SUPPLEMENTARY INFORMATION:
Evidence for category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex

Daniel McNamee[1,2], Antonio Rangel[2,3], John P. O'Doherty[1,2,3]

[1] Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland

[2] Computation and Neural Systems, California Institute of Technology, Pasadena, USA

[3] Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, USA

Contact:

Daniel McNamee

Caltech, MC 136-93,

Pasadena, CA 91125, USA

E-mail: dmcnamee@caltech.edu

CONTENTS

SUPPLEMENTARY FIGURES

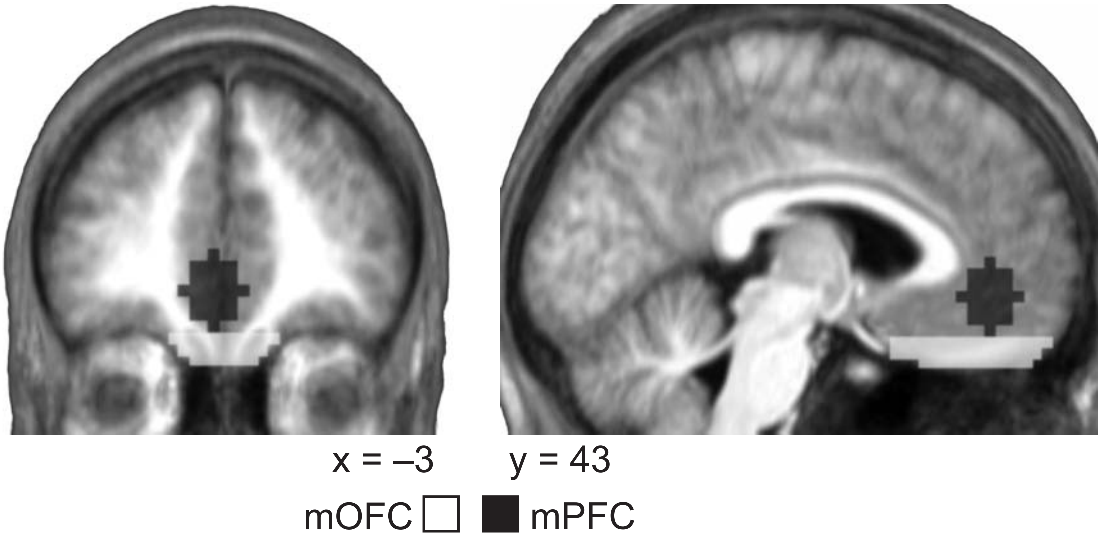

Figure S1. Masks Covering Distinct Subregions of vmPFC. Based on previous functional and anatomical results, our a priori hypothesis was that distributed and univariate value encoding signals would be found in vmPFC extending from the orbital surface to more dorsal regions up to and including parts of Brodmann areas 10 and 32. Due to the similarity of the experimental design, we used univariate peak coordinates from a related study (Chib et al., 2009) to construct a medial prefrontal cortex (mPFC) mask as a sphere of radius of 9mm surrounding these peak coordinates (corresponding to the size of the multivariate searchlight sphere). A similar functional mask did not exist for the medial orbitofrontal cortex (mOFC); most likely due to the distributed nature of the value codes found there and the relative scarcity of MVPA studies in value-based decision-making thus the mOFC mask was constructed according to anatomical descriptions used previously in the literature (Beckmann et al., 2009). This mask encompassed the medial orbital and olfactory sulci bilaterally with the anterior and posterior limits defined by the extents of these sulci. The vmPFC mask was defined as a union of these two masks.

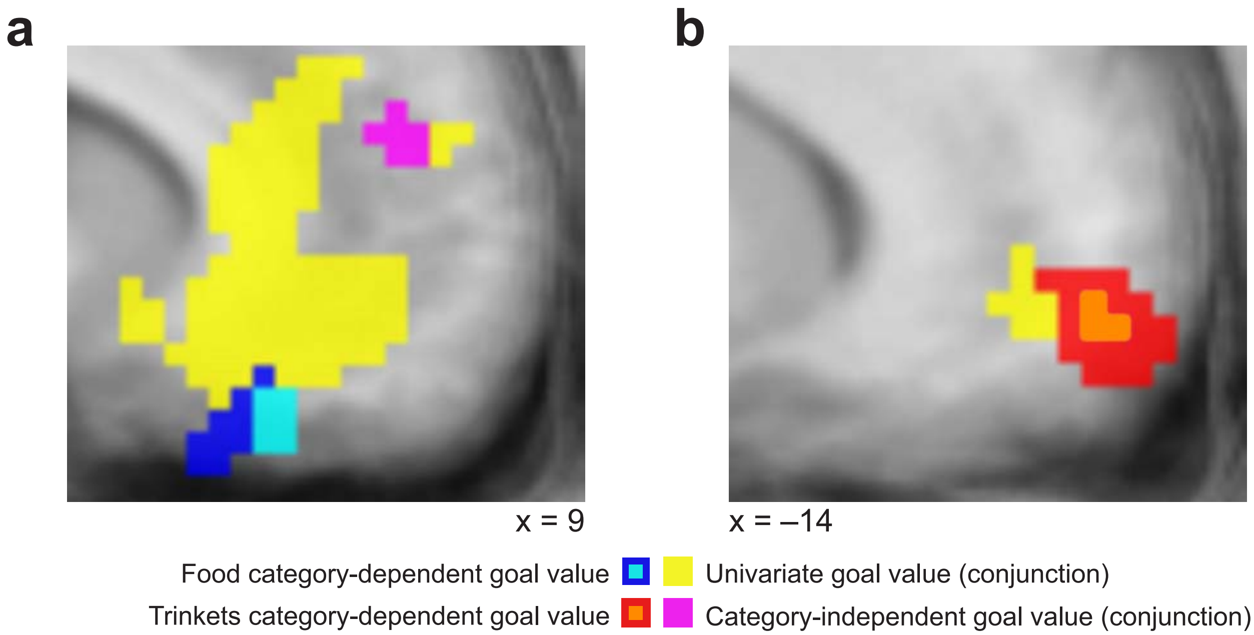

Figure S2. Independent Replication of Main Results.
For an independent replication of our results, we applied our analysis procedures to the data acquired for a previous study (Chib et al., 2009) with a similar task paradigm but with some important differences. This study also used a BDM auction process to elicit the participants' willingness-to-pay (WTP) on an integer scale from $0 to $4 for a variety of items drawn from three categories (food, monetary sums, and “trinkets”). However, the WTP bids (that is, the goal values) for all the items were recorded before the participant entered the scanner. Subsequently, on each trial in the scanner, participants were required to make binary choices between an item and a fixed reference sum of money (the median bid over all items placed during the pre-scan behavioral experiment). The motor response performed was a left or right button press and was completely uncorrelated with both the choice and the value of the item since the item and the reference sum of money were randomly presented to the left and right of a fixation cross. Choosing the item meant that the participant paid the reference price in exchange for an 80% chance of receiving the item. If they chose the reference amount of money, they would neither pay anything nor have the opportunity to play the lottery.

The analyses in the original study indicated that the value of the lottery item on each trial was commonly represented (as a smoothed univariate BOLD response) in a dorsal portion of vmPFC for all three item categories. This value representation was interpreted as a “decision value” signal (as opposed to a goal value in the paradigm used in the current study) since it is being computed in order to make a binary decision choice. In light our results, we hypothesized that distributed value signals, both category-dependent and category-independent, would accompany this smoothed value signal in the ventral and dorsal portions of vmPFC respectively. More specifically, we expected to see an anterior/posterior dissociation in category-dependent value signals along the medial orbital surface, whereby food value would be located more posteriorly and trinkets more anteriorly. We performed the same value decoding analyses as described in the main text on this dataset (19 participants; 15 male; mean age, 23.7; age range, 18-47).

a, A food-category-dependent value representation was located in posterior mOFC (peak [x = 3, y = 33, z = –24], $t = 2.86$; $P < 0.05$ SVFWE, small volume familywise error corrected, within a 20mm-radius sphere centered on the peak food value coordinates from the study reported in the main text). A category-independent value signal (conjunction across training/testing on food/trinkets and trinkets/food respectively) was located in mPFC (peak [x = 6, y = 57, z = 12], $t = 1.98$, $P < 0.05$ uncorrected). **b**, A trinket category-dependent value representation was located in anterior mOFC (peak [x = –15, y = 57, z = –9], $t = 2.94$; $P < 0.05$ SVFWE within a 20mm-radius sphere centered on the peak trinket value coordinates from the study reported in the main text). No clusters were present in any unanticipated ROI (e.g. a trinket category-dependent value signal where food category-dependent signals were found in the primary dataset). Results are shown at $P < 0.005$ and $P < 0.05$ uncorrected and overlaid on an averaged structural image. These results provide a robust independent replication of the anterior vs. posterior vmPFC value coding effects observed in the main study, with the ventral vs. dorsal effect also evident (although the multivariate category-independent result in dorsal vmPFC was uncorrected for multiple comparisons).

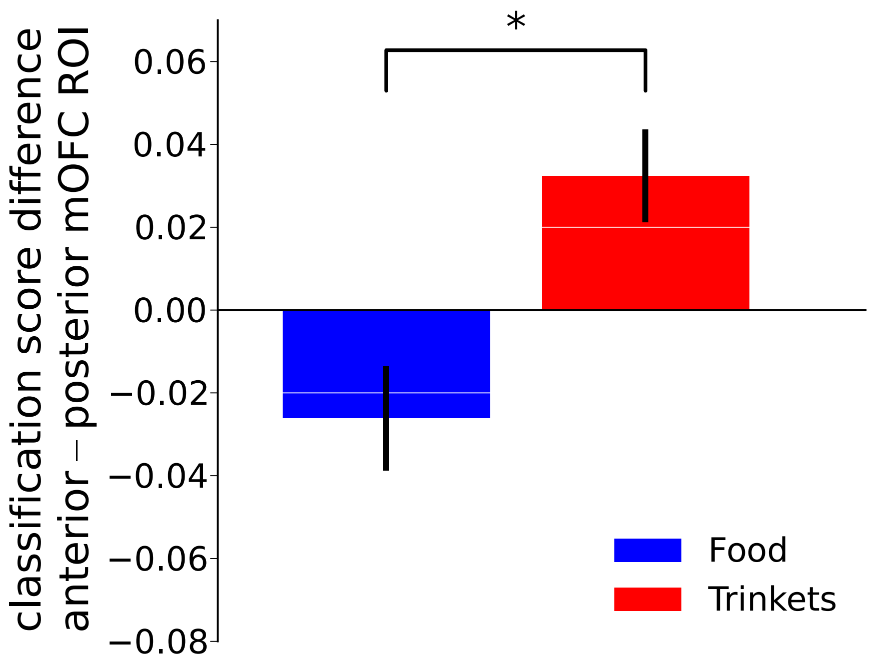

Figure S3. Leave-One-Participant-Out Anterior/Posterior mOFC Gradient Analysis. Here we replicate the anterior/posterior mOFC gradients identified in the main text in a completely non-circular manner using ANOVA interaction tests applied to per-subject classification scores derived using a leave-one-participant-out approach.

For each subject and item category, we first performed second-level mass-univariate *t*-tests on the classification maps for 12 participants only (leaving one participant out). The peak *t*-score coordinate within the mOFC ROI was identified and the classification score for the left-out participant at the peak coordinate was recorded. In addition, the peak *t*-score from the alternative item category analysis within a searchlight sphere of voxels (restricted to the mOFC ROI) surrounding that peak coordinate was also taken. For example, for each subject we recorded two food value classification scores: (1) one based on the peak coordinate in mOFC and (2) the other based on the peak coordinate within a searchlight sphere of the peak coordinate from the trinket value decoding. Similarly, two trinket value classification scores were also acquired for each subject. In this way, for each item category and subject, we independently derived a classification score and then also recorded a classification score for the alternative item category within the same locality. This process was repeated for every subject in both analyses being contrasted. The end result was a dataset comprised of four classification scores for each subject derived in a completely independent manner.

The data was entered into a repeated measures 2 × 2 ANOVA design (spatial location x item category) and there was a significant interaction between the two factors ($P = 0.039$) whereby the trinket-category-dependent value encoding signal was stronger in the region identified more anteriorly but not posteriorly and vice versa for the analogous food-related signal. This replicates the corresponding result in the main text (**Fig. 2b**) in a completely independent manner.

In this figure, the simple main effect of spatial location on classification score is plotted across item category i.e. the distribution of the relative differences in *t*-scores between the anterior and posterior ROIs (food items in blue, trinkets in red). Error bars reflect standard error of the mean.

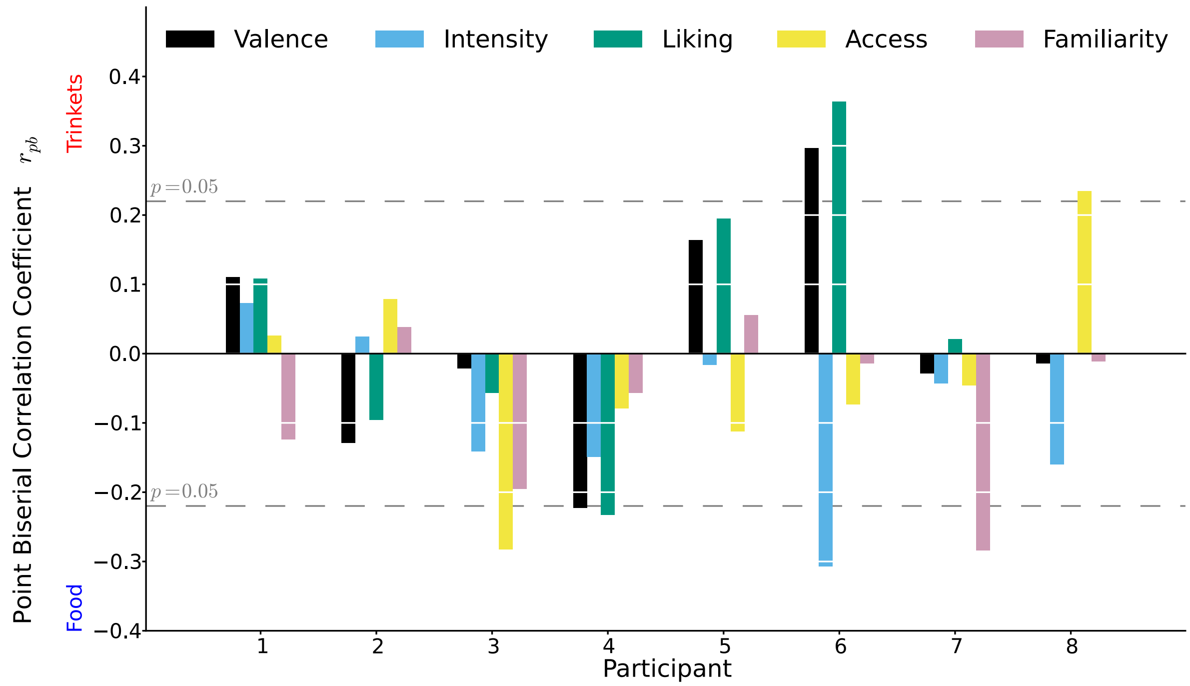

Figure S4. Item Ratings.

We acquired post-hoc behavioral ratings of the food and trinket items used from 9/13 of the original participants. One participant did not complete the questionnaire leaving 8/13 to be analyzed. The items were rated on five scales – “valence”, “intensity”, “liking”, “accessibility”, and “familiarity” from a score of 1 to 7. Items were presented in a random fashion across categories. Specifically, the questions were:

LIKING – how much do you like this item? A score of 1 means “I do not like this item at all”, a score of 4 means “I neither like nor dislike this item”, while a score of 7 means “I really like this item a lot”.

FAMILIARITY – how familiar are you with this item? A score of 1 means “This item is unknown to me”, a score of 4 means that “I am somewhat familiar with this item”, while a score of 7 means “I’m completely familiar with this item”.

INTENSITY – how intense are the feelings evoked by this item? A score of 1 indicates “This item evokes no feelings or emotion for me”, a score of 4 “I have some feelings towards this item”, while a score of 7 means “I have very intense feelings towards this item”. Note that for this question, it is irrelevant whether the feelings/emotions you have are positive or negative.

ACCESSIBILITY – how easy do you feel it is for you to obtain this item? A score of 1 means “It is almost impossible for me to get this item”, a score of 4 means “I can get this item without much difficulty”, while a score of 7 means “I would have no problem getting this item”.

VALENCE – how pleasant or unpleasant is this item? A score of 1 means “It is a very unpleasant item”, a score of 4 means “This item is neither pleasant nor unpleasant”, while a score of 7 means “This is a very pleasant item”.

The point-biserial correlation $\mathrm{r_{pb}}$ is the Pearson correlation between item ratings and the dichotomous variable indicating whether the item is a food item or a trinket. It describes to what extent higher or lower ratings are correlated with trinkets or food items. Positive correlations indicate that higher ratings correlate with trinkets; negative correlations indicate that higher ratings correlate with food items. A zero correlation implies that the ratings are evenly matched across items.

Results of statistical analyses can be seen in **Table S2, Table S3,** and **Table S4**. At $P > 0.05$, there was no significant difference between food and trinket items with respect to any rating (across subjects or across items). In two ratings (“intensity” and “familiarity”), there was a trend towards higher ratings in the food category. The subject-level point-biserial correlation showed that this was a weak effect within individual subjects with only one subject reaching a $P < 0.05$ significance threshold for each rating.

The bar chart in this figure reflects the point-biserial correlation coefficients r_{pb} for each subject between item ratings and a dichotomous variable which indicated whether the item was drawn from the food or trinket category. Repeated-measure statistical tests of any ratings difference between the food and trinkets category were not significant ($P > 0.05$). As can be seen from this figure, there is a high degree of variability within and across subjects in these ratings indicating that they are unlikely to account for the gradient effects reported in the main analyses.

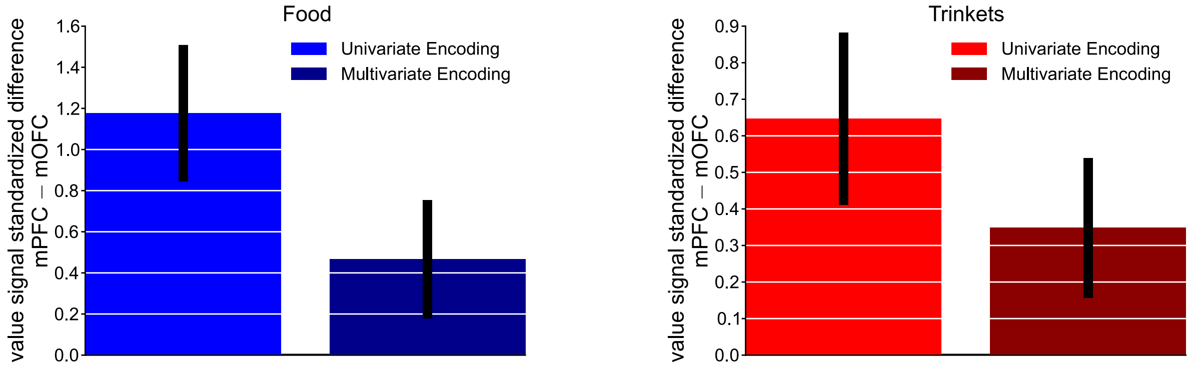

Figure S5. Leave-One-Participant-Out Ventral/Dorsal vmPFC Gradient Analysis. Here we replicate the ventral/dorsal vmPFC gradients identified in the main text in a completely non-circular manner using ANOVA interaction tests applied to per-subject value representation scores derived using a leave-one-participant-out approach.

For each analysis, we first performed second-level mass-univariate *t*-tests on the multivariate classification maps and general linear modeling beta maps for 12 participants only, leaving the 13th subject out. The peak *t*-score coordinate within each vmPFC ROI was identified and a value representation score (classification score for the multivariate analyses or first-level GLM *t*-score for the univariate analyses) for the left-out participant at the peak coordinate was recorded. This process was repeated for every subject in for both the food and trinket item categories. The end result was a dataset comprised of four classification scores for each subject derived in a completely independent manner.

Since we seek to compare results across encoding strategies, we standardized these results by computing the distribution of standardized value signal differences between the ventral and dorsal ROIs for each item category and encoding strategy. That is, we subtracted the mPFC scores from the mOFC scores and then divided by the standard deviation across both ROIs. This data is plotted in this figure. The data was then entered into a repeated measures 2 × 2 ANOVA design (spatial location × encoding strategy) and there was a significant interaction between the two factors for both categories (P = 0.0008 for food, P = 0.0005 for trinkets) whereby there was a greater drop in signal strength in mOFC compared to mPFC for univariate encoding as opposed to multivariate encoding. This replicates the corresponding result in the main text (**Fig. 4a**) in a completely independent manner.

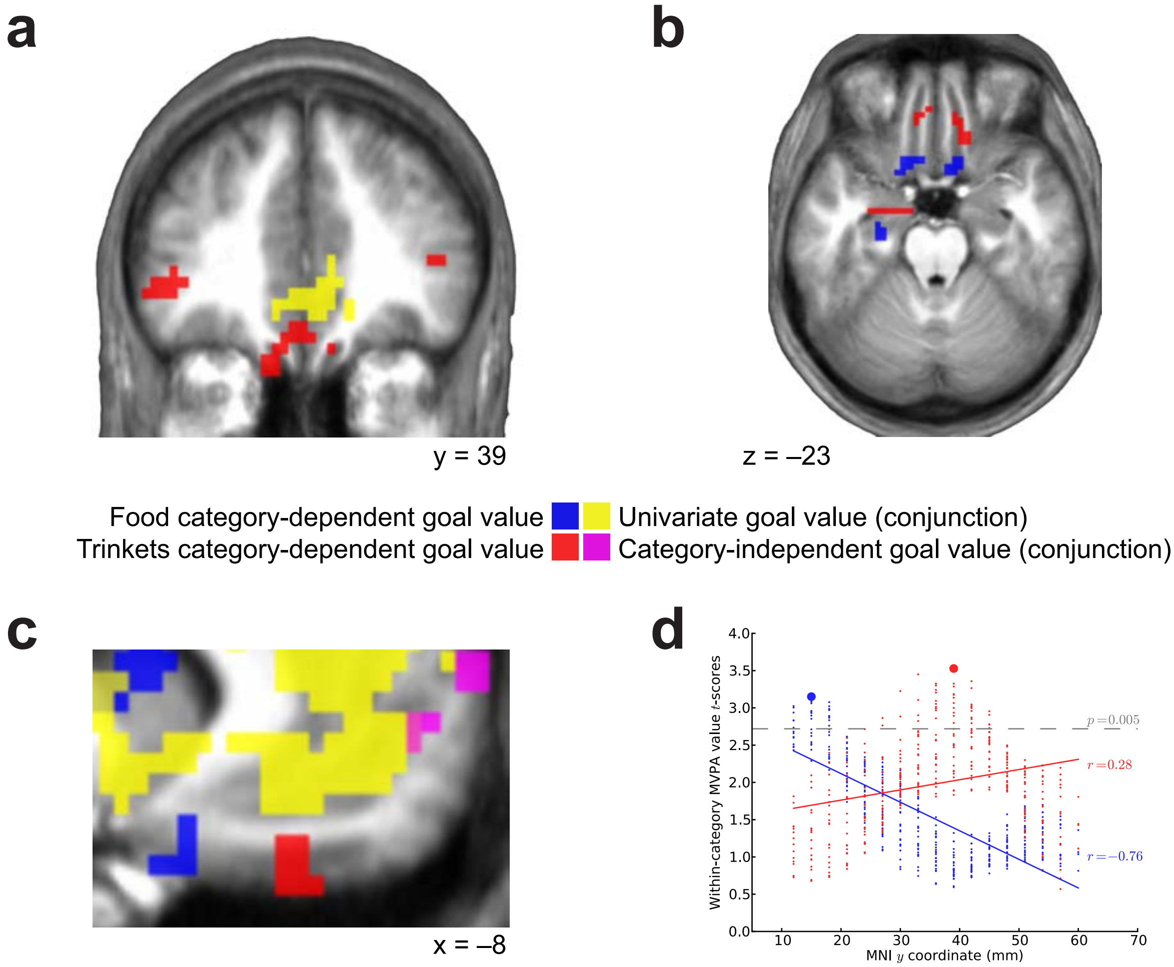

Figure S6. Value Decoding Based on "Mean-Subtraction" Searchlight.

We have used the terms "univariate" and "multivariate" to refer to signals identified using mass-univariate general linear modeling and MVPA (after orthogonalization with respect to the univariate signals) respectively. An alternative interpretation of "univariate" and "multivariate" signals in the context of a multivariate searchlight algorithm would be the signal identified using the mean and "mean-subtracted" activity respectively for each sample in the searchlight. The mean-subtracted activity is the voxel responses in a searchlight after subtracting the mean voxel response across the searchlight. We repeated the value decoding analyses using this alternative approach. This involved applying the same classification procedures as in the main text except with two crucial differences (1) the smoothed GLM-estimated goal value signal was not projected out and (2) the mean activity was subtracted and the variance across voxels normalized on a per-sample basis in every searchlight sphere (in effect, the neural pattern was standardized for every sample/sphere).

We repeated both the category-dependent and category-independent value decoding analyses in this manner. To ensure that this different methodology identified the same signals as previously in the main text, we tested for a significant activation (at $P < 0.05$ SVFDR, 10 voxel extent threshold) within ROIs defined as 20mm-radius spheres (Chib et al., 2009) surrounding peaks defined by the equivalent analyses. We also checked that no activations were unexpectedly present in an alternative ROI.

Significant clusters of voxels overlayed on an averaged brain template are presented in this figure. Food category-dependent goal value coding was observed in posterior mOFC (peak [x = 9, y = 14, z = –22], $t = 3.15$, $P < 0.005$ SVFDR) and trinket category-dependent goal value coding in anterior mOFC (peak [x = –3,y = 41,z = –11], $t = 4.20$, $P < 0.005$ SVFDR). Category-independent value representations (conjunction across pairwise cross-category analyses) were located more dorsally in mPFC (peak [x = –3, y = 47, z = –4], $t = 2.74$, $P < 0.05$ SVFDR). No results were "mismatched" between the two analysis methodologies. That is, no trinket category-dependent value representation was found in the food ROI and vice versa, and no category-dependent value decoding was present in the category-independent ROI and vice versa.

We also implemented an average signal searchlight whereby we attempted to decode cross-category value signals based on the mean activity within a searchlight sphere only. Out of six training/testing data combinations (e.g. train to decode value on monetary sums, test on food items), four resulted in a significant cluster in dorsal vmPFC at $P < 0.005$ uncorrected (10 voxel extent threshold) though they did not reach the corrected threshold $P < 0.05$ SVFDR. The ROI was defined as a 20mm-radius sphere surrounding the peak coordinates [x = –6, y = 53, z = –4] from the category-independent value decoding conjunction in the main text.

Each panel refers to an equivalent panel in the main text: **Fig. 2a (a, b)**, **Fig. 3a (c)**, **Fig. 2b (d)**. Results are displayed at $P < 0.005$ uncorrected and overlaid on an averaged structural image.

SUPPLEMENTARY TABLES

Table S1. fMRI Results

Category	Region	Hemi	x	y	z	t	p
Univariate Value Representation							
Food*	Medial prefrontal cortex	L	–3	38	–4	4.35	< 0.001
Money	Medial prefrontal cortex	L	–3	32	–7	3.24	0.001
Trinkets*	Medial prefrontal cortex		0	41	–7	4.35	< 0.001
Conjunction	Medial prefrontal cortex		0	35	–7	3.14	0.001
Distributed Category–Dependent Value Representation							
Food*	Medial orbitofrontal cortex	L	–9	17	–22	3.05	0.002
Trinkets*	Medial orbitofrontal cortex	L	–3	41	–11	3.86	< 0.001
Distributed Category–Independent Goal Value Representation							
Conjunction*†	Medial prefrontal cortex	L	–6	53	–4	2.88	0.002
Conjunction	Medial prefrontal cortex	L	–3	41	3	2.40	0.008
Distributed Goal Category Representation‡*							
Conjunction	Medial orbitofrontal cortex	L	–3	20	–22	6.12	< 0.001
Conjunction	Medial prefrontal cortex	L	9	29	0	7.56	< 0.001
Conjunction	Central orbitofrontal cortex	L	–21	38	–11	11.14	< 0.001
Conjunction	Frontopolar cortex	R	6	65	–11	6.89	< 0.001
Conjunction	Frontopolar cortex	L	–12	71	14	6.78	< 0.001
Conjunction	Dorsolateral prefrontal cortex	L	–60	17	14	11.34	< 0.001
Conjunction	Dorsolateral prefrontal cortex	R	45	32	21	5.84	< 0.001
Conjunction	Insula	R	45	5	3	5.27	< 0.001
Conjunction	Middle frontal gyrus	L	–33	5	35	6.84	< 0.001
Conjunction	Middle frontal gyrus	R	30	2	28	6.29	< 0.001
Conjunction	Middle frontal gyrus	L	–18	2	60	5.68	< 0.001
Conjunction	Anterior cingulate cortex	R	15	8	32	6.90	< 0.001
Conjunction	Intraparietal sulcus	L	–48	–31	42	11.65	< 0.001
Conjunction	Intraparietal sulcus	R	33	–70	42	7.94	< 0.001
Conjunction	Precuneus	L	–6	–64	14	5.54	< 0.001
Conjunction	Posterior cingulate cortex	R	3	–43	42	7.43	< 0.001
Conjunction	Parahippocampal gyrus	R	36	–10	–33	6.56	< 0.001
Conjunction	Inferior temporal gyrus	L	–45	–64	–22	7.64	< 0.001
Conjunction	Inferior temporal gyrus	R	30	–73	–15	7.36	< 0.001

Category	**Region**	**Hemi**	**x**	**y**	**z**	**t**	**p**
Distributed Goal Category Representation‡ (continued)*							
Conjunction	Fusiform	R	24	–43	–29	6.90	< 0.001
Conjunction	Fusiform	L	–27	–43	–22	7.18	< 0.001
Conjunction	Extrastriate cortex	L	–12	–79	32	11.34	< 0.001
Conjunction	Extrastriate cortex	R	63	–61	10	6.99	< 0.001
Conjunction	Extrastriate cortex	R	9	–70	–7	5.15	< 0.001
Conjunction	Striate cortex	L	–21	–79	14	9.67	< 0.001

Results thresholded at $P < 0.05$ FDR. Voxelwise FDR correction was performed within a ventromedial prefrontal ROI for all value-related results (i.e. SVFDR).

* Results which survive at $P < 0.005$ FDR or SVFDR.

† Conjunction across five binary category permutations (all except training value on money and decoding value on trinkets).

‡ Conjunction across all three binary category combinations.

Table S2. Item Ratings, Subject-Level Analysis

	P1	P2	P3	P4	P5	P6	P7	P8	Mean	SEM	p
Valence (r_{pb})	0.110	–0.129	–0.022	–0.223	0.163	0.296	–0.028	–0.014	0.019	0.058	0.750
Intensity (r_{pb})	0.073	0.025	–0.141	0.149	–0.016	–0.307	–0.043	–0.160	–0.090	0.043	0.077
Liking (r_{pb})	0.108	–0.095	–0.056	–0.232	0.195	0.364	0.021	0.000	0.038	0.065	0.579
Access (r_{pb})	0.026	0.079	–0.282	–0.079	–0.112	–0.073	–0.046	0.235	–0.032	0.053	0.572
Familiarity (r_{pb})	–0.124	0.038	–0.196	–0.056	0.056	–0.014	–0.284	–0.011	–0.074	0.042	0.122

Positive r_{pb} values indicate stronger positive rating correlations with the trinket category; negative r_{pb} values indicate stronger positive rating correlations with the food category. Grey background indicates significance at p = 0.05 (r_{pb} = $\pm$0.2199).

Table S3. Item Ratings, Distribution Per Category Across Subjects

	Food		Trinkets		Food > Trinkets
	Mean	**SEM**	**Mean**	**SEM**	**Repeated measures t-test**
Valence	4.450	0.199	4.456	0.107	p = 0.973, t = –0.035
Intensity	4.028	0.166	3.712	0.146	p = 0.059, t = 2.250
Liking	4.400	0.242	4.494	0.126	p = 0.692, t = –0.412
Access	5.309	0.265	5.225	0.270	p = 0.640, t = 0.489
Familiarity	5.534	0.214	5.278	0.207	p = 0.111, t = 1.822

Table S4. Item Ratings, Distribution Per Category Across Items

	Food		Trinkets		Food > Trinkets
	Mean	**SEM**	**Mean**	**SEM**	**Independent t–test**
Valence	4.450	0.126	4.456	0.109	p = 0.970, t = –0.037
Intensity	4.028	0.100	3.712	0.162	p = 0.101, t = 1.661
Liking	4.400	0.133	4.494	0.129	p = 0.614, t = –0.507
Access	5.309	0.141	5.225	0.134	p = 0.666, t = 0.440
Familiarity	5.534	0.144	5.278	0.190	p = 0.285, t = 1.077

Table S5. Items Used Organized By Category.

Food Items	Money Items	"Trinket" Items
Ambrosia	20c	1984, George Orwell (book)
Apple Pies	30c	A Brief History of Time, Stephen Hawkings (book)
Bombay Mix	40c	A Portrait of the Artist as a Young Man, J. Joyce (book)
Cashews	60c	Abbey Road, The Beatles (music CD)
Choco Chip Cookies	70c	Alarm Clock
Coco Pops	90c	Batteries
Cornflakes	€1.2	Blade Runner (movie DVD)
Cream Crackers	€1.3	Blank DVDs
Crunchies	€1.5	Bourne Ultimatum (movie DVD)
Digestives	€1.6	Calendar
Doritos	€1.8	Combination Lock
Elevenses	€1.9	Dracula (book)
Fig Rolls	€2	Family Guy Season 7 (TV series DVD)
Fingers	€2.1	Golden Compass, Philip Pullman (book)
Frosties	€2.2	Harry Potter (book)
Fruit Pastilles	€2.3	Indiana Jones (movie DVD)
Gherkins	€2.4	James Bond, Quantum of Solace (movie DVD)
Granola Bar	€2.5	Joshua Tree, U2 (music)
Spam	€2.6	Kings of Leon Live (music/movie DVD)
Jaffa Cakes	€2.7	Lord of the Rings (movie DVD)
Bacon Fries	€2.8	Monopoly (boardgame)
Liquorice All Sorts	€2.9	OK Computer, Radiohead (music CD)
Mikado	€3	Playing Cards
Mini Rolls	€3.1	Shampoo
Beetroot	€3.2	Sherlock Holmes (book)
Pineapple Rings	€3.3	Slumdog Millionaire (movie DVD)
French Fancies	€3.4	Socks
Pickled Onions	€3.5	The Departed
Green & Black Chocolate	€3.6	The Hitchhiker's Guide to the Galaxy (book)
Rice Krispie Squares	€3.7	Stapler
Riesen	€3.8	T-Shirt
Salted Peanuts	€3.9	The Dark Knight (movie DVD)
Sesame Sticks	€4.1	The Wire Season 4 (TV series DVD)
Pringle's Original	€4.2	Travel Plug Adaptor
Fox's Shortcakes	€4.3	Trinity College Key Chain
Tea Cakes	€4.4	Trinity College Mug
Terry's Orange	€4.6	Trinity College Sweatshirt
Pickled Eggs	€4.7	Umbrella
Walkers Crisps	€4.8	USB Key 2GB
Werthers Sweets	€4.9	Water Bottle

The items used were similar to those used in Chib et al, 2009, although they were customized to be familiar to participants in Ireland, where the study was performed. Our motivation for using these specific goods is that we wanted to include large varied groups of items that were approximately similar in their average values to participants so as to control for the effects of value per se when doing between category comparisons. In addition, we required that the items be “everyday” items to ensure that all the subjects would be similarly exposed to the items and thus would be able to reasonably evaluate them. All subjects reported a high degree of familiarity with each of the items in a post-scan verbal report. Monetary amounts were selected in the range 10c to 5 euros, in 10c increments. Item order was randomly determined at the beginning of each experiment. Forty items were used in each class.

Subjects were only allowed to bid using a discrete set of values (Chib et al., 2009, Plassmann et al., 2007) thus the bids recorded are an approximation to the true values for the items. A true WTP of €1.20 is measured as €1, if the subject values an item at €4.50, we record a value of €4. However, a correlation analysis reported in Plassmann et al., 2007 showed that this discretized WTP distribution strongly reflects how much a subject likes the items, and thus can be taken as a good approximation to their true subjective goal-valuations.

Table S6. Items Used In Chib et al., 2009 Organized By Category.

Food Items	Money Items	Trinket Items
Chocolate Chip Cookies	20c	300 (movie DVD)
Chocolate Pudding	40c	A Brief History of Time, S. Hawkings (book)
Cookies	60c	Batman Begins (movie DVD)
Doritos Chips	80c	Blade Runner (movie DVD)
Fig Rolls	$1	Bourne Ultimatum (movie DVD)
Ghiradelli Chocolate Bar	$1.2	Caltech Backpack
Hershey's Milk Chocolate Bar	$1.4	Caltech Cap
Ho-Ho Chocolate Cake Rolls	$1.6	Caltech Flag
Kit Kat Chocolate Bar	$1.8	Caltech Key Chain
Lindt Chocolate Bar	$2	Caltech Mug
Mrs Fields Cookies	$2.2	Caltech Travel Mug
Oreo Cookies	$2.4	Caltech Sack
M & Ms	$2.6	Caltech Straw Hat
Powdered Donuts	$2.8	Freakonomics, S. Levitt & S. Dubner (book)
Pringles Chips	$3	Indiana Jones Boxed Set (movie DVDs)
Reeses Peanut Butter Cups	$3.2	Stapler
Rice Krispie Squares	$3.4	Spiderman (movie DVD)
Skittles Sweets	$3.6	The Big Lebowski (movie DVD)
Sweetarts Hard Candy	$3.8	The World is Flat, T. Friedman (book)
Twix Chocolate Bar	$4	Transformers (movie DVD)