-2	20	52		8.22
Left caudate nucleus	-14	10	8		8.17
Right middle frontal gyrus	48	32	30	4875	10.14
Right superior frontal gyrus, dorsolateral	30	62	-4		9.93
Right middle frontal gyrus	44	12	48		7.44
Right cerebellum crus I	16	-82	-28	3798	9.65
Left cerebellum crus I	-32	-60	-34		9.20
Left cerebellum crus I	-10	-80	-26		7.90
Right putamen	30	12	-6	1692	6.98
Right pallidum	26	-8	-4		5.20
Right angular gyrus	36	-60	48	2064	8.15
Right inferior parietal gyrus	52	-42	54		6.42
Left inferior parietal gyrus	-34	-52	44	2267	7.18

Left superior parietal gyrus	-22	-72	54		5.40
Left supramarginal gyrus	-56	-54	28		4.95
Right middle temporal gyrus	62	-32	-8	867	6.11
Right inferior temporal gyrus	56	-56	-14		4.27
Right precuneus	8	-60	48	428	5.19

All reported clusters were significant at the $P < 0.05$, FWE-corrected threshold at the cluster level after peak thresholding at $P < 0.001$. The critical cluster extent threshold for each contrast is given by the value of k.

Supplementary Table 3. fMRI activations from model behavior-based GLM, from Figure 7 and supplementary figures 4 and 5

Region (AAL2)	MNI Coordinates			Number of voxels	Peak *t*-value
	x	y	z		
Model-P(yes) (k = 155)					
Left superior frontal gyrus, medial	-2	58	14	1739	5.39
Right superior frontal gyrus, medial orbital	8	64	-8		4.80
Model confidence (k = 163)					
Right middle temporal gyrus	46	-38	0	3374	8.71
Right middle temporal gyrus	60	-14	-10		6.06
Right supramarginal gyrus	62	-40	42		5.82
Left cuneus	-4	-92	28	2592	8.18
Left cuneus	-16	-70	22		6.08
Left superior frontal gyrus	-64	-30	18	3107	7.31
Left superior frontal gyrus	-56	4	-14		6.00
Right anterior cingulate	16	48	14	1659	6.28
Right postcentral gyrus	24	-40	64	644	6.35
Right precentral gyrus	18	-18	68		4.45
Right lingual gyrus	14	-66	-6	347	5.95
Right lingual gyrus	16	-46	-6		3.64
Right putamen	28	-2	-4	241	5.17
Confidence residuals (k = 158)					
Left cuneus	-2	-88	26	2558	8.82
Right lingual gyrus	14	-70	-2		6.57
Left anterior cingulate	-6	36	-6	3109	8.32
Left superior frontal gyrus, medial	-4	62	10		6.57
Left superior frontal gyrus, medial	-10	48	42		5.87
Right precentral gyrus	14	-28	66	526	6.25
Left paracentral lobule	-14	-30	74		4.36
Left middle temporal gyrus	-64	-6	-18	647	6.13
Left middle temporal gyrus	-64	-26	-8		5.64
Left supramarginal gyrus	-60	-24	14		4.55
Left precuneus	-4	-44	38	252	6.08
Right precuneus	4	-56	24		3.87
Right superior temporal gyrus	54	-32	16	615	5.82
Right middle temporal gyrus	66	-48	12		3.81
Left angular gyrus	-50	-68	30	407	5.56
Right middle temporal gyrus	-34	-58	18		3.53

Right Rolandic operculum	64	0	12	733	4.11
Right precentral gyrus	48	-14	54	173	5.09
Model surprise (k = 152)					
Right middle frontal gyrus	46	30	32	3304	11.06
Right superior frontal gyrus, dorsolateral	28	56	-10		10.20
Right superior frontal gyrus, dorsolateral	30	60	12		7.98
Left middle frontal gyrus	-34	56	14	5187	10.06
Left lateral orbital gyrus	-38	48	-14		8.18
Left precentral gyrus	-42	6	48		8.03
Right angular gyrus	36	-62	50	1591	9.01
Right inferior parietal gyrus	50	-44	56		6.70
Right cerebellum crus I	14	-80	-30	3848	8.94
Left cerebellum crus I	-10	-82	-26		8.12
Left putamen	-18	8	10	1556	7.53
Right putamen	24	10	-8		6.11
Left inferior parietal gyrus	-34	-58	50	2809	7.13
Right precuneus	8	-60	46		6.34
Right inferior temporal gyrus	60	-40	-16	413	5.92
Right superior temporal gyrus	52	-26	-4		3.67
Surprise residuals (k = 152)					
Left insula	-38	16	-10	641	7.970
Left inferior frontal gyrus, triangularis	-56	16			5.71
Right insula	38	16	-10	492	7.61
Left cerebellum crus I	-18	-78	-28	715	7.40
Left cerebellum crus I	-34	-58	-32		7.30
Left middle frontal gyrus	-32	54	8	839	7.28
Right angular gyrus	32	-60	40	1304	7.12
Right inferior parietal gyrus	56	-40	50		5.32
Right angular gyrus	54	-58	36		4.48
Right middle cingulate	10	30	32	672	6.86
Left supplementary motor area	-10	22	64		5.44
Right superior frontal gyrus, dorsolateral	20	52	28	2285	6.73
Right superior frontal gyrus, dorsolateral	26	58	-10		6.22
Right middle frontal gyrus	50	30	32		6.13
Left inferior parietal gyrus	-42	-54	54	699	6.52
Left middle occipital gyrus	-26	-62	38		4.91
Left angular gyrus	-52	-58	36		4.50
Right middle temporal gyrus	58	-38	-10	205	4.87

All reported clusters were significant at the $P < 0.05$, FWE-corrected threshold at the cluster level after peak thresholding at $P < 0.001$. The critical cluster extent threshold for each contrast is given by the value of k.

Supplementary Table 4. fMRI activations for model-based values GLM, from Figure 9

Region (AAL2)	MNI Coordinates			Number of voxels	Peak *t*-value
	x	y	z		
Unfamiliarity (k = 164)					
Left fusiform gyrus	-36	-44	-20	5360	11.10
Left inferior occipital gyrus	-28	-84	-10		8.81
Left parahippocampal gyrus	-28	-26	-24		7.25
Right fusiform gyrus	36	-56	-12	4216	11.08
Right lingual gyrus	24	-88	-6		9.38
Right fusiform gyrus	32	-38	-22		7.73
Left inferior frontal gyrus, orbitalis	-40	30	-12	3795	10.97
Left inferior frontal gyrus, opercularis	-40	4	28		9.16
Left inferior frontal gyrus, triangularis	-42	24	14		8.56
Left supplementary motor area	-8	24	48	2195	9.29
Left superior frontal gyrus, medial	-8	52	44		7.15
Left superior frontal gyrus, medial	-6	50	24		5.37
Right anterior orbital gyrus	38	40	-18	196	7.37
Left middle occipital gyrus	-24	-66	36	625	6.79
Left gyrus rectus	-4	48	-18	465	6.61
Right gyrus rectus	2	28	-20		3.94
Schema value x participant-P(yes) (k = 175)					
Right Rolandic operculum	52	-24	20	581	6.83
Right supramarginal gyrus	64	-28	38	581	4.27
Right superior temporal gyrus	68	-12	8	581	4.15
Right hippocampus	38	-18	-22	266	6.69
Right inferior temporal gyrus	54	-12	-36	266	3.79
Right inferior frontal gyrus, triangularis	54	34	10	383	6.48
Left paracentral lobule	-16	-12	66	176	6.42
Right middle temporal gyrus	66	-52	-2	616	6.06
Left superior temporal gyrus	-64	-30	20	483	5.88
Left postcentral gyrus	-40	-36	64	593	5.69
Left postcentral gyrus	-20	-36	74	593	5.22
Left paracentral lobule	-6	-18	68	593	3.60
Right gyrus rectus	8	44	-22	330	5.37
Right anterior cingulate	4	38	-2	330	4.42
Left lingual gyrus	-8	-68	-4	196	4.75
Right postcentral gyrus	30	-40	64	282	4.72
Left superior temporal gyrus	-60	2	-12	212	4.42
Right lingual gyrus	8	-60	4	221	4.39
Right calcarine fissure	16	-78	10	221	3.84

Episodic value x participant-P(yes) (k = 194)					
Left superior frontal gyrus, medial	-4	60	12	1439	6.92
Right superior frontal gyrus, medial	14	52	18	1439	5.08
Left gyrus rectus	-8	38	-16	1439	4.88
Left middle occipital gyrus	-34	-86	36	477	4.57
Left superior frontal gyrus	-20	32	36	201	5.01

All reported clusters were significant at the $P < 0.05$, FWE-corrected threshold at the cluster level after peak thresholding at $P < 0.001$. The critical cluster extent threshold for each contrast is given by the value of k.

Supplementary Table 5. Whole brain seed-to-voxel connectivity effects from Figures 11 and 12

Region (AAL2)	MNI Coordinates			Number of voxels	Peak *t*-value
	x	y	z		
Recipe Positive > Recipe negative, presentation 1, correct — MTC seed (k = 96)					
Left superior frontal gyrus, medial orbital	-2	46	-12	351	6.42
Left middle temporal gyrus	-52	-68	12	112	5.30
Correct > Error, independent presentations 2 and 3, correct — HPC seed (k = 83)					
Left middle frontal gyrus	-40	52	20	253	6.14
Left superior frontal gyrus, dorsolateral	-18	54	-4	253	5.66
Right insula	36	18	0	100	5.76
Right superior frontal gyrus, dorsolateral	30	54	14	122	5.76

All reported clusters were significant at the $P < 0.05$, FWE-corrected threshold at the cluster level after peak thresholding at $P < 0.001$. The critical cluster extent threshold for each contrast is given by the value of k.

Supplementary Table 6. Comparison of four parameter models with three parameter base model.

Model	Added parameters	Parameter range	Parameter mean (SD)	Mean difference in AIC from base model (SD)	AIC t-test (*P*-value)	Participants with lower AIC (N)
Base model	*d* (decay)	0-1	0.01 (0.01)			
	β_e (episodic value temperature)	0-80	5.08 (0.23)			
	β_s (schema value temperature)	0-80	2.48 (1.49)			
Log (recipe dimensions)	*c* (number of eigenvalues to retain – log-transformed)	2-58 (linear space)	2.85 (0.89)	-10.17 (9.65)	0.003	15
Asymmetric schemas	β_{sn} (separate negative schema value temperature)	0-80	1.91 (1.39)	-9.38 (16.32)	0.006	13
Schema similarity weighting	β_r (schema similarity temperature)	-15-15	0.21 (0.46)	-8.89 (15.56)	0.006	10
Unfamiliarity bias	β_m (unfamiliarity bias)	-5-5	0.005 (0.30)	-11.08 (19.93)	0.007	13
Fixed bias	β_0 (intercept term)	-5-5	0.09 (0.22)	-11.16 (20.36)	0.008	13
Schema discounting model	*k* (exponential discounting factor for schema values)	0-80	4.12 (7.93)	-5.87 (14.57)	0.03	12
Recall model	*Pr* (fixed probability of recall for cached episodic values)	0-1	0.07 (0.10)	1.17 (1.68)	0.99	3

Supplementary Table 7. Comparison of five parameter models to preferred four parameter model

Model	Added parameters	Parameter range	Parameter mean (SD)	Difference in AIC from best four parameter model (mean (SD))	AIC t-test (*P*-value)	Participants with lower AIC (N)
Schema similarity weighting*	β_r (schema similarity temperature)	-15-15	0.13 (0.41)	-7.51 (11.04)	0.002	15
Fixed bias*	β_0 (intercept term)	-5-5	0.08 (0.21)	-6.64 (12.29)	0.008	16
Unfamiliarity bias*	β_m (unfamiliarity bias)	-5-5	-0.006 (0.29)	-6.27 (12.19)	0.01	17
Asymmetric schemas	β_{sn} (separate negative schema value temperature)	0-80	1.57 (0.97)	-4.54 (12.27)	0.04	11
Schema discounting model	k (exponential discounting factor for schema values)	0-80	4.25 (9.13)	-1.57 (5.85)	0.5	12

Supplementary Table 8. Comparison of six parameter models adding to preferred five parameter model

Model	Added parameters	Parameter range	Parameter mean (SD)	Difference in AIC from best five parameter model (mean (SD))	AIC t-test (*P*-value)	Participants with lower AIC (N)
Unfamiliarity bias*	β_m (unfamiliarity bias)	-5-5	-0.33 (0.63)	-9.31 (17.48)	0.009	17
Fixed bias	β_0 (intercept term)	-5-5	-0.02 (0.38)	-2.53 (6.31)	0.03	13
Asymmetric schemas	β_{sn} (separate negative schema value temperature)	0-80	1.51 (1.03)	-0.71 (4.16)	0.2	17

Supplementary Table 9. Comparison of seven parameter model adding to preferred six parameter model

Model	Added parameters	Parameter range	Parameter mean (SD)	Difference in AIC from best five parameter model (mean (SD))	AIC t-test (*P*-value)	Participants with lower AIC (N)
Memory bias*	β_m (memory strength bias)	-5-5	0.55 (0.83)	-6.58 (10.28)	0.003	14

Supplementary references

Barron, H. C., Dolan, R. J., & Behrens, T. E. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. *Nat Neurosci, 16*(10), 1492-1498. doi:10.1038/nn.3515

Dumais, S. T. (2004). Latent semantic analysis. *Annual Review of Information Science and Technology, 38*, 189-230.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25*(2-3), 259-284. doi:Doi 10.1080/01638539809545028

Yonelinas, A. P. (1994). Receiver-Operating Characteristics in Recognition Memory - Evidence for a Dual-Process Model. *Journal of Experimental Psychology-Learning Memory and Cognition, 20*(6), 1341-1354. doi:Doi 10.1037/0278-7393.20.6.1341