

Is the whole the sum of the parts? Neural correlates of consumer bundle valuation in humans

Logan Cross, Ryan Webb, John P. O'Doherty

January 23, 2025

Abstract

Previous research has investigated the neural representation of valuation for goods during value-based decision-making, identifying decision value coding in the ventromedial prefrontal cortex (vmPFC) and elsewhere. However, previous studies have focused only on single items, as opposed to goods composed of multiple items (i.e., bundles). In this study, we investigated how participants evaluate consumer items both individually and in bundles at behavioral and neural levels. Participants underwent a deep-fMRI scanning protocol in which they were scanned on three separate days while we elicited valuations for varying single and bundled items. Behaviorally, we find that bundle values are systematically discounted using a sub-additive function, compared to the sum of individual item values. At the neural level, we found that a distributed network of brain areas including the vmPFC but also other parts of prefrontal cortex, computes the value of a bundle with the same value code employed for individual item evaluation, suggesting that these general value regions contextually adapt within this hierarchy. Additionally, we tested various models of how this value code may move between levels in this hierarchy. Our findings suggest that the value representation undergoes a divisive normalization process that actively re-scales the code based on the distribution of values within the current context, as opposed to utilizing an absolute value code. This study provides novel insights into the neural mechanisms that govern multi-attribute decision-making and value normalization in consumer bundle evaluations.

Introduction

In daily life, humans must evaluate options that contain multiple components, such as a meal, a bundle of cable TV channels, an investment fund, or a vacation package. Because decisions over bundles necessitate trade-offs between their distinct components, bundles are the fundamental primitive of economic analysis of consumer choice (Mas-Collel). From a normative standpoint, bundle evaluation can be accomplished via a hierarchical process of valuing the bundle's constituent items and then aggregating them to value the bundle: the whole is the sum of its parts. This process is formalized in attribute integration theories of value, which propose that the value of a stimulus is constructed by assigning values to its attributes and then aggregating them (Bettman et al., 1998). In the human brain, there is accumulating evidence for partially distinct neural representations of individual features of a stimulus, which are proposed to be integrated to compute an overall stimulus value (Suzuki et al., 2017; O'Doherty et al., 2021). This evidence is consistent with a large literature establishing a value representation for the objects of choice in regions of frontal cortex (Bartra et al., 2013). However, to date, we do not know how the brain evaluates bundles of such objects. In the present study, we examine whether the attribute integration theory of value generalizes to valuing bundles, for which the components that need to be integrated simply represent a different level of abstraction: they are also choice objects. This raises fundamental questions regarding how the value of bundles of items are represented in the brain, and how these representations influence human choice behaviour.

Neuro-imaging studies have identified specialized cortical areas in lateral orbitofrontal cortex (lOFC) that encode stimulus attributes and exhibit functional connectivity to medial orbitofrontal cortex (mOFC) and ventromedial prefrontal cortex (vmPFC) for value integration (Suzuki et al., 2017; Lim et al., 2013). Furthermore, distinct value representations for categories of goods, such as food and noncomestible consumer items, are separately encoded in posterior regions of the medial orbitofrontal cortex and more anteriorly in the vmPFC, respectively, along with an integrated value signal more dorsally within the vmPFC consistent with a form of "common currency" across categories (McNamee et al., 2013; Clithero and Rangel, 2014). Thus, if the individual items in a bundle are considered constituent features that are combined to compute an integrated bundle value, a similar integrative process can potentially be found for bundles. We might therefore expect to find that values of items and bundles exhibit a hierarchical representation over cortical areas, perhaps continuing the previously-noted lateral-to-medial or ventral-dorsal dissociations. Alternatively, the same region may code for the value of items and bundles, suggesting a more generalized valuation process across different decision contexts. We seek to answer whether there are non-overlapping value representations for single items distinct from bundles, or whether there is a shared representation used in both contexts. We test these hypotheses by establishing whether spatially distinct brain regions separately correlate with the values of individual

items and bundles.

A related question concerns how single-item and bundle value codes are scaled, particularly if items and bundles are evaluated in the same region. Recent behavioral and neuro-physiological evidence suggests that valuation regions in the primate brain implement a relative value code that adapts to the temporal and spatial context of the choice set (Padoa-Schioppa, 2009; Louie et al., 2015; Rustichini et al., 2017). This compression of the neural code is necessary because neural activity is bounded above and below by physiological limits, and must therefore be scaled to fall within these limits in a manner that maintains sufficient resolution for accurate discrimination. For instance, Cox and Kable (2014) find that the BOLD signal in the brain’s valuation network adapts to the distribution of item stimuli across wide and narrow value conditions. Behavioral evidence also suggests that a similar re-scaling takes place for attributes (relative to other alternatives) before they are aggregated in the valuation of an item (Dumbalska et al., 2020). However, all of these studies of adaptation have examined this code in the context of valuing single items. When items are combined together, does the neural code simply scale linearly to accommodate the expanded range of possible valuations from the linear combination of items? Or does the neural code adapt across bundle and single item contexts, re-scaling each separately?

Given the prior evidence for context-dependent value scaling, we hypothesize that value regions switch between single item and bundle contexts by re-scaling the value code according to the distribution of values in the current context, rather than encoding the absolute value of an option using the same scale across single item and bundle trials. This context-specific normalization would allow for optimal encoding and discrimination of values within each context, aligning with the brain’s tendency to optimize its resources. Context-dependent normalization would also support the remarkable behavioral flexibility observed in human decision-making, allowing individuals to tailor their value representations to the specific demands of the current decision problem, even when choices from multiple contexts are presented in rapid succession. Divisive normalization, a canonical cortical computation for encoding relative information (Carandini and Heeger, 2012), offers one biologically plausible mechanism for value re-scaling across contexts. We therefore compare theoretical normalization codes (including divisive normalization) to the neural data in order to shed light on the computational mechanisms underlying the contextual adaptation of the value signal when items are combined together into a bundle.

In our study, participants valued single food items, non-comestible consumer goods, and bundles of these same items using an incentive compatible mechanism. Participants then made choices between these items or bundles and a reference monetary amount in randomly interleaved trials while being scanned with functional MRI (fMRI). In order to address these questions and cater for variation across participants in the nature of the value codes, we adopted an fMRI experimental design that was optimized for multivariate representational analyses at the level of individual subjects. To maximize our capacity to probe neural representations of single and bundle values, we aimed to maximize

the number of trials of each type we could obtain per participant. To accomplish this, we implemented an elongated fMRI experimental design in which each participant performed the task on three separate days of scanning, allowing us to obtain over 900 separate trials in each participant, many more trials than would typically be found per participant in a conventional fMRI study design. We collected this extensive data in a total of 14 participants, which was also a sufficient sample size to enable us to perform group-level inference.

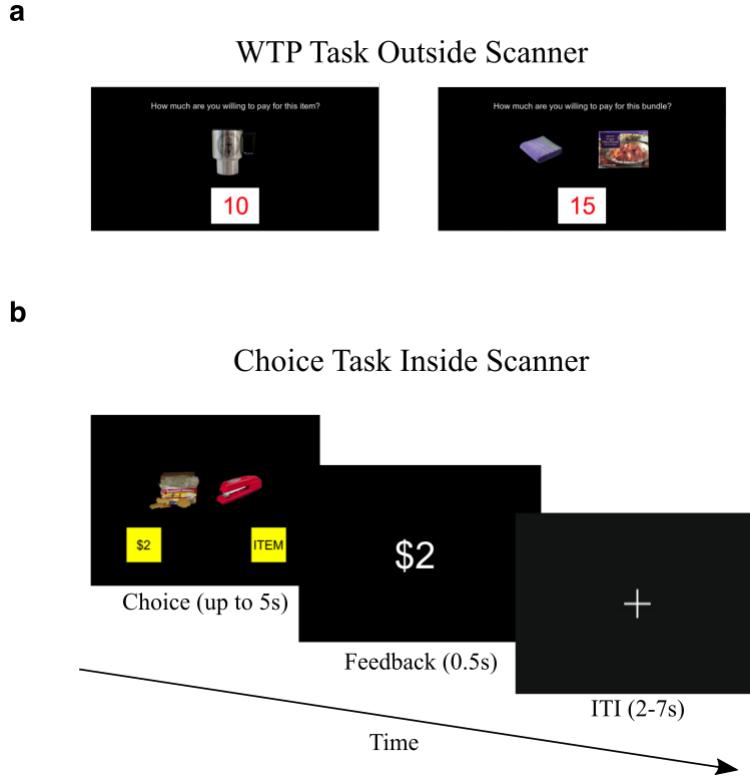


Figure 1: Experimental Design

- a.** WTP task. Participants reported how much they would be willing to pay for items and bundles of two items in a BDM auction. This was untimed outside the scanner.
- b.** Choice task. Inside the scanner, participants made choices with the items and bundles. During single item trials, a choice was made between an item and a reference monetary amount equal to the median bid of single item trials in the WTP task. Similarly during bundle trials, a choice was made between a bundle and a reference monetary amount equal to the median bid of WTP bundle trials. Participants had up to 5s to make a choice indicated by a right or left button press. The experiment involved 3 days of scanning 5 runs of the task for a total of 15 runs.

Results

We recruited participants into a 3 day experiment. On each day, a participant first reported how much they would be willing to pay (WTP) for various foods and non-comestible consumer goods (items) and pairs of these items (bundles) (Figure 1a). These WTP ratings represent an item's/bundle's subjective value and are recorded with an incentive

compatible procedure validated in previous studies (see Methods for details) (Becker et al., 1964; Chib et al., 2009; McNamee et al., 2013; Suzuki et al., 2017). Participants were allocated a budget of \$0-\$20 and submitted a wide distribution of WTP bids across categories, albeit with most items and bundles valued at small dollar amounts (Figure 2a; mean WTP value for each category: individual item food = $\$3.34 \pm 3.30$ s.d., individual item trinket = $\$3.30 \pm 3.50$ s.d., food bundle = $\$5.96 \pm 4.44$ s.d., trinket bundle = $\$4.92 \pm 4.81$ s.d., mixed bundle $\$5.394 \pm 4.65$ s.d.).

After the WTP task on each day, participants were scanned using functional MRI (fMRI) while they performed a choice task with the same items and bundles. Three participants were scanned with a high-resolution fMRI protocol (voxel size = 1.5mm isotropic) designed to record from medial prefrontal cortex (mPFC) regions with high fidelity (Figure S1). The remaining eleven participants were scanned with a standard wholebrain protocol (2.5 mm isotropic). On each trial in the choice task, a participant made a choice between an item (or a bundle) vs a reference monetary amount equal to their median WTP bid on that category (Figure 1b). This ensured that participants would value the item/bundle in isolation and choose the item/bundle about half the time. We observed a slight bias to choosing the money on individual item trials (Figure S2).

Bundle value is a sub-additive function of the individual item values

We first examined the values that participants assigned to a bundle related to the values of its constituent items. This relationship can be qualitatively visualized in the density plot in Figure 2b. As the sum of the individual item values increases, bundle value tends to also monotonically increase, as expected. However, there is substantial density below the diagonal, suggesting that bundle value is a sub-additive function of the individual item values.

To test this hypothesis, we modeled the value for a bundle, $v_{i,j}$, of items i and j as a function of the constituent item values. The theory of attribute integration specifies the value of a bundle as a linear combination of values of the individual items ($v_{i,j} = \beta_0 + \beta_1 v_i + \beta_2 v_j$). Using a linear mixed effects regression, we found that both parameter estimates for the individual items were significantly less than one ($\hat{\beta}_1=0.73$, $p < 10^{-10}$, $t=6.24$; $\hat{\beta}_2=0.73$, $p < 10^{-11}$, $t=6.72$; $\hat{\beta}_0=0.83$; $R^2=0.78$), consistent with the hypothesis that bundle value is a sub-additive function of the individual item values (Figure 2b). We additionally tested four non-linear mixed-effects models (see Methods). The power model and logarithmic model fit the data better than the linear model, while a degree-2 polynomial model did worse, controlling for the number of parameters (Table 1). The best-performing model was a divisive normalization model which normalizes the bundle components relative to each other (we will discuss this model further below). All of these nonlinear models are concave (Figure 7b), thus participants discount their subjective value of a bundle more as the individual item values increase.

Table 1: **Behavioral models of bundle value.** Each model fits random effects parameters per participant. BIC scores reflect fit across participants. See Methods for model details.

Model	BIC
Normalization	38996
Power	39049
Logarithmic	39115
Linear	39178
Polynomial	39230

We next analyzed how the relation between bundle and item value varies depending on the category of the bundle. The linear model fit to the data of each type of bundle is shown in Figure 2c. The slopes of this model are smallest for bundles that are duplicates of the same item ($\beta_1 = 0.61 \pm 0.01\text{sem}$, $\beta_2 = 0.61 \pm 0.01\text{sem}$, Intercept $\beta_0 = 0.98 \pm 0.13\text{sem}$), which is consistent with the economic principle that the utility of an item decreases with each additional unit (diminishing marginal utility). Moreover, food and mixed bundles have larger slopes than same item bundles and trinket bundles (Food: $\beta_1 = 0.74 \pm 0.04\text{sem}$, $\beta_2 = 0.77 \pm 0.04\text{sem}$, Intercept $\beta_0 = 1.07 \pm 0.26\text{sem}$; Mixed: $\beta_{Food} = 0.71 \pm 0.05\text{sem}$, $\beta_{Trinket} = 0.71 \pm 0.06\text{sem}$, Intercept $\beta_0 = 0.77 \pm 0.26\text{sem}$; Trinket: $\beta_1 = 0.67 \pm 0.06\text{sem}$, $\beta_2 = 0.63 \pm 0.06\text{sem}$, Intercept $\beta_0 = 0.70 \pm 0.19\text{sem}$).

Altogether, these behavioral results demonstrate that the value of a bundle is computed as a subadditive combination of the individual item values.

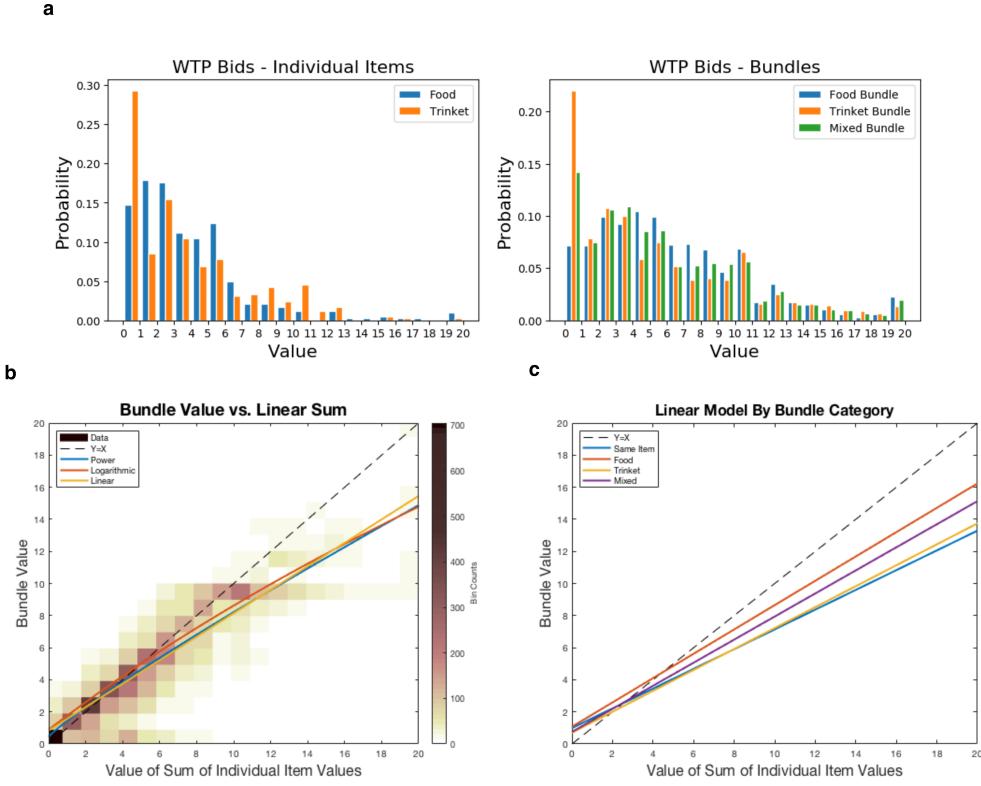


Figure 2: **WTP Behavior**

a. Histograms of WTP bids for individual items and bundles across all subjects.

b. Density plot of the WTP value of a bundle vs the sum of the values of the constituent items in a bundle in all subjects. If bundle value was equal to a linear addition of the constituent item values, bundle values would lie along the diagonal $Y=X$. Three models were constructed to predict the value of a bundle as a function of the individual item values: a linear model, a nonlinear power model, and a nonlinear logarithmic model. The fitted curves for each model are plotted along the density plot. All three models display that bundle value is a subadditive function of the individual item values, as they extend below the $Y=X$ line.

c. The fitted linear model stratified by bundle type.

Neural representation of subjective value and choice

We next investigated how subjective value and choice are encoded in the brain during the choice task. To do so, we first constructed a univariate general linear model (GLM) that included regressors for a stimulus's WTP rating time-locked to the onset of the trial, and a regressor for the choice made on that trial time-locked to the choice (in addition to a regressor for trial type and other covariates of no interest; see Methods). Several regions were more active when participants choose an item or bundle vs. when they choose the reference monetary amount, including clusters in dmPFC areas such as superior frontal gyrus (SFG), anterior cingulate gyrus (ACC), vmPFC, and the angular gyrus ($P < 0.001$ FDR corrected cluster-level, Figure 3a). By including the regressor for the WTP of the item/bundle shown in a trial in the same GLM, we were also able to isolate subjective value signals independent of choice. After cluster-level false discovery rate correction ($P < 0.001$, FDR), one large cluster in the anterior portion of vmPFC and frontal

pole showed a positive correlation to subjective value (Figure 3b). These results are consistent with previous results (Clithero and Rangel, 2014; Bartra et al., 2013), and set the stage for further probing how the representations in these regions are modulated trial type.

Next we aimed to test whether anatomically distinct brain regions represent the value of individual items and the value of bundles. One possibility is that additional regions are recruited when computing the value of bundles, since this involves a hierarchical process of valuing the individual items and then integrating them to evaluate the bundle.

In order to examine whether the brain recruits additional circuitry when evaluating a bundle versus valuing a single item, we tested for the interaction of value and trial type in the previously described GLM (contrasts: bundle value > single item value and single item value > bundle value). At the group level, no clusters survived correction for either comparison, thereby providing no evidence for a topography of separable single item value and bundle value codes at the univariate level in the current dataset (Figure S3).

We next aimed to determine whether the value of bundled items might be represented differently to single items at a more distributed level. For this we utilized multivariate pattern analysis (MVPA) in order to detect such distributed codes. We implemented decoding analyses across conditions to test for the existence of distinct bundle value codes, single item value codes, or general value regions. A ridge regression decoder was trained on distributed voxel patterns in several regions of interest (ROIs) in PFC. The decoder was trained on samples from one trial type (ie. single item trials) and then tested on both trial types separately on a held out run (cross validated with the leave-one-run-out method). Decoder performance was assessed with a Pearson correlation between the predicted values and the true values in the test set. To test for condition-independent general value regions, we analyzed if the decoder could predict the value of samples drawn from the opposite condition from that it was trained on. This would identify general value regions that utilize the same distributed code when computing the value of both single items and bundles. To test for condition-dependent value regions, we analyzed if the decoder could only predict samples within the condition it was trained on, while failing to predict samples from the opposite condition. This would identify distinct single item value regions or bundle value regions.

Value could be predicted above chance at the group level on all four types of train/test splits ('train and test on single items,' 'train and test on bundles,' 'train on single items/test on bundles,' 'train on bundles/test on single items') in vIPFC, dmPFC, dlPFC, MFG, and IFG (Figure 3c, two-sided one-sample Wilcoxon signed rank test $P < 0.05$ and FDR-corrected for multiple comparisons $q = 0.05$). By being able to decode value across conditions in these regions, this analysis suggests that they possess distributed general value codes that are independent of condition. Additionally, value was decoded above chance for the 'train on single items/test on bundles' partition in rACC, dACC, vmPFC, anterior OFC, lateral OFC, and medial OFC, providing some evidence for generalized value codes in these regions

as well (although prediction accuracies were not significant for the reverse partition ‘train on bundles/test on single items’). In order to depict how value decoding varies across individuals, individual subject results are also shown for two representative subjects in Figure 3c, and the remaining subjects in Supplementary Figure 4. To identify a condition-dependent value region, we assessed if a region’s decoder only had significant prediction accuracy when trained and tested on a single condition (ie. only significant for ‘train and test on single items’ but not for the other three partitions). No region met this criteria at the group level. Additionally, there were no significant differences in a decoder’s prediction accuracy between conditions for any ROI (two-sided two-sample Wilcoxon signed rank test $p < 0.05$ and FDR-corrected for multiple comparisons $q = 0.05$). For example, although there is a 0.06 difference in pearson correlation between ‘train on bundles/test on single items’ and ‘train and test on bundles’ for dACC, this difference is not significant after correcting for multiple comparisons ($p < 0.426$). Thus, the cross decoding analyses do not yield any evidence for the existence of a distributed item and/or bundle-specific value code in any ROI.

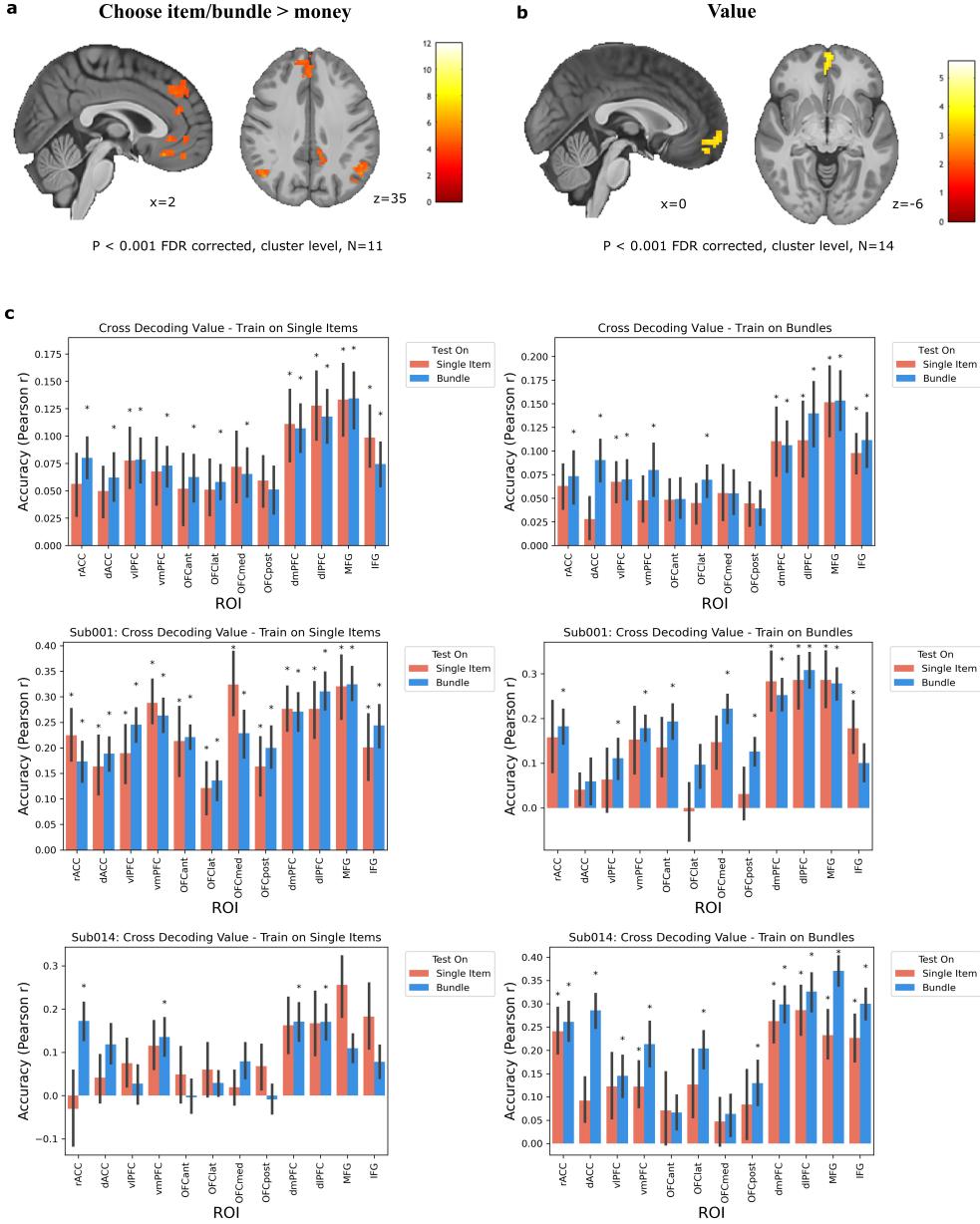


Figure 3: Neural representation of subjective value and choice.

- a.** Areas more active when the item or bundle was chosen than when the reference monetary amount was chosen. Significant clusters in dmpFC/SFG, ACC, vmPFC, and angular gyrus (N=11). Clusters are defined by a $p < 0.001$ uncorrected cluster forming threshold followed by $p < 0.001$ FDR cluster-level correction. The 3 participants that underwent the high resolution scanning protocol are absent from this analysis because that protocol was focused on ventral frontal areas (with a limited field of view) with no coverage in dorsal frontal areas.
- b.** Neural correlates of the value of the item/bundle presented. Significant cluster in anterior vmPFC (N=14). Clusters are defined by a $p < 0.001$ uncorrected cluster forming threshold followed by $p < 0.001$ FDR cluster-level correction.
- c.** MVPA cross decoding analysis results. Ridge regression decoders were trained on samples from a condition and tested on samples from a held out run in that conditions. Left: decoders trained on trials of single items. Right: decoders trained on bundle trials. Top row: Group result. Bottom rows are the results from the first and the last subject. Asterisks * represent significant prediction accuracies on a test partition (two-sided one-sample Wilcoxon signed rank test $p < 0.05$ and FDR-corrected for multiple comparisons $q = 0.05$). At the group level, there were no significant paired differences in prediction accuracies between test conditions for any ROI (two-sided two-sample Wilcoxon signed rank test $p < 0.05$ and FDR-corrected for multiple comparisons $q = 0.05$). Error bars reflect SE across participants. See Supplementary Figure 4 for the individual subject results for the other subjects.

Normalization of the value code

Thus far we have demonstrated that the brain's value regions represent value in a manner that generalizes across evaluating individual item values and bundles of those items. A subsequent question emerges from this pattern: how does the value code adaptively normalize to trial type? The common currency hypothesis suggests that options are encoded with the same value scale so they can be compared (Padoa-Schioppa and Assad, 2008; Chib et al., 2009; Lin et al., 2012; Levy and Glimcher, 2011), thereby value regions may code for items and bundles on the same range that scales with their WTP ratings. However, given the biological constraints of neurons, in order to evaluate decisions about which house to buy or which entree to buy with the same valuation system, the brain must also adaptively normalize the value code to the distribution of values available in a decision-making context (Louie et al., 2015). Since the distribution of bundle values is systematically larger than the distribution of single item values (Figure 7a), the neural code may normalize to the context of the current trial. A value region could therefore encode a \$4 rated item very differently than a \$4 rated bundle if options are appraised based on their utility relative to the other options within a condition (Figure 4a). Relative value coding has been observed empirically in experiments where a different distribution of values is presented in each block (Padoa-Schioppa, 2009; Louie et al., 2011), presumably because the value code adapts over time to the distribution in each block. But crucially in our experiment, bundle trials and single item trials are randomly intermixed within a block, therefore any adaptation must be instantaneous.

To test whether the neural representation codes for value in absolute or relative terms, we constructed GLMs with different value regressors. One GLM included a regressor for value according to the WTP bid in dollars for the item or bundle displayed in that trial, while in a separate GLM value was normalized (with a z-score) by trial type. Value contrasts were then compared at the second level (normalized value > absolute value and absolute value > normalized value). A significant cluster in vmPFC emerged for the normalized value > absolute value comparison (peak voxel: $x=-2, y=62, z=-7; t_{13}=5.75; P < 0.001$ uncorrected cluster forming threshold followed by $P < 0.001$ FDR cluster-level correction), suggesting that the neural representation of value in vmPFC is normalized. No significant clusters emerged after correction for the opposite analysis (absolute value > normalized value).

An alternative explanation for this result is that vmPFC is computing the difference in value between the item/bundle presented in a trial and the reference monetary amount it is choosing against. This computation would also produce a relative value code, but is a simpler form of adaptation that does not put the single item value and bundle value distributions on the same scale. Thus, another GLM was built with a modified value regressor equal to the WTP bid minus the reference monetary amount, and the resulting value contrast was tested against the fully normalized value contrast. Again, a significant cluster in vmPFC resulted from the normalized value > value difference contrast (peak voxel: $x=0, y=52, z=-14; t_{13}=4.83; P < 0.001$ uncorrected cluster forming threshold followed by $p < 0.005$ FDR cluster-level correction), and the opposite contrast yielded no significant clusters. These results demonstrate that at

least at the univariate level, vmPFC normalizes the value code by condition.

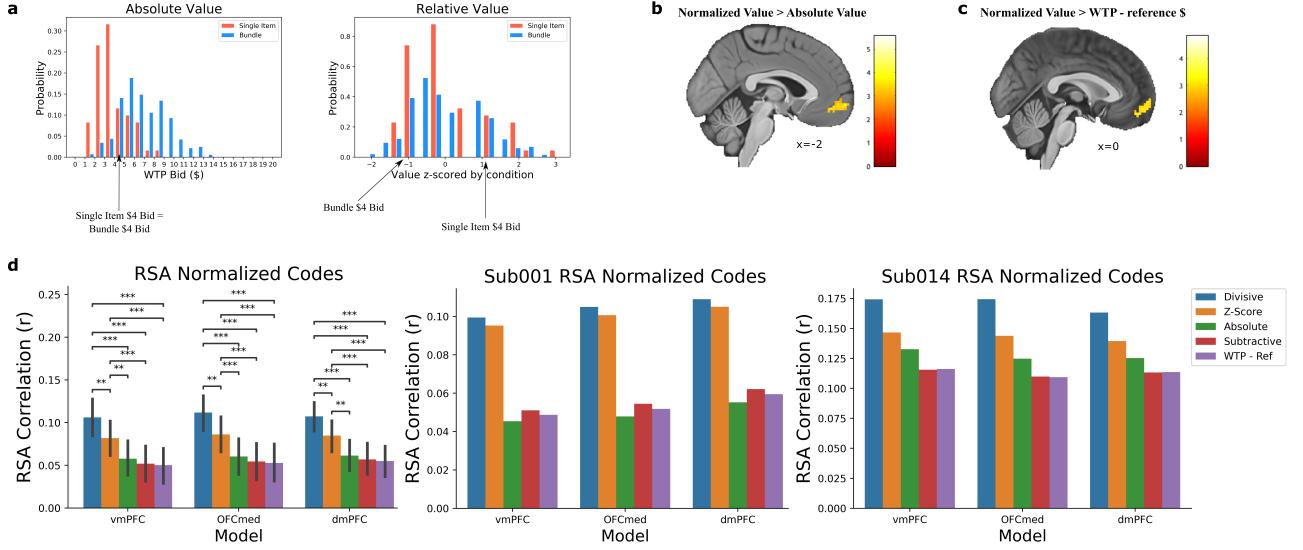


Figure 4: Normalization of the value code.

- a.** A depiction of the distributions of value by condition with absolute value and relative value codes. An absolute value code (left) represents value according to the participants' WTP bids, an incentive compatible measure of the subjective value of an item or bundle. A relative value code normalizes value by condition, plotted here after z-scoring individual item values and bundle values separately for an example subject (right). A relative code adapts to the context of the task and puts values from different distributions on the same scale.
- b.** Contrast of normalized value > absolute value. A cluster in vmPFC is better fit by the normalized value model, indicating that the value representation in this region is normalized. Clusters are defined by a $P < 0.001$ uncorrected cluster forming threshold followed by $P < 0.001$ FDR cluster-level correction.
- c.** Contrast of normalized value > value difference, with a cluster in vmPFC emerging similarly to the previous pane. Value difference is the WTP bid \$ amount - the reference \$ amount. $P < 0.005$ FDR cluster-level corrected ($P < 0.001$ uncorrected cluster forming threshold).
- d.** Representational Similarity Analysis (RSA) results comparing several representations of value, from absolute to value to different models of value normalization. Asterisks * represent significant differences between models (two-sided one-sample Wilcoxon signed rank test *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, all FDR-corrected for multiple comparisons). Left group result, middle and right two representative subjects (see Supplementary Figure X for remaining subjects).

To test different hypothetical value representations at the multivariate level, we performed representational similarity analysis (RSA). RSA provides a more data-driven method to examine the structure of neural representations by constructing dissimilarity matrices (DSMs) according to how the multidimensional voxel space changes from trial to trial or from condition to condition (Kriegeskorte et al., 2008). These data-driven DSMs can then be compared to model DSMs that encode how the features of the task, such as value, evolve from trial to trial. Here, trial by trial DSMs were built for every ROI and correlated to model DSMs for five potential value representations: absolute value, the WTP bid - the reference monetary amount, subtractive normalization (subtracting the mean value of a stimulus in a category), z-score normalization, range and divisive normalization. As shown in Figure 4d, the divisive normalized value DSM had significantly higher correlations to the neural DSMs than the other models in three main value ROIs at

the group level, vmPFC, OFCmed, and dmPFC (two-sided Wilcoxon signed rank tests, $P < 0.01$ and FDR-corrected for multiple comparisons ($q = 0.05$) except for range normalization which was not significantly different from divisive normalization. This suggests that the neural geometry in these value regions is most reflective of a normalized value code, consistent with either divisive or range normalization. The results from two representative participants are also shown in Figure 4 with the remaining participants displayed in Supplementary Figure X.

Discussion

In real-world decision-making, consumers often have to make choices between options that are each made up of multiple goods. However, it is unknown how the human brain constructs the value of bundles of multiple items. To investigate this question, we used a BDM auction procedure to elicit subjective values for food items, non-comestible consumer goods, and bundles of these items. Then participants made choices with these items and bundles during fMRI scanning.

WTP behavior could be predicted by modelling bundle value with a weighted integration model that uses the individual item values as attributes. We found that bundle value is computed with a sub-additive function of the individual item values, therefore bundles are systematically discounted in relationship to the sum of the individual item values. Concave non-linear models capture this behavior, and this concavity suggests that participants discount the subjective value of a bundle more as the values of its constituent items increase.

These results are compatible with the expectation most consumers have that bundles of multiple goods are usually discounted in comparison to the total price of the constituent products purchased separately (Thorne, 2004; Nagle and Holden, 1987). Value meals at fast food restaurants, snack variety packs, vacation packages, and season tickets for sports teams all represent discounted bundles in the real-world marketplace. However, the incentives of buyers and sellers in the marketplace are usually different than the one in our experimental setup. Sellers typically offer bundles at a discount to strategically price-discriminate to ensure that high-willingness to pay customers pay a higher price, as in discounting an item when it is purchased in bulk. Or alternatively, they increase the probability of a buyer purchasing an additional good they would not have purchased otherwise. For example, a customer at a fast-food restaurant will often have the option to buy a hamburger and fries separately, so a bundle of the two needs to be discounted in order to become an attractive package. In our experiment, each item and bundle is bid on separately and a BDM auction trial is selected at random, so when evaluating a bundle in the experiment, the option to purchase them separately is not available. Therefore, the results of our experiment suggests that bundle discounting may be a more general phenomenon of multi-item valuation and not just the result of market equilibrium prices. This pattern is consistent

with the law of diminishing marginal utility, which holds that an additional unit of a good has a lower value. In our experiment, bundles of the same item were discounted slightly more than the other bundles, which suggests that bundles of different items are discounted similarly to bulk quantities of the same item, but not as much.

Microeconomic theory also accounts for the possibility that bundles of goods are substitutes or complements (Varian, 2014). For example, tea and coffee could be viewed as substitutes as they both offer a caffeine boost. A consumer typically does not buy both at a cafe, so their utility bundled together is likely less than the sum of their independent utilities. Complements are products which are typically bought and used together, such as pasta and pasta sauce. Other examples, like left and right shoes, can exhibit superadditivity, where the utility of one shoe by itself is lower than half of the value of both shoes together. However, in economic theory, these concepts are typically defined in terms of a consumer's response to price changes, which are not present in our experiment. Moreover, the concepts of diminishing marginal utility and substitution/complementarity are not intrinsically linked theoretically. It is possible to define a utility function over goods which are perfect substitutes which also exhibits diminishing, constant, or increasing marginal utility. Though we did not construct our experiment to analyze substitution or complementary due to price changes, it is possible that the sub-additivity we observe in behaviour may arise from subject's evaluating a large majority of our stimuli as substitutes with diminishing marginal utility. Investigating how the brain evaluates and represents the value of bundles containing such items, and how their attributes are traded off against each other, is an interesting avenue for future brain research.

At the neural level, we tested for condition-dependent and condition-independent value signals at the univariate and multivariate levels. Previous research identified a spatial topography in OFC where food value was represented in posterior mOFC and value codes for consumer goods were represented in anterior mOFC (McNamee et al., 2013). In the present study we investigated whether a spatial topography would also exist in relation to the valuation of single items vs bundled items. One possibility that we tested is that bundle items, being more complex in the sense of being composed of multiple items as features, would be represented in a distinct region of the vmPFC compared to the value of single items. However, we found no evidence for separate single item value regions or bundle value regions or for a topography of value complexity. An anterior portion of vmPFC showed a correlation to subjective value across both individual item trials and bundle trials. In a cross-decoding analysis similar to methods previously used to identify category-dependent and category-independent value codes (McNamee et al., 2013), a decoder was trained on samples from one trial type and tested to predict value for the other trial type. Distributed condition-independent value codes were revealed throughout PFC, including in dmPFC, vIPFC, and MFG.

While we did not find evidence for any spatial topography with regard to the representation of the value of single and bundled items, we did find evidence to suggest that value is encoded in these regions in a relative as opposed to absolute manner. Specifically, we found that a form of normalization of value representation is occurring in these

areas, such that value representations are rescaled depending on whether the trial involves a single or a bundle item. Using representational similarity analysis, we compared neural activity against different forms of normalized and non-normalized value signals. Our findings suggest that value normalization provides the best account of the neural data compared to other forms of non-normalized value signals. Furthermore, our findings particularly favor one of two different forms of algorithmic normalization: divisive and range normalization. However, these two different implementations could not be further distinguished from each other in the current study. To differentiate a divisive normalization from range normalization scheme would have necessitated an experimental manipulation with three or more options available on every trial as opposed to the two options available in the current experiment (the target item and the reference bid). As a consequence, further investigation will be necessary to resolve the specific algorithmic form of normalization involved.

Our findings have implications for the common currency hypothesis of value coding (McNamee et al., 2013; Clithero and Rangel, 2014). Partly consistent with the notion of a common currency, we found evidence that the value regions we identified in our study encode value in a general way across both single and bundle items. However, a common currency would also require the value of every stimulus to be represented using the same invariant value scale. On the contrary, while our results show that value is represented by the same distributed voxel patterns across all trial types, this value code normalizes to the context and does not encode absolute value through a common currency representation. Thus, it is clear that our findings do not support a time and context invariant common currency representation. Instead, the brain appears to be utilizing a value representation that is scaled by context(Louie et al., 2011; Padoa-Schioppa and Assad, 2008).

To conclude, in the present study we show that the brain represents the value of bundles in the same brain areas that encode the value of single items. We found no evidence for a topographical arrangement of value representations within the ventromedial prefrontal cortex or elsewhere, when comparing single item and bundled item value signals, suggesting that these value signals are not hierarchically structured at least with respect to neuroanatomical organization. However, single and bundled items are represented differently with respect to value-scaling within those brain regions, such that the scale of the value representations appears to be specifically adjusted as a function of the single or bundle trial context. Our findings thus indicate that value representations for complex combinations of stimuli utilize overlapping neural representations that are scaled differently depending on the context.

Methods

Participants

Participants (N=14) were recruited from the general population through the Caltech Brain Research Participant System (7 females, 7 males, 24.9 ± 3.74 years, mean \pm s.d.). They did not have any food allergies and were not dieting at the time of the experiment. They were given a participation fee of \$40 (\$20 per hour), in addition to receiving monetary, food, and non-comestible consumer goods as rewards depending on their choices in the experiment. Each subject gave their informed consent, and the study was approved by the Institutional Review Board of the California Institute of Technology.

Stimuli

Across participants, 70 food items and 40 non-comestible consumer goods were used as stimuli in the experiment. Food items included fruits, snacks, and mains (including microwaveable meals) that are available at local grocery stores. Consumer goods included a diverse array of items under \$40 in price, including cell phone chargers, kitchen items, Caltech memorabilia, and books. Many of these items have been used in previous studies (Suzuki et al., 2017). The full list of items can be found in the Supplementary Table 1.

Experimental tasks

Participants performed the experiment in three sessions on three separate days in order to maximize the amount of fMRI data within-subject. On each day, participants first performed a willingness-to-pay (WTP) task outside the scanner, then performed the choice task in the scanner where they were asked to choose between an item or bundle of two items versus a reference monetary amount. Participants were asked to refrain from eating 4 hours before the experiment in order to ensure that food items were valuable to them. Compliance was confirmed by self-reports.

Each participant performed the willingness-to-pay and choice tasks on a total of 40 unique items presented throughout the experiment. On a given day, the participant was presented 20 items, composed of 10 food items and 10 non-food items. 10 of these items were presented in all three days, and 10 new items were introduced every day. To construct bundles, every item was paired with each other, including pairs of the same items. Thus, 210 bundles were included each day ($20 \text{ choose } 2 = 190 + 20$ pairs of the same item). On the final day of the experiment after the choice task, outside the scanner participants rated how familiar they were with each of the 40 items. For each item, the participant indicated their familiarity with the item on a continuous scale from ‘not at all’ to ‘very much’ by moving a red pointer,

with no time constraint as described previously (Suzuki et al., 2017).

WTP Task (outside the MRI scanner)

Participants completed an un-timed BDM auction task to measure their willingness-to-pay for items and bundles, with a procedure similar to previous studies (Chib et al., 2009; McNamee et al., 2013; Suzuki et al., 2017). The BDM auction is a reliable incentive-compatible method to elicit subjective values for items (Becker et al., 1964). Participants were endowed with a \$20 budget in cash, and instructed that they can use this cash to purchase items from our laboratory store (and keep the money they do not use). In each trial, an item or bundle was shown and the participant was asked to type in how much they would be willing to pay from \$0-\$20 for that item/bundle (Figure 3.1a). Participants first bid on the individual items (20 each day) and then bid on the bundles (210 each day).

Each trial was to be treated independently, as a random trial from the entire experiment was selected at the end of the experiment. If the selected trial was from the WTP task, the participants' bid on that trial was then compared against a randomly generated price (uniform probability from \$0-\$20), and if their bid was greater than or equal to that price, they received the item(s) and paid the corresponding price with their \$20 budget. If their bid was less than the price, they did not receive the item(s), and they did not have to pay anything. Participants were explicitly instructed about this auction procedure, and about how the optimal strategy is to bid their true subjective value for a given item/bundle. With a questionnaire, we confirmed that participants understood the mechanism of the auction.

Choice Task (inside the MRI scanner)

In the scanner, participants made choices involving the items and bundles previously bid on in the WTP task. Each trial involved a binary choice between an item or bundle and a reference monetary amount (Figure 3.1b). The reference monetary amount was equal to the participant's median WTP bid for that category (individual items were chosen against the median bid on individual items and bundles were chosen against the median bid on bundles). This ensured that the participants would choose the reference monetary amount about half the time and choose the item or bundle half the time. For a trial, the stimulus appeared in the middle of the screen, and the word 'ITEM' appeared on the bottom left or right with equal probability while the reference monetary amount in '\$X' appeared on the other side. Participants selected their choices with a button box, with the leftmost button indicating choosing the left option and vice versa for the rightmost button. Participants had 5 seconds to make a decision, after which their choice was presented on the screen for 0.5 seconds and followed by a jittered intertrial interval (ITI phase, 2–7s).

On each of the three days, participants were scanned for five runs. Each run included 62 trials, where each of the 20 individual items was presented once per run, and each of the 210 bundles in a day was presented once on that day. On day 1, anatomical scans were collected after the choice task.

fMRI Data Acquisition

The fMRI data was acquired on a Siemens Prisma 3T scanner at the Caltech Brain Imaging Center (Pasadena, CA) with a 32-channel radio frequency coil. At the end of the first day of scanning, T1 and T2 weighted anatomical high-resolution scans were collected with 0.9mm isotropic resolution.

High resolution data

A high-resolution partial volume slab was collected in three participants with a 1.5mm isotropic voxel size (Figure S1) and the following parameters: multiband acceleration = 4, 64 slices, TR = 1100ms, TE = 26ms, flip angle = 63°, FOV= 192mm x 192mm, in-plane GRAPPA ($R = 2$), echo spacing = 0.68ms. The protocol is optimized to view mPFC in high-resolution and therefore the partial volume slab cuts off portions of the motor cortex and parietal lobe. EPI-based fieldmaps of positive and negative polarity were also collected before each run with similar parameters as the sequence.

Standard resolution data

A wholebrain multiband echo-planar imaging (EPI) protocol was collected in eleven participants with a 2.5 mm isotropic voxel size and the following parameters: multiband acceleration = 4, 72 slices, TR = 1120ms, TE = 30ms, A-P phase encoding, -30 degrees slice orientation from AC-PC line, flip angle = 54°, FOV= 192mm x 192mm. EPI-based fieldmaps of positive and negative polarity were also collected before each run with similar parameters as the sequence.

fMRI Preprocessing

Data was preprocessed using a standard pipeline for preprocessing of multiband data. Using FSL (Smith et al., 2004), images were brain extracted, realigned, high-pass filtered (100s threshold), and unwarped. Images were denoised by ICA component removal. Components were extracted using FSL's Melodic, classified into signal or noise with a classifier trained on separate datasets or manually classified for the high-resolution dataset. T2 images were aligned to T1 images with FSL FLIRT, and then both were normalized to standard space using ANTs (using CIT168 high resolution T1 and T2 templates (Avants et al., 2009; Tyszka and Pauli, 2016)). The functional data was first co-registered to anatomical images using FSL's FLIRT, then registered to the normalized T2 using ANTs. For univariate analyses, data was spatially smoothed in FSL with a 5-mm FWHM Gaussian kernel. For multivariate analyses, data was spatially smoothed with a 2-mm FWHM Gaussian kernel.

Behavioural Predictions of Normalization

Let the vector $\mathbf{v} = [v_i, v_j] \in \mathbb{R}_+^2$ with components v_i and v_j denote the values of item i and j individually, and let $v_{i,j}$ denote the value of the bundle of these items. Theories of linear value integration would imply that $v_{i,j} = v_i + v_j$. However if value is constructed via a relative coding, this relationship is sub-additive. To see why, consider a normalization function $Z(\mathbf{v}) = \frac{\beta\mathbf{v}}{\sigma + \sum_{n \in N} v_n}$ with scale parameter $\beta > 0$ and contrast parameter $\sigma > 0$ which implements a linear constraint on neural activity (?). If the valuation of a bundle is $v_{i,j} = z_i + z_j$, then $v_{i,j} = \beta \frac{v_i + v_j}{\sigma + v_i + v_j}$ which is a concave function of the sum of the item values (see Figure 5).

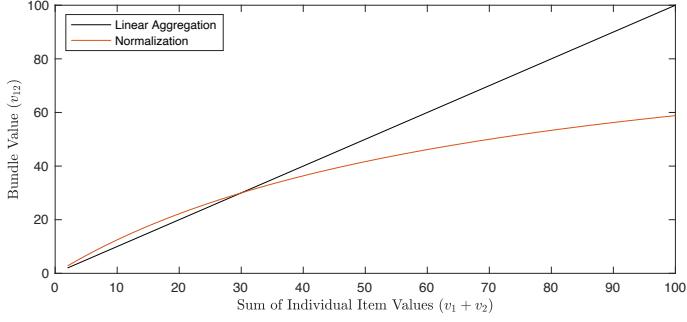


Figure 5: A normalization model predicts that the bundle value should be a concave function of item values if the reported bid is the sum of the normalized values.

Behavioral Analyses

Linear and nonlinear regression analyses were performed to model how bundle value is computed as a function of the values of the constituent items in the bundle ($v_{ij} = f(v_i, v_j)$). The linear model represents bundle value as a linear combination of the individual item values:

$$v_{i,j} = \beta_0 + \beta_1 v_i + \beta_2 v_j$$

Item i and item j simply correspond to the item shown on the left and right, respectively, during stimulus presentation. A mixed-effects model was estimated across all subjects, with subject-specific random effects terms for intercept and slope. Nonlinear mixed effects models were also constructed:

$$\text{Polynomial: } v_{i,j} = \beta_0 + \beta_1 v_i^2 + \beta_2 * v_i^3 + \beta_3 * v_j^2 + \beta_4 * v_j^3$$

$$\text{Power: } v_{i,j} = \beta_0 + \beta_1 * v_i^{\beta_2} + \beta_3 * v_j^{\beta_4}$$

$$\text{Logarithmic: } v_{i,j} = \beta_0 + \beta_1 * \log(\beta_3 + v_i) + \beta_4 * \log(\beta_5 + v_j).$$

All models and model statistics were estimated with Matlab, with the fitlme and nlmefit functions, and model fits were evaluated with BIC and R² (Table 1). To plot the fitted line/curve for each model in Figure 7.b, bundle value was computed from the fitted function parameters with v_i set equal to v_j for all values from 0-10 (with a 0.01 step size). The sum of these values (double of the value set both for ItemValue1 and ItemValue2) is represented by the x-axis.

The linear models were also separately fit on data from each type of bundle, food, trinket, mixed, and duplicates of the same item. Due to a small amount of data per subject, random effects terms could not be properly estimated for bundles of the same item, and therefore only a fixed effects model was estimated.

Regions of interest

Regions of interest (ROIs) were defined using the AAL database (Tzourio-Mazoyer et al., 2002). The labels used in the paper are mapped to the original ROI names as follows: ACC_pre: rACC, ACC_sup: dACC, Frontal_Inf_Orb_2: vlPFC, Frontal_Med_Orb: vmPFC, OFCant: OFCant, OFClat: OFClat, OFCmed: OFCmed, OFCpost: OFCpost, Frontal_Sup_Medial: dmPFC, Frontal_Sup_2: dlPFC, Frontal_Mid_2: MFG, Frontal_Inf_Tri: IFG.

Univariate Analyses

Univariate analyses were conducted in SPM12. General linear models (GLMs) were constructed to examine how subjective value and choice are encoded in the brain during the choice task. This GLM included a regressor for value, time locked to the onset of the trial as a parametric modulator, which was modified in separate GLMs to test hypotheses about value representation. The stimulus onset regressor additionally had another parametric modulator for trial type (-1 for single item trials and 1 for bundle trials). Thus, a contrast corresponding to the representation of bundle value or single item value could be computed as the interaction between value and trial type. Reaction times were additionally included as a third parametric modulator. Regressor for the choice made on a trial was time locked to the time of choice, with separate regressors for choosing an item/bundle and choosing the reference monetary amount. Regressors of no interest were included: left and right button presses (duration=0), motion regressors, and run. Missed trials were not modeled. Data across all three days in a subject were included in the same model, with different days and separate regressors per day entered as different sessions.

Different representations of value were included in the value regressor. The value contrast in Figure 4b,c used a normalized version of value, where value was z-scored by trial type (item or bundle). The absolute value model

uses the raw WTP \$ amount that the subject rated an item/bundle outside the scanner. WTP minus reference (value difference) used WTP minus the reference monetary amount on that trial.

Multivariate Analyses (MVPA)

To examine the nature of the bundle value code as the fine-grained distributed level, we implemented MVPA and RSA analyses. All analyses used the PyMVPA toolbox (Hanke et al., 2009), Scikit-learn functions, and custom Python code.

MVPA samples

To prepare for MVPA and RSA, we extracted trial-by-trial voxel-wise fMRI responses. A GLM was designed for each participant that modeled the onset and duration of each trial separately to extract the voxel responses that were unique to each trial. Other regressors of no interest modeled the other events and were not separated by trial (one regressor across an entire day): outcome phase (onset and duration), left and right button presses (duration=0). As in the univariate analyses, data across all three days in a subject were included in the same model, with different days and separate regressors per day entered as different sessions. After estimating the models, the parameter estimate maps (beta maps) for each trial were concatenated into a 4D dataset, with length equal to the number of trials a subject completed across the experiment.

Cross-Decoding Analysis

To test for distinct bundle value codes, single item value codes, and general value regions, we implemented a cross-decoding analysis similar to previously used methods (McNamee et al., 2013) at the ROI level. The 4D dataset of beta maps for every trial were loaded with PyMVPA functions. Value was z-scored by trial type (item or bundle) and included as the targets to predict in the PyMVPA dataset. Then the voxels for each ROI were entered as features for the cross decoding procedure. Ridge regression was used as the decoder in all analyses (Scikit-learn's, `linear_model.Ridge`), with $\alpha = 10^3$ (this parameter was optimized with sweeping). Each run was used as a separate cross-validation fold. Decoders were trained on the training samples from one trial type and tested on samples from the held out run (leave-one-run-out cross-validation). Decoders were tested on the samples from both trial types in the test set, even though they were trained on one type. However, decoder predictions were quantified separately for individual item samples in the test set and bundle samples in the test set. This ensured that performance could be compared across trial type separately from within trial type. Prediction accuracy was quantified by the Pearson correlation between the predictions and true value labels and averaged across cross-validation folds. This resulted in four decoder prediction accuracy metrics for each ROI and subject according to the four train/test splits: ‘train and test on single

items,’ ‘train and test on bundles,’ ‘train on single items/test on bundles,’ ‘train on bundles/test on single items’. The average prediction accuracy across participants for each of these train/test splits is plotted in Figure 8c. Significance was assessed for each split vs. chance level ($r=0$) with two-sided one-sample Wilcoxon signed rank tests at $P < 0.05$ and FDR-corrected for multiple comparisons (ROIs) at $q = 0.05$. To assess if a decoder’s prediction accuracy was significantly different between conditions in the test set, we used two-sided two-sample Wilcoxon signed rank tests at $P < 0.05$ and FDR-corrected for multiple comparisons (ROIs) at $q = 0.05$.

Cross-Decoding — Normalization of the value code

A similar cross decoding analysis was performed to test whether the neural representation codes for value in absolute or relative terms at the multivariate level. Ridge regression decoders were trained on 14/15 runs in one condition as described above. The only difference in the decoder procedure is that absolute value codes were used as targets rather than the relative/normalized values used in the first cross decoding analysis. The decoder is then tested on the samples from both trial types in the test set. In this analysis the decoder’s is tested on both the individual item samples and bundle samples at the same time, in order to test if the decoder ranks the samples across trial types in an absolute or relative fashion. This procedure thus outputs predictions on all trials in the held out run, and these predictions are correlated to 1. the true labels in absolute value (WTP \$ amount) and 2. the true labels normalized by condition (relative value). These prediction accuracy metrics were then averaged across participants and plotted in Figure 3.4d. Differences in prediction accuracy between absolute value and relative value were tested against chance with a nonparametric version of the Paired T-test, the two-sided two-sample Wilcoxon signed rank test $P < 0.05$ and FDR-corrected for multiple comparisons (ROIs) ($q = 0.05$).

Representational Similarity Analysis (RSA)

To examine whether the representational geometry of the regions of interest correlated to an absolute value or relative value code, we conducted representational similarity analysis (RSA). The 4D dataset of beta maps for every trial were loaded with PyMVPA functions as in the cross decoding analysis. Trial by trial neural dissimilarity matrices (DSMs) were constructed for every ROI by computing pairwise comparisons of the beta map across trials with PyMVPA’s PDist function. Euclidean distance was used as the distance metric. Two model DSMs were constructed: 1. trial by trial pairwise distances according to the difference in absolute \$ value of the stimuli between trials 2. trial by trial pairwise distances according to the difference in normalized value of the stimuli between trials (where value was z-scored by trial type (item or bundle)). For all DSMs, within day comparisons were removed to avoid potential confounds due to similarity being driven by patterns being in the same run or day. Neural DSMs and model DSMs were then compared with Pearson correlations. Differences in the correlations between absolute value and relative value were tested against

chance with a non-parametric version of the Paired T-test, the two-sided two-sample Wilcoxon signed rank test $P < 0.05$ and FDR-corrected for multiple comparisons (ROIs) ($q = 0.05$).

a

WTP Task Outside Scanner



b

Choice Task Inside Scanner

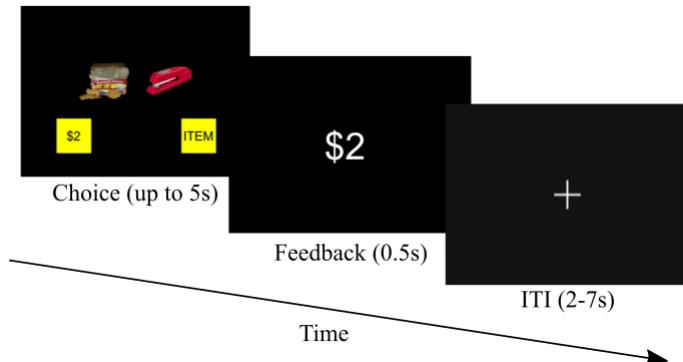


Figure 6: **Experimental Design**

- a.** WTP task. Participants reported how much they would be willing to pay for items and bundles of two items in a BDM auction. This was untimed outside the scanner.
- b.** Choice task. Inside the scanner, participants made choices with the items and bundles. During single item trials, a choice was made between an item and a reference monetary amount equal to the median bid of single item trials in the WTP task. Similarly during bundle trials, a choice was made between a bundle and a reference monetary amount equal to the median bid of WTP bundle trials. Participants had up to 5s to make a choice indicated by a right or left button press. The experiment involved 3 days of scanning 5 runs of the task for a total of 15 runs.

Model	R ²	BIC
Linear	0.7771	39151
Polynomial	0.7722	39510
Power	0.7797	39049
Logarithmic	0.7777	39097

Table 1. Behavioral models of bundle value

R² and BIC scores reflect fit across participants. Each model fits random effects parameters per participant. See Methods for model details.

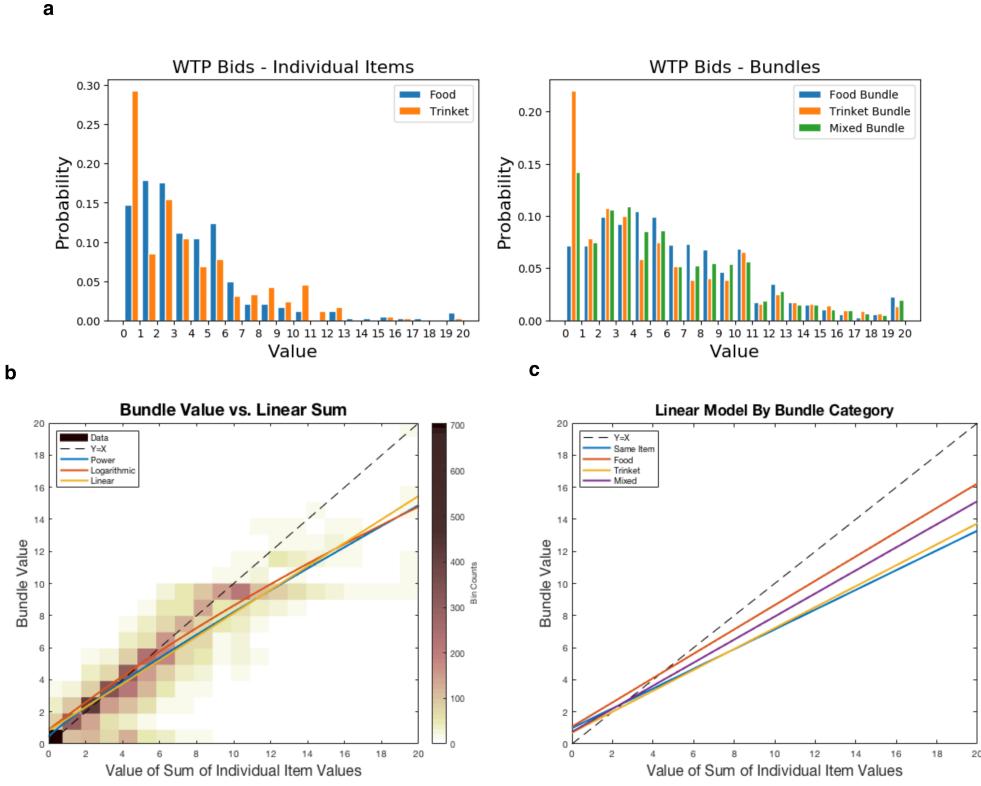


Figure 7: **WTP Behavior**

- a. Histograms of WTP bids for individual items and bundles across all subjects.
- b. Density plot of the WTP value of a bundle vs the sum of the values of the constituent items in a bundle in all subjects. If bundle value was equal to a linear addition of the constituent item values, bundle values would lie along the diagonal $Y=X$. Three models were constructed to predict the value of a bundle as a function of the individual item values: a linear model, a nonlinear power model, and a nonlinear logarithmic model. The fitted curves for each model are plotted along the density plot. All three models display that bundle value is a subadditive function of the individual item values, as they extend below the $Y=X$ line.
- c. The fitted linear model stratified by bundle type.

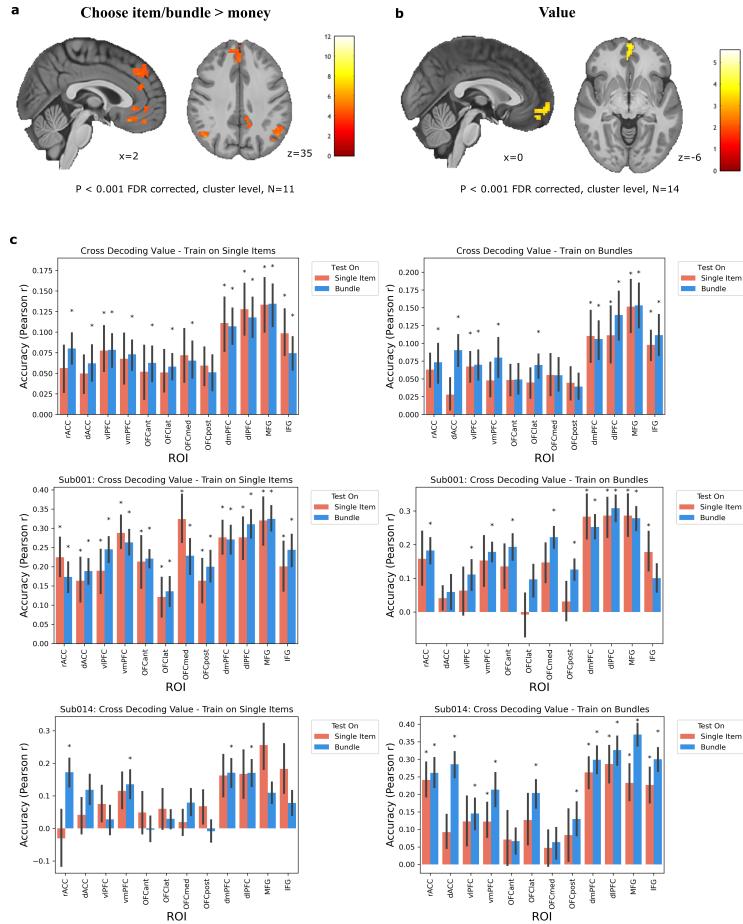


Figure 8: **Neural representation of subjective value and choice.**

- a.** Areas more active when the item or bundle was chosen than when the reference monetary amount was chosen. Significant clusters in dmPFC/SFG, ACC, vmPFC, and angular gyrus (N=11). Clusters are defined by a $P < 0.001$ uncorrected cluster forming threshold followed by $P < 0.001$ FDR cluster-level correction. Second level result shown here without the 3 participants that underwent the high resolution scanning protocol because that protocol was focused on ventral frontal areas (with a limited field of view) and did not provide coverage in dorsal frontal areas.
- b.** Neural correlates of the value of the item/bundle presented. Significant cluster in anterior vmPFC (N=14). Clusters are defined by a $P < 0.001$ uncorrected cluster forming threshold followed by $P < 0.001$ FDR cluster-level correction.
- c.** MVPA cross-decoding analysis results. Ridge regression decoders were trained on samples from one condition and tested on samples from a held out run in both conditions. Top row: Group result. Second and third rows represent the results of individual subjects, the first and the last subject. Left: decoders trained on trials of single items. Right: decoders trained on bundle trials. Asterisks * represent significant prediction accuracies on a test partition (two-sided one-sample Wilcoxon signed rank test $P < 0.05$ and FDR-corrected for multiple comparisons $q = 0.05$). At the group level, there were no significant paired differences in prediction accuracies between test conditions for any ROI (two-sided two-sample Wilcoxon signed rank test $P < 0.05$ and FDR-corrected for multiple comparisons $q = 0.05$). Error bars reflect SE across participants. See Supplementary Figure 4 for the individual subject results for the other subjects.

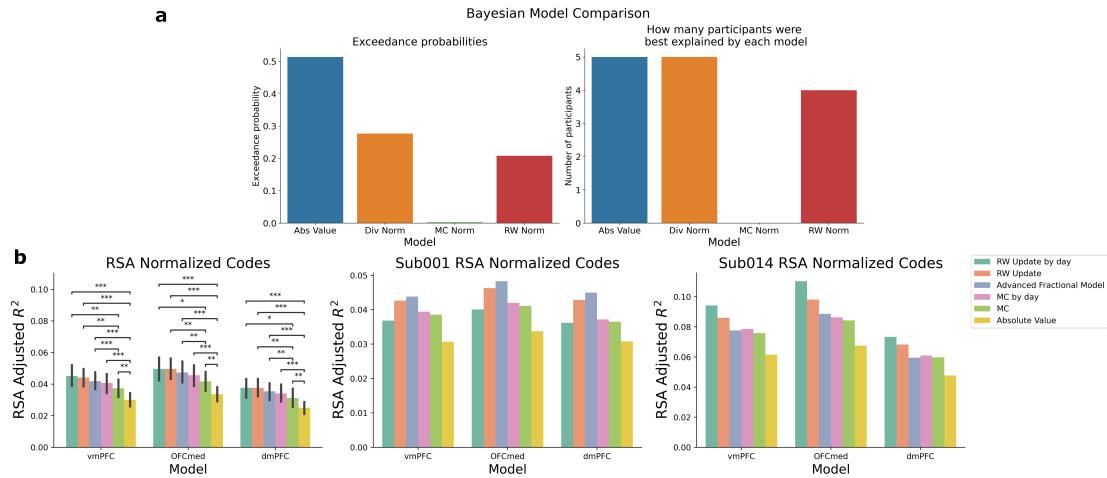


Figure 9. Comparing different normalization models.

a. Bayesian model comparison of behavior comparing different ways of computing value. Absolute value representations are compared with different ways of implementing divisive normalization. RW - Rescorla Wagner, MC - Monte Carlo.

b. RSA results comparing several representations of divisively normalized value (and absolute value) to the neural data. Asterisks * represent significant differences between models (two-sided one-sample Wilcoxon signed rank test *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, all FDR-corrected for multiple comparisons). Left group result, middle and right two representative subjects (see Supplementary Figure X for remaining subjects).

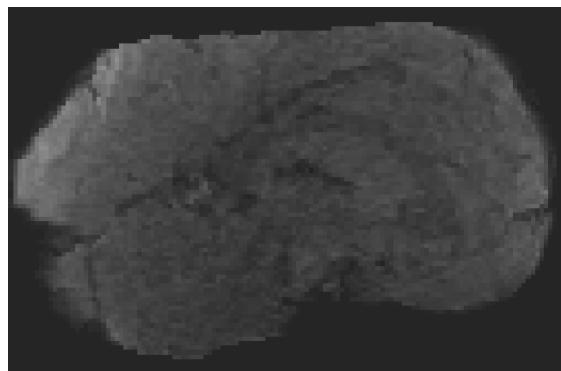


Figure 10: **Supplementary Figure 1. High Resolution Sequence**

High-resolution partial volume scans were collected in three participants with a 1.5mm isotropic voxel size.

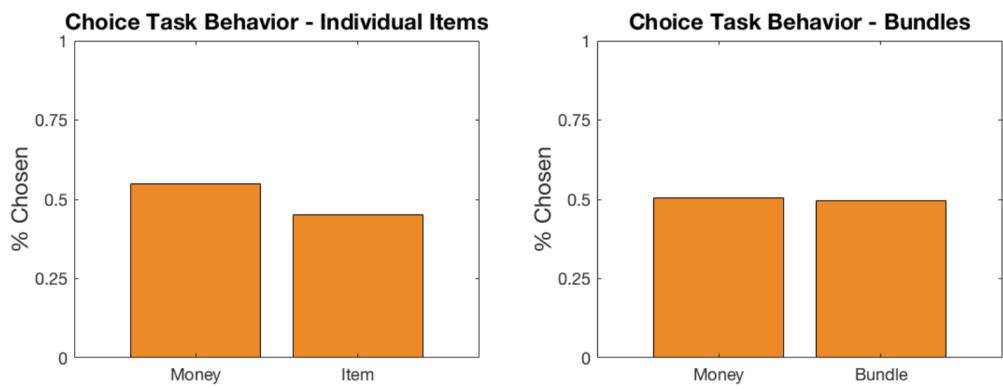


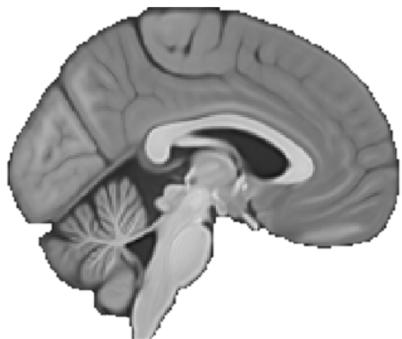
Figure 11: **Supplementary Figure 2. Behavior on the choice task.**

Percentage of trials in which the item or bundle was chosen vs. the reference monetary amount.

Single item value > Bundle value



Bundle value > Single item value



$P < 0.001$ FDR corrected, cluster level

Figure 12: **Supplementary Figure 3. Bundle Value vs Single Item Value.**

Univariate contrasts testing the interaction of value and trial type. No clusters survived in either comparison after multiple comparisons correction.

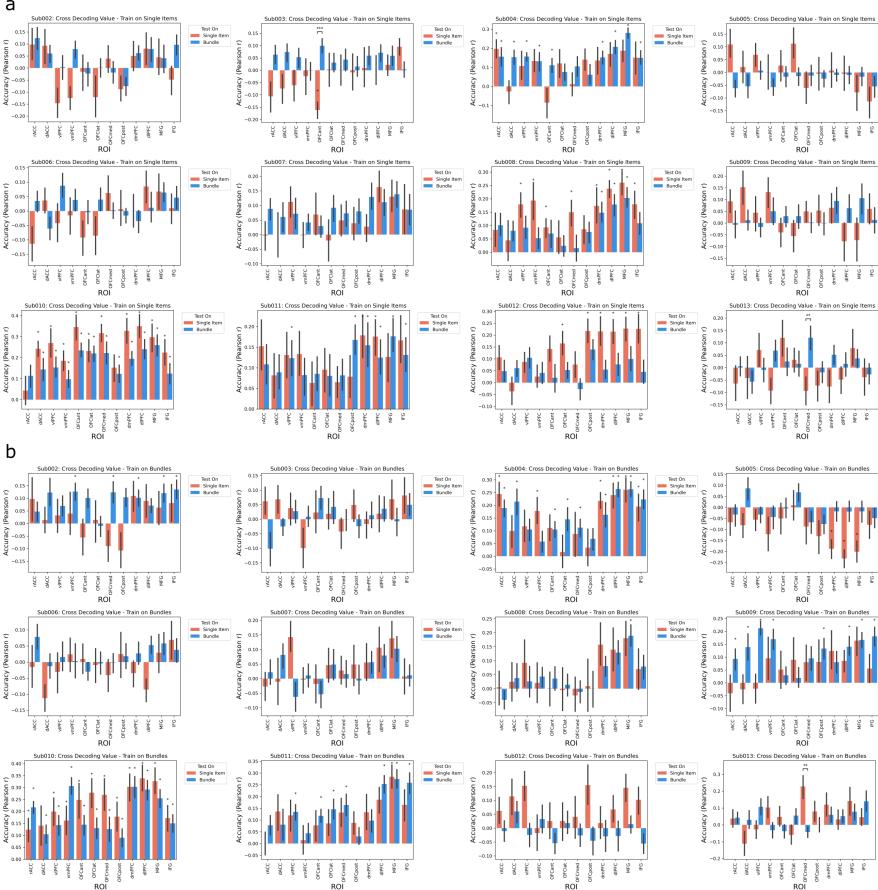


Figure 13: **Supplementary Figure 4. Bundle Value vs Single Item Value.**

MVPA cross decoding analysis individual subject results. Left: decoders trained on trials of single items. Right: decoders trained on bundle trials. Asterisks * represent significant prediction accuracies on a test partition (two-sided one-sample Wilcoxon signed rank test $P < 0.05$ and FDR-corrected for multiple comparisons $q = 0.05$). At the group level, there were no significant paired differences in prediction accuracies between test conditions for any ROI (two-sided two-sample Wilcoxon signed rank test $P < 0.05$ and FDR-corrected for multiple comparisons $q = 0.05$). Error bars reflect SE across participants.

Food items used	
1. 3 Musketeers	36. Sun Chips
2. Barnum's Animal Crackers	37. Dole Mixed Fruit
3. Doritos Nacho Cheese	38. Grapefruit
4. Chips Ahoy!	39. Banana Chips
5. Kit Kat	40. Dark Chocoloate Bananas
6. Pop-Tarts Brown Sugar Cinnamon	41. Crispy Apple
7. Pop-Tarts Brown Sugar Strawberry	42. Vegetable Chips
8. Ghiradelli Chocolates	43. Sweet Potato Chips
9. Twix Cookie Bars	44. Chopped Salad Chicken
10. Hershey's Whatchamacallit Candy	45. Mexicali Salad
11. Apple Pie	46. Caesar Salad
12. Avocado	47. Veggie Wrap
13. Blackberries	48. Super Burrito
14. Cauliflower	49. Chocolate and Berry
15. Ritz Crackers'n Cheese Dip	50. Green Beans Chips
16. Cherry Pie	51. Salami
17. Chocolate Muffins	52. Smoked Turkey
18. Powdered Donuts	53. American Cheese
19. Granny Smith Apple	54. Chicken and Roasted Beet
20. Green Grapes	55. Mozzarella Cheese
21. Mango	56. Roast Beef
22. Milano Cookies	57. Caprese Sandwich
23. Orange	58. Tuna Salad Wrap
24. Raspberries	59. Smoked Salmon
25. Red Velvet Cake	60. Plain Yogurt
26. Quaker Chewy Granola Bar	61. Strawberry Yogurt
27. Starburst	62. Blueberry Yogurt
28. Strawberry	63. Deviled Eggs
29. Crunchy Donuts	64. Smore's Chewy Bars
30. Chicken Tikka Masala	65. Gnocci
31. Lamb Vindaloo	66. Magherita Pizza
32. Pollo Asado Burrito	67. Macarons
33. Bean and Cheese Burrito	68. Blueberry Crisp Clif Bars
34. Chocolate Chip Clif Bars	69. Yogurt Pretzels
35. Ferrero Chocolates	70. Chocolate Pretzels
Consumer goods used	
1. A Brief History of Time book	21. Lock
2. Freakonomics book	22. Notebook
3. 1984 book	23. Bathroom scale
4. Water bottle	24. Playing cards
5. Wireless mouse	25. Honey clementine candle
6. Yoga mat	26. Roses candle
7. Hitchhikers book	27. Umbrella
8. Lord of the Rings book	28. Android charger
9. Caltech backpack	29. iPhone charger
10. Caltech hat	30. Clothes hangers
11. Caltech banner	31. Beach towel
12. Caltech keychain	32. Cooking supplies
13. USB stick 16GB	33. Kitchen utensils
14. Caltech mug	34. Pens
15. Caltech drawstring bag	35. Plates
16. Desk lamp	36. Portable charger
17. Stapler	37. Portable speaker
18. Over the ear headphones	38. Screwdrivers
19. Head backpack	39. Sunglasses
20. Batteries	40. Surge Protector

Table 2: Items used in experiment.

References

- Bartra, O., McGuire, J. T., and 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.
- Becker, G. M., DeGroot, M. H., and Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science*, 9(3):226–232.
- Bettman, J. R., Luce, M. F., and Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, 25(3):187–217.
- Carandini, M. and Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62.
- Chib, V. S., Rangel, A., Shimojo, S., and O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *Journal of Neuroscience*, 29(39):12315–12320.
- Clithero, J. A. and Rangel, A. (2014). Informatic parcellation of the network involved in the computation of subjective value. *Social Cognitive and Affective Neuroscience*, 9(9):1289–1302.
- Cox, K. M. and Kable, J. W. (2014). BOLD Subjective Value Signals Exhibit Robust Range Adaptation. *The Journal of Neuroscience*, 34(49):16533–16543.
- Dumbalska, T., Li, V., Tsetsos, K., and Summerfield, C. (2020). A map of decoy influence in human multialternative choice. *Proceedings of the National Academy of Sciences*, 117(40):25169–25178.
- Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., and Pollmann, S. (2009). Pympva: A python toolbox for multivariate pattern analysis of fmri data. *Neuroinformatics*, 7(1):37–53.
- Kriegeskorte, N., Mur, M., and Bandettini, P. A. (2008). Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience*, 2:4.
- Levy, D. J. and Glimcher, P. W. (2011). Comparing apples and oranges: Using reward-specific and reward-general subjective value representation in the brain. *Journal of Neuroscience*, 31(41):14693–14707.
- Lim, S.-L., O'Doherty, J. P., and 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. *Journal of Neuroscience*, 33(20):8729–8741.

- Lin, A., Adolphs, R., and Rangel, A. (2012). Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience*, 7(3):274–281.
- Louie, K., Glimcher, P. W., and Webb, R. (2015). Adaptive neural coding: From biological to behavioral decision-making. *Current Opinion in Behavioral Sciences*, 5:91–99.
- Louie, K., Grattan, L. E., and Glimcher, P. W. (2011). Reward value-based gain control: Divisive normalization in parietal cortex. *Journal of Neuroscience*, 31(29):10627–10639.
- McNamee, D., Rangel, A., and O’Doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature Neuroscience*, 16(4):479–485.
- Nagle, T. T. and Holden, R. K. (1987). *The strategy and tactics of pricing*, volume 3. Prentice Hall Englewood Cliffs, NJ.
- O’Doherty, J. P., Rutishauser, U., and Iigaya, K. (2021). The hierarchical construction of value. *Current opinion in behavioral sciences*, 41:71–77.
- Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *Journal of Neuroscience*, 29(44):14004–14014.
- Padoa-Schioppa, C. and Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience*, 11(1):95–102.
- Rustichini, A., Conen, K. E., Cai, X., and Padoa-Schioppa, C. (2017). Optimal coding and neuronal adaptation in economic decisions. *Nature communications*, 8(1):1208.
- Suzuki, S., Cross, L., and O’Doherty, J. P. (2017). Elucidating the underlying components of food valuation in the human orbitofrontal cortex. *Nature Neuroscience*, 20(12):1780–1786.
- Thorne, J. (2004). Discounted bundling by dominant firms. *George Mason Law Review*, 13:339.
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., and 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.
- Varian, H. R. (2014). *Intermediate microeconomics: A modern approach: Ninth international student edition*. WW Norton & Company.