Coexistence of sensory qualities and value representations in human orbitofrontal cortex

Yoshimoto, Takaaki

Doctor of Philosophy

National Institute for Physiological Sciences

Department of System Neuroscience

Division of Cerebral Integration

SOKENDAI (The Graduate University for Advanced Studies)

School of Life Science

Department of Physiological Sciences

Contents

Abstract

The orbitofrontal cortex (OFC) supposedly plays a critical role in value-based decision-making. Abundant evidence exists on the role of the OFC in reinforcer devaluation tasks, which assess the ability to represent identity, sensory qualities, and subjective values of the expected outcomes. However, it remains unclear which aspect is specifically represented in this region. Here, by conducting functional magnetic resonance imaging experiments, wherein participants rated the palatability of 128 food items using photos, investigated whether the human OFC represents object identity, sensory qualities, or value. Employing many items helped me dissociate object identity from sensory qualities and values; the inferred sensory qualities of the identical items were manipulated by a change in the metabolic state. Moreover, value differences between items were analytically controlled by a technique similar to age adjustment. The palatability ratings for food items significantly decreased after the meal. By applying representational similarity analysis, I found that the OFC represents value, as reported previously. Moreover, identical items were represented similarly in the lateral OFC in a given metabolic state; however, these representations were altered post feeding. Importantly, this change was not explained by subjective value, suggesting that the OFC represents sensory quality and value, but not object identity.

1. Introduction

Decision making is a cognitive process in which multidimensional information must be integrated. For example, when people decide what to eat at a restaurant, they infer gustatory experience from the photos and descriptions on the menu while considering our current metabolic state (i.e., satiated or hungry). Even if the dessert options themselves are identical, inferred gustatory experiences are not the same at the time of the first course as when dessert is ordered, suggesting that the inferred sensory qualities and identity can be dissociated under certain conditions. As in the old proverb, "hunger is the best sauce," the sensory qualities of food are dynamically altered by changes in one's internal metabolic state while the food identity is stably represented. By integrating such information, an individual estimates the value of each option in deciding what to eat. The orbitofrontal cortex (OFC) integrates sensory and visceral information, providing a basis for such decision-making (Kringelbach, 2005). Specifically, a growing body of research has shown that the OFC encodes the predicted outcomes that follow either sensory events or behavioral choices (Rudebeck et al., 2013; Schoenbaum et al., 2009). In particular, studies on rodents, monkeys, and humans have consistently demonstrated that the OFC plays a critical role in reinforcer devaluation, in which specific food rewards are provided until satiated or paired with illness (Gallagher et al., 1999; Gremel & Costa, 2013; Izquierdo, 2004; Pickens et al., 2005; Reber et al., 2017; Rudebeck et al., 2013; West et al., 2013). While healthy participants prefer an option that leads to the food with which they are not sated, OFC-lesioned animals or patients show no such preference.

However, the specific function of this region remains under debate (Rudebeck & Murray, 2014; Schoenbaum et al., 2011; Stalnaker et al., 2015), partially because of the complexity of the devaluation paradigm. Although behavior is driven by composite representations of identity, sensory qualities, and subjective values of expected outcomes in

this paradigm (Howard et al., 2015; Rudebeck & Murray, 2014), the emphasized aspect of the task varies across studies. Although several studies have focused on the involvement of the OFC in value representations (Gallagher et al., 1999; Kringelbach et al., 2003; Rudebeck et al., 2013), others have emphasized representations of the sensory qualities or food identity (Burke et al., 2009; Gremel & Costa, 2013; Schoenbaum et al., 2011). Indeed, there are several difficulties in dissociating these representations in the devaluation paradigm. For example, because different taste rewards evoke different sensory qualities, differences in identity and sensory qualities cannot be dissociated should the reward type given be small, as in animal experiments. Moreover, devaluation alters the sensory qualities of expected outcomes, as well as their subjective value (Rudebeck & Murray, 2014), although their identity remains the same (e.g., orange juice can be identified as orange juice even if it does not taste good after devaluation). Thus, it remains unclear whether the OFC represents identity, sensory quality, or the subjective value of the expected outcome.

In this study, I sought to answer this question using a non-selective devaluation paradigm with many reward types. In our functional magnetic resonance imaging (fMRI) experiment, participants rated the palatability of 128 various food photos before and after meals on two experimental days. By employing representational similarity analysis (Kriegeskorte, 2008), I investigated where in the brain the representations of identical objects were maintained through pairs of the same or different metabolic states. If the OFC represents object identity, the representations of identical objects would be invariant across different metabolic states, whereas if the OFC represents the inferred sensory qualities but not identity, those representations would be maintained only in the same metabolic state. Furthermore, I investigated whether such representations could be explained by differences in the subjective value of the food. These procedures allowed me to specify whether the OFC represents identity, sensory qualities, or the subjective value of the expected outcomes in the devaluation paradigm.

2. Materials and Methods

2.1. Participants

Twenty-four right-handed, healthy, non-obese male participants [mean ± standard error of mean (SEM) age 27.4 ± 5.9 years, range 25–39; body mass index (BMI) 22.1 ± 2.9 kg/m², range 17.0–27.6] with no history of neurological or psychiatric problems were recruited from the local area via email. This study was approved by the ethical committee of the National Institute for Physiological Sciences of Japan. All participants provided written informed consent. No statistical test was performed to determine the sample size a priori. The sample size that I chose is similar to those used in previous studies (Chikazoe et al., 2014; Haxby et al., 2011; McNamee et al., 2013).

2.2. Experimental design

2.2.1. Experimental procedure

The same experimental procedure was repeated twice in two days; the two experiments were spaced 7–41 days apart (17.0 ± 9.0, mean ± standard deviation [SD]). Participants were instructed to abstain from the intake of any food and caffeinated or alcoholic beverages after 8 pm on the day preceding each experiment. The schedule on the experiment day was controlled

in the same manner. Between the first and second fMRI sessions, the participants ate the standard Japanese lunch box (776 kcal), which contained rice seasoning, fried egg, grilled salmon, simmered vegetables, meatballs, fried tofu fritter, and simmered black bean. If they were not satiated, they could eat as much of the three types of snacks (chocolate cookie, salt cracker, and rice cracker) as they wished. Total calorie consumption was 802.2 ± 64.8 (mean ± SD).

2.2.2. Visual stimuli

I collected 128 food photos from the Internet. Half of the images depicted the main dishes (mainly salty and savory), and the other half depicted desserts (mainly sweets). The descriptions are provided in the Supplementary Table. 1. No food items in these photos were the same as those that the participants actually ate in the experiment. Every photo was resized to 600 × 400 pixels. Each food type photo set was reordered according to the average of the luminance values of all the pixels, and the luminance values in all pixels of each dessert photo were controlled by matching their average with that of the main dish photo in the same luminance rank. To generate control stimuli, the pixels in each food image were reordered, resulting in 128 scrambled images. Sixty-four control images were constructed from the average image of the scrambled images of different food types with the same luminance rank.

2.3. fMRI task

The same fMRI experiments were conducted before and after a meal, and the experiments were repeated on day 2. Thus, the participants completed four fMRI sessions. In an fMRI session, 128 food photos, 64 scrambled images, and 64 resting images (white cross in the center of a black background) were presented through four runs. The order of the 256 stimuli was fixed across all the fMRI sessions. While participants were asked to rate the palatability of the food using a ten-point scale during the presentation of the food image, they were asked to rate the brightness of each image during the presentation of the scrambled image (1=not at all; 10=most ever). Each stimulus without a resting image was presented for 2,900 ms, during which participants considered the rating on a ten-point scale. This was followed by the presentation of a fixation image (red cross in the center of a black background) for 1,100 ms, during which participants verbally reported their rating. These ratings were recorded manually by the experimenter, who was listening via an audio system connected to a microphone near the participant's mouth. During the presentation of the resting image for 4,000 ms, the participants were given a break (Fig. 1A).

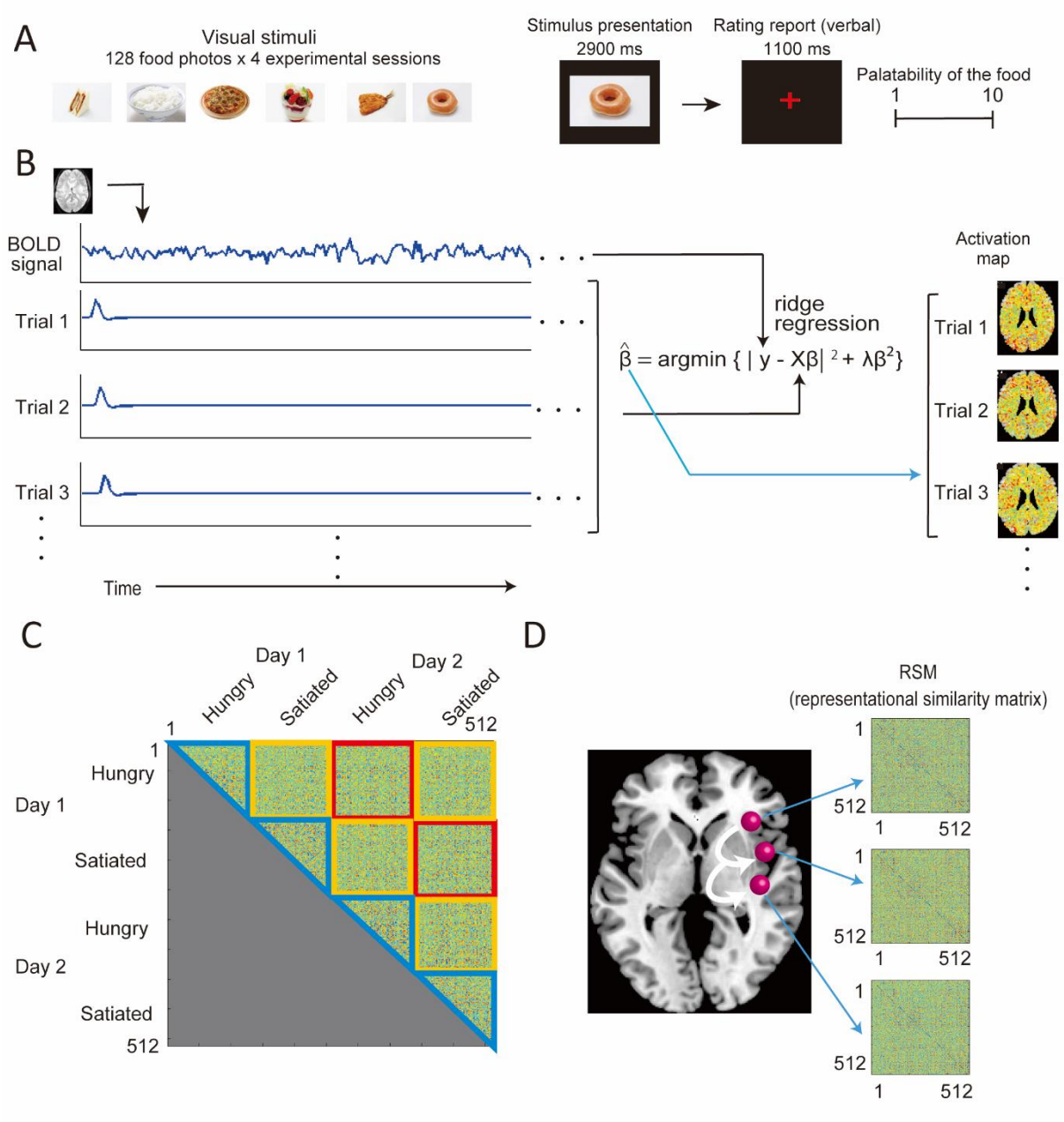


Fig. 1. Experimental task, analytical procedures, and data structure

(A) fMRI task procedure. Twenty-four healthy male adults completed four fMRI sessions in 2 days. Before and after the meal, the same fMRI experiments were conducted, and this was repeated on day 2. A total of 128 food photos were presented during the fMRI sessions.

Participants were instructed to determine their rating for palatability of the food using a 10-point scale during the presentation of each photo. This was followed by fixation, during which the participants reported their rating verbally. (B) Analytical procedures to generate the activation map in each trial. BOLD signal in each voxel was extracted from the preprocessed fMRI data. The presentation period of every food stimulus was modeled in every experimental session as well as regressors of no interest. To estimate the precise regression coefficient weights in adjacent trials, I adopted a ridge regression method. Totally, 128 (food photos) × 4 (experimental sessions) activation maps were acquired for each participant. (C) Schematic illustration of the session pair structure. To conduct the representational similarity analysis, I first constructed a representational similarity matrix on a trial-by-trial basis. The 4 × 4 matrix indicates the combinations of experimental sessions (due to redundancy, only the upper triangle is shown). Each element represents an experimental session pair. Each color surrounded by a square or upper triangle represents a 128 × 128 representational similarity matrix (RSM) for each session pair. Because there were four experimental sessions in total, I analyzed 10 session pairs, including four within-session pairs (shown in blue) and six cross-session pairs (shown in red and orange). Moreover, cross-session pairs could be divided into the same metabolic state pairs (shown in red) and different metabolic state pairs (shown in orange). (D) Schematic illustration of the searchlight analysis. Representational similarity matrix (RSM) was constructed by computing the correlation coefficient between the neural activation patterns of

each trial pair in a 5 mm radius sphere searchlight, in the whole brain.

2.4. Imaging parameter

MR images were collected using a 3.0T-fMRI system (Verio; Siemens Erlangen, Germany) with a 32-element phased-array head coil. T2*-weighted gradient echo-planar imaging (EPI) was used to obtain functional images. The sequence was as follows: repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, TA = 900 ms, field of view = 192 mm, 60 slices with isotropic 3-mm voxels, multiband EPI: multiband factor = 3. For anatomical imaging, a T1-weighted three-dimensional (3D) magnetization-prepared rapid-acquisition gradient echo (MP-RAGE) sequence was employed (TR = 1800 ms; TE = 1.98 ms; flip angle = 9°; field of view = 256 mm; and voxel dimensions = 1.0 × 1.0 × 1.0 mm).

2.5. Data analysis

2.5.1. Behavioral analysis

The experiment consisted of four sessions: day 1 hungry, day 1 satiated, day 2 hungry, and day 2 satiated. I conducted two-way repeated-measures analysis of variance (ANOVA) on the palatability of the food by day (days 1 and 2) and metabolic state (hungry and satiated) as factors, using SPSS 23 (See Supplementary Results and Supplementary Fig. 1). All ANOVAs

in this study were performed by using SPSS23 including Supplementary results. Palatability was used as a metric for the subjective value of food in this study. The palatability rating scores were further analyzed separately for the main dishes and desserts, as shown in the Supplementary Results and Supplementary Fig. 2. Furthermore, using the generalized linear model, I investigated whether the amount of snack consumption during the meals affected the palatability after the meals. The results of this analysis are also reported in Supplementary Results.

2.5.2. fMRI data preprocessing for the representational similarity analysis

Standard image preprocessing was performed using the Statistical Parametric Mapping (SPM12) package in MATLAB (R2014a). The functional images were realigned to correct for head motion. Each participant's T1-weighted image was co-registered with the mean image of all functional images for each participant. Accordingly, all functional images were processed by segmentation and normalization to the MNI-152 space using the unified segmentation-normalization tool in SPM12. To obtain precise voxels in the gray and white matter of each participant, the first EPI volume in the first run of day 1 was segmented into gray and white matter masks using the tissue probability map of SPM12. To construct a representational similarity matrix (RSM) that was associated with the identical objects, neural activations were estimated on a trial-by-trial basis. To estimate the brain activity evoked at each trial by

regression, the response and explanatory variables were computed as follows: I extracted the BOLD signal time series for each voxel. Nuisance variables (including six head motion components encoding the x, y, and z directions and pitch, roll, and yaw for each run) and the dummy variables encoding each run, were regressed out from the data. The residual in this regression was high-pass filtered at 1/128 Hz and then used as a response variable in each voxel. To construct a design matrix for the explanatory variables, each presentation period of 128 food and 64 scrambled images was separately modeled for each stimulus, and the verbal reporting period was modeled as a regressor of no interest. The design matrix, which was high-pass filtered at 1/128 Hz and whitened, was used as the explanatory variable. By applying the ridge regression with the empirically determined ridge parameter (see Supplementary Methods and Results for details), the brain activity evoked by each food photo stimulus was estimated on a trial-by-trial basis, resulting in $128 \times 4 = 512$ regression coefficients (see Fig. 1B). Finally, these coefficients were transformed to t-values as estimates of the neural activity.

2.5.3. Representational similarity analysis for identical objects and metabolic states

Because it is difficult to directly model the effects of object specificity, metabolic states, values, and their interactions on neural responses, I employed representational similarity analysis, a method for uncovering representational properties underlying multivariate data (Kriegeskorte, 2008). This method is a type of pattern analysis that compares the neural and model the RSMs.

Rather than directly modeling the relationships between psychological/physiological factors and neural responses, this method compares the representations at the similarity matrices level. By establishing a correspondence between the relationships among the stimuli and their representations, a higher-level functional perspective can be obtained. All data processing and analysis of the multivariate analysis in this study were performed using MATLAB (R2014a), except for the statistical tests and multiple comparisons in the group analysis of the whole brain. Our first aim in this study was to investigate which brain areas are associated with the representations of identical objects and whether these representations are affected by metabolic state changes. A schematic explanation of this analysis is shown in Supplementary Fig. 5. For each session pair (e.g., day 1 hungry and day 2 satiated), I estimated the neural representational similarity on a trial-by-trial basis, resulting in a 128 × 128 RSM. Because there were four experimental sessions in total, six cross-session RSMs (i.e., 4 × 3 /2 = 6) and four within-session RSMs were computed (Fig. 1C). To examine how metabolic states affected neural representations, cross-session RSMs were divided into pairs of same and four different metabolic states (Fig. 1C). Following the basic scheme of representational similarity analysis, I estimated how well the model RSM (e.g., the identity pair model) explained the neural RSM by comparing two matrices (Kriegeskorte, 2008). Prior to the construction of neural RSMs, to standardize the spatial pattern of activity, 512 (128 × 4) estimates of the neural activity were de-meaned for each voxel. For the computation of the neural RSMs, only voxels within the

grey matter mask were used. Using searchlight analysis (Kriegeskorte et al., 2006), correlation coefficients of activation patterns for each trial combination (128 × 128) were calculated in a spherical searchlight (radius = 5 mm), yielding the neural RSM for each session pair (Fig. 1D). Additionally, voxels in the white matter mask were used to compute the noise RSMs from a spherical searchlight (radius = 5 mm) for each session pair. Then, these noise RSMs were averaged across white matter voxels and z-scored for each session pair. After regressing out the white matter noise component, neural RSMs were converted to vectors and rank-ordered (Supplementary Fig. 5); as a linear relationship between the two RSMs could not be assumed (Kriegeskorte, 2008). To explore the brain regions associated with the representations of identical objects, percentiles of the elements in the neural RSM were averaged separately for identical and non-identical object corresponding elements in the identical object model RSM (Supplementary Fig. 5). Because I were interested in the effect of the metabolic state change, I separately averaged these metrics across pairs of same (two session pairs, Fig. 1C: red) and different metabolic states (four session pairs, Fig. 1C: orange). Using the procedures described above, I obtained four representational similarity metrics, categorized as identical object pair in the same metabolic state (ISS), non-identical object pair in the same metabolic state (NSS), identical object pair in the different metabolic state (IDS) and non-identical object pair in the different metabolic state (NDS) in each gray matter voxel. To investigate the identical object specificity effect in the same and different metabolic state pairs, I calculated [ISS – NSS] and

[IDS – NDS], and then the data were spatially smoothed (full width half maximum = 8 mm) for each participant. In this study, I adopted the same full-width half maximum value for smoothing in both univariate and multivariate analyses. Because I focused on representations in the cerebral cortices, I masked the information map with a cerebral cortex mask created using the SUIT anatomy toolbox (http://www.diedrichsenlab.org/imaging/propatlas.htm). Under the null hypothesis of no difference in the representational similarities between identical and non-identical object pairs, the smoothed data were subjected to a non-parametric one-sample t-test across participants, using FSL's Randomize function (v2.1, 5,000 permutations). The threshold-free cluster enhancement (TFCE) correction for multiple comparisons (Smith & Nichols, 2009) was used, and the statistical threshold was set at $p < 0.05$, family wise error (FWE) (Fig. 2A). All statistical tests and multiple comparisons for group analysis of the whole brain in this study, including those reported in Supplementary Results, were performed using the same procedures as described above. To visualize the representations of identical objects and the effect of metabolic states on them, I conducted a region of interest (ROI) analysis. Representational similarity metrics were spatially smoothed separately for each category, as described above. I selected the primary visual area (V1) and fusiform gyrus (FG) as the ROIs in addition to the OFC to investigate identical object representations in detail. As previous studies reported that V1 encoded concrete visual features such as shape or color of the objects (Grill-Spector, K., & Malach, R., 2004) and FG encoded semantic categories (Mummery, C. J.

et al., 1998; McCandliss, B. D. et al., 2003), identical object representations in these areas were assumed to be independent of the metabolic state change. I generated three anatomically defined ROIs (V1: primary visual area, FG: fusiform gyrus, l OFC: lateral orbitofrontal cortex) based on the standard automated anatomical labeling (automated anatomical labeling; AAL) template (Tzourio-Mazoyer et al., 2002) within the gray matter voxels of at least 12 participants. The anatomical labels used for defining the three ROIs were as follows: V1: calcarine gyrus, 43, 44; FG: fusiform gyrus, 55, 56; l OFC: superior orbitofrontal cortex 5, 6; middle orbitofrontal cortex 9, 10; and inferior orbitofrontal cortex 15, 16. In ROI analysis, it is well-known that the use of the same dataset for both selection of the ROI and further statistical analyses entails invalid statistical inferences (i.e., double-dipping problem (Kriegeskorte et al., 2009)). To avoid this problem, I applied a leave-one-participant-out procedure (Chikazoe et al., 2019). After excluding each of the 24 participants, I conducted non-parametric one-sample t-tests on the identical object specificity effect in the same metabolic state pair with the remaining 23 participants, resulting in 24 t-value maps. I then identified the peak voxel in the three ROIs for each map and extracted four representational similarity metrics (i.e., ISS, NSS, IDS, and NDS) in the averaged data from a spherical searchlight (radius = 5 mm) region, around the peak from the left-out participant's data (Fig. 2B). I conducted paired t-tests on the identical object specificity effect in the same and different metabolic states and the difference between these effects in the data. All t-tests on behavioral, physiological, and ROI data in this study

were two-sided and Bonferroni corrected, including those in Supplementary Results. A Kolmogorov-Smirnov test was performed prior to the t-test to confirm the normal distribution of the data.

2.5.4. Univariate analysis for value representations

There was a difference in value distance distribution between identical and non-identical object pairs in some experimental session pair RSMs as shown in Fig. 3A (Because the term "distance" is used to denote a dissimilarity measure in the context of representational similarity analysis (Kriegeskorte, 2008), hereafter I refer to the absolute difference of subjective values between each pair of stimuli as the value distance). Indeed, two-way repeated-measures ANOVA on the value distance with object specificity and metabolic state as factors revealed that the main effects of both factors were significant (object specificity: $F_{1,23} = 59.27$, $p < 0.001$; metabolic state: $F_{1,23} = 46.35$, $p < 0.001$), as was their interaction ($F_{1,23} = 45.03$, $p < 0.001$) (Fig. 3B). The post hoc simple t-tests revealed significant object specificity effects in both the same and different metabolic state pairs (same state: $t_{23} = 9.06$, $p < 0.001$; different state: $t_{23} = 2.88$, $p = 0.025$) (Fig. 3B). These results indicate the possibility that the value distance distribution was correlated with the object specificity effect on neural representations in the OFC. To explore this possibility, I first investigated whether the OFC represented subjective values in both the univariate and multivariate analyses. First, in the univariate parametric modulation analysis, all

data processing and analysis were performed using the SPM12 package in MATLAB (R2014a). Value representations were analyzed separately in the hungry state, in the satiated state, and in all sessions. Functional images, which were processed by the segmentation and normalization described above, were spatially smoothed. In individual analysis, a design matrix was constructed from the combinations of the 4×4 runs (run number × session) in all sessions, and 4×2 runs in each state along the time series for each participant. I fitted a general linear model (GLM) to the fMRI data for each participant. Neural activity during each condition was modeled with box-car functions convolved with the canonical hemodynamic response function. A design matrix included task-related regressors (the main dish rating in each state, the dessert rating in each state, and verbal response in the fixation period), and one regressor for parametric modulation of the demeaned subjective values. The time series for each voxel was high-pass filtered at 1/128 Hz. Assuming a first-order autoregressive model, the serial autocorrelation was estimated from the pooled active voxels with the restricted maximum likelihood procedure and was used to whiten the data. Motion-related artifacts were minimized by incorporating six parameters (three displacements and three rotations) from the rigid-body realignment stage into each model. Because each participant was involved in two separate days experiments, the two parameter estimates for the subjective values were averaged for each participant. These averaged parameter estimates were subjected to a non-parametric one-sample t-test. The statistical threshold was set at $p < 0.05$ and FWE-corrected. The table reporting the clusters and

their information was constructed using FSL's cluster function (all sessions: Fig. 3C, whole brain map; Supplementary Fig. 7A, cluster list; Supplementary Table 2. Hungry state: whole brain map; Supplementary Fig. 8A, cluster list; Supplementary Table 3. Satiated state: whole brain map; Supplementary Fig. 9A, cluster list; Supplementary Table 4).

2.5.5. Representational similarity analysis for value representations

I also investigated whether the OFC represented the subjective values using representational similarity analysis. A schematic explanation of this analysis is shown in Supplementary Fig. 6. To use independent data from the analysis above, I used the data in the within-session pairs only (Fig. 1C, blue), which did not include data used in the analysis of identical objects. I used data from the upper triangle of the RSM only because the diagonal element in the within-session RSM was always one and the matrix was symmetrical about the diagonal. Thus, these data do not contain information about identical objects. I examined how well the value distance model explained the neural representational dissimilarity. As in the identical object analysis, correlation coefficients of activation patterns in each trial combination were calculated in a spherical searchlight (radius = 5 mm) for each session pair in each cerebral cortex voxel. I constructed a neural representational dissimilarity matrix (RDM) whose element was one minus the corresponding correlation coefficient. Then, the elements in the neural RDMs were rank-ordered. As in the identical object, white matter noise signals were regressed out from the

neural RDMs. On the other hand, I computed a value distance model matrix in which elements encoded absolute values of differences between subjective values corresponding to each trial pair. After conversion to a vector, the model data were z-scored. In a given searchlight, the neural RDM was converted to a vector and rank-ordered. To explore the brain regions associated with value representations, the correlation coefficient between a vector from the percentile of the neural RDM and a vector from the model RDM was calculated in a given searchlight. The relatedness of these RDMs was examined by 1,000 random permutations of the trial pairs (see Supplementary Fig. 6). To investigate value representations in all sessions, only in the hungry state, and only in the satiated state, I averaged the RDM relatedness metrics separately across all four within-session pairs for all sessions in the hungry and satiated states, and across two within-session pairs in hungry–hungry state pairs (hungry state) and satiated–satiated state pairs (satiated state) (Fig 1C, blue). All the data were spatially smoothed for each participant. I masked the information map using the same mask used in the identical object analysis. Under the null hypothesis of no relatedness between the value distance model RDM and neural RDM, the bias maps calculated by the permutation test were subjected to a non-parametric one-sample t-test across participants. The statistical threshold was set at $p < 0.05$ and FWE-corrected. The table reporting the clusters and their information was constructed using FSL's cluster function (all sessions: Fig. 3D, whole brain map; Supplementary Fig. 7B, cluster list; Supplementary Table 5. Hungry state: whole brain map; Supplementary Fig. 8B,

cluster list; Supplementary Table 6. Satiated state: whole brain map; Supplementary Fig. 9B, cluster list; Supplementary Table 7).

2.5.6. Value distribution adjustment

Because differences in value distance are accompanied by identity differences, to rigorously estimate the identical object specificity effect, the value distance must be controlled before being subjected to statistical tests. Although several studies have applied the representational dissimilarity decomposition technique to dissociate multiple representations using a GLM (Alink et al., 2015; Chikazoe et al., 2014), I did not adopt this method because associations between object specificity and value distance (see Fig. 3B) might cause a multicollinearity problem. Alternatively, I developed a method in which the value distance distribution was adjusted between pairs of identical and non-identical objects. This novel method was inspired by the basic idea of age adjustment, which has been broadly used in the fields of epidemiology and public health (Neison, 1844; Curtin & Klein, 1995). In the age adjustment technique, the difference in age distributions between two different aged populations, is controlled for. For example, if the specific population with cancer has a higher mortality rate than the healthy population, it is possible that this effect is confounded by the age distribution difference between the two populations since the population with cancer is older than the healthy population. In this technique, a weighting approach was adopted to control for this confounding

effect, as described below. In the population with cancer, younger groups are weighted more heavily, and older groups are weighted less heavily to match their age distribution with that of the healthy population. In detail, the mortality rate in the population with cancer is multiplied by the proportion of the age distribution in the healthy population separately for each age distribution bin and then summed across the age distribution. This “age-adjusted” mortality rate allows an adequate comparison between two different aged populations. In particular, I employed a weighting approach in age adjustment that controls for the difference in age distributions between two populations in order to adjust for the difference in the value distance distributions between identical and non-identical pairs in our data. First, the proportions of trial pairs of identical and non-identical object pairs were separately calculated for each value distance (see Fig. 4A, right column). To control for the effect of value distance, I adjusted for the difference in the value distance distribution between identical and non-identical object pairs by replacing the value distance distribution of the non-identical object pair with that of the identical object pair. On the other hand, after regressing out the white matter noise component as described above for the identical object analysis, the percentile of the rank of the elements of the neural RSM were averaged as a neural representational similarity metric, separately, for each value distance bin and for identical and non-identical pairs. Furthermore, I calculated the product between the proportion of trial pairs and the neural representational similarity metric for each value distance bin. The products were summed across all value distances, resulting in

a summary statistic of representational similarity of identical and non-identical object pairs. These procedures were similarly repeated, once more, except that the value distance distribution of the identical object pair was replaced by that of the non-identical object pair. The two summary metrics from the different distribution adjustment directions were averaged. In this way, I obtained neural similarity metrics while controlling for the value effect in identical and non-identical object pairs in each session pair in each cerebral voxel. Because I was interested in the effect of the metabolic state change, I separately averaged these metrics across pairs of the same metabolic states (two session pairs, Fig 1C: red) and different metabolic states (four session pairs, Fig 1C: orange). Using the procedures described above, I obtained four representational similarity metrics categorized into ISS, NSS, IDS, and NDS. To estimate the difference in identical object specificity effects between the same and different metabolic state pairs, I calculated [(ISS – NSS) – (IDS – NDS)] metrics in each gray matter voxel. After these similarity metrics were spatially smoothed for each participant, they were subjected to a one-sample t-test. The statistical threshold was set at $p < 0.05$ and FWE-corrected. The table reporting the clusters and their information was constructed using FSL's cluster function (Fig 4B, whole brain map; Supplementary Fig 10, cluster list; Supplementary Table 8). To visualize representations of identical objects and the effect of metabolic states on them, I conducted an ROI analysis. Each representational similarity metric in the four categories described above was spatially smoothed. Using a leave-one-participant-out procedure and the

l OFC ROI as in the identical object analysis, I extracted four representational similarity metrics (i.e., ISS, NSS, IDS, and NDS) from the averaged data from a spherical searchlight (radius = 5 mm) region around the peak in each participant (Fig 4C). I conducted a paired t-test on identical object specificity effects in the same and different states and the difference between these effects in the data. A Kolmogorov-Smirnov test was performed prior to the t-test to confirm the normal distribution of the data. To visualize the value representations in this searchlight region, I computed the neural representational similarities in relation to the value distance. In all within-session pair data (Fig 1C, blue), I extracted neural RDM elements based on value distance (0, 1, 2, > 2) without value distribution adjustment in this region, and then averaged them in each bin. All the data were spatially smoothed (Fig. 4D). I selected these bins as the larger value distance was relatively rare, and thus, the larger value distance data were derived from a few participants. These RDM elements in all value distance bins (from 0 to 9) are plotted in Supplementary Fig. 12. Furthermore, I investigated whether the food type of the visual stimuli affected the representations of object specificity in the OFC. The details of the analyses are reported in the Supplementary Results.

2.5.7. Nasal oxytocin administration

All data in this study were derived from a project aimed at investigating the effect of nasal oxytocin on neural responses to subjective value changes. Participants took nasal oxytocin on

an experimental day and placebo on the other day, in a randomized crossover, double-blind manner. Previous studies have demonstrated that nasal oxytocin administration alters food intake in healthy male adults (Lawson et al., 2015; Ott et al., 2013). Before the experiment began, six 0.1-mL puffs (three per nostril) of intranasal oxytocin (Syntocinon) or placebo and vehicle were administered intranasally, amounting to a total dose of 24 IU oxytocin (0.6 mL). Moreover, the psychological and physiological effects of the drug were examined. Before and after scanning, participants completed 10-cm visual analogue scales (VAS) of appetite including the question: 'How hungry do you feel right now?' (Supplementary Fig. 13). Thus, participants completed the VAS rating four times on each experimental day. Before and after a meal, the experimenter used a lancet to obtain a drop of blood from the side of the finger and measured the plasma glucose level using a glucometer (Supplementary Fig. 14). Because I did not observe any effects of oxytocin administration in behavioral, physiological, and neural data analysis, I analyzed the data assuming that the same experimental procedure was repeated on both days. The results of these analyses are provided in Supplementary Results.

3. **Results**

3.1.Behavioral Results

Twenty-four healthy male adults completed four fMRI sessions in two days before and after a meal with a total of 128 food photos presented during the fMRI sessions (Supplementary Table 1). Participants were instructed to rate the palatability of the food using a 10-point scale during the presentation of the photo in each trial (Fig. 1A). Palatability was used as a metric of the subjective value of food. The rating scores were subjected to a two-way repeated measures ANOVA with the factors of experimental day (days 1 and 2) and meal (before and after a meal) (day: $F_{1,23} = 4.06$, $p = 0.056$; meal: $F_{1,23} = 81.37$, $p < 0.001$; day × meal: $F_{1,23} = 2.45$, $p = 0.13$). Post-hoc simple t-tests revealed that while eating a meal significantly decreased the subjective value of the food ($t_{23} = 9.02$, $p = 1.03\times 10^{-8}$), the experimental day did not have a significant effect ($t_{23} = -2.02$, $p > 0.05$). Although having eaten a meal decreased the palatability for all foods, the experimental day (first or second day) had no effect on palatability (Supplementary Fig. 1).

3.2.Representations of identical objects in the OFC is sensitive to metabolic state change

The first aim of this study was to determine which brain region contributed to identical object representations and its association with the metabolic state change. To investigate the effect of metabolic state change, the representational similarity metrics of the identical and non-identical object pairs were averaged separately for pairs of the same metabolic state (two session pairs, Fig. 1C: red) and different metabolic states (four session pairs, Fig. 1C: orange) (See Supplementary Fig.5 in detail). To estimate the object specificity effects in pairs from the same and different metabolic states, the differences in these metrics between identical and non-identical object pairs were subjected to a one-sample t-test for each voxel, separately, in two metabolic state pairs. I found that identical objects were similarly represented in the lateral OFC only in pairs of the same metabolic state, whereas identical objects were similarly represented in the V1 and FG in pairs of both the same and different metabolic states (Fig. 2A).

To visualize representations of identical objects and the effect of metabolic states on them, I conducted an ROI analysis in V1, FG, and lateral OFC. Representational similarity metrics for the four categories described above are plotted in Fig. 2B. I observed the object specificity effect in the same metabolic state pair in all three regions (V1: $t_{23} = 12.11$, $p = 3.68 \times 10^{-11}$, FG: $t_{23} = 14.54$, $p = 8.76 \times 10^{-13}$, l OFC: $t_{23} = 5.10$, $p = 7.28 \times 10^{-5}$) and the object specificity effect in a different metabolic state pair in V1 and FG, but not in OFC (V1: $t_{23} = 12.45$, $p = 2.10 \times 10^{-11}$, FG: $t_{23} = 12.88$, $p = 1.07 \times 10^{-11}$, l OFC: $t_{23} = -0.43$, $p > 0.05$). Furthermore, I found a difference in the object specificity effect across metabolic states only in the OFC (V1: $t_{23} = -0.45$, $p > 0.05$, FG: $t_{23} = -0.99$, $p > 0.05$, l OFC: $t_{23} = 3.34$, $p = 0.0028$). These results demonstrated that the representations of identical objects in V1 and FG were not affected by the metabolic states, suggesting that object identity was represented in the primary and higher visual cortices. On the other hand, representations in the OFC do not reflect pure visual features, but rather the sensory qualities of the expected outcome, which are adaptively altered, reflecting changes in the metabolic states.

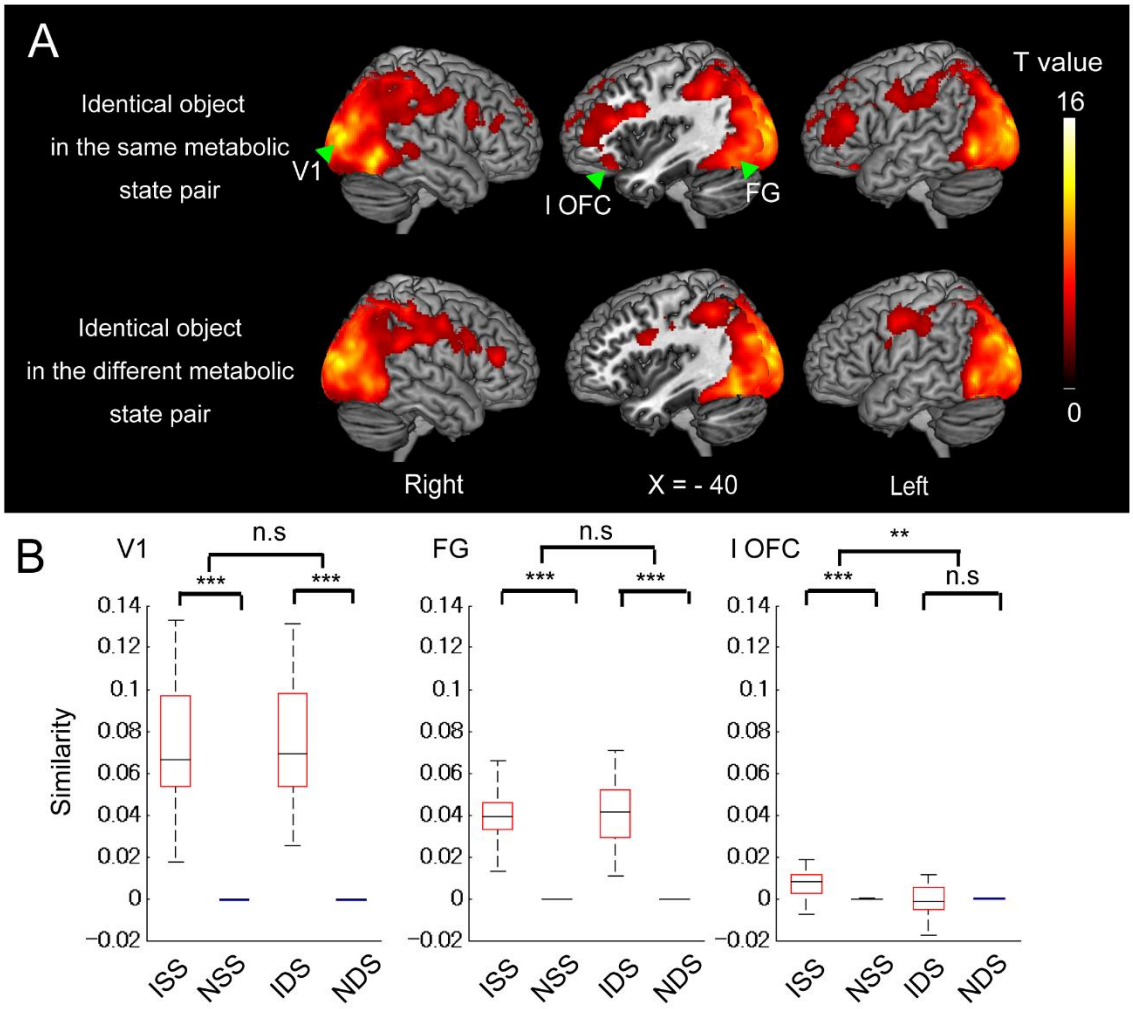


Fig. 2. Representations of identical objects in the same and different metabolic state pairs

(A) The brain regions that showed higher similarity for identical objects in comparison to non-identical objects in the same (upper) and different (lower) metabolic state pairs. Color denotes t value thresholds at $p < 0.05$ (FWE). N = 24. (B) Neural representational similarity in three regions of interest: V1, primary visual area; FG, fusiform gyrus; and l OFC, lateral orbitofrontal cortex. N =24. ISS, identical object pair in the same metabolic state; NSS, non-identical object

pair in the same metabolic state; IDS, identical object pair in the different metabolic state; NDS, non-identical object pair in the different metabolic state; n.s. not significant. Boxes represent the median and 25^{th}/75^{th} percentiles, and whiskers represent the minimum and maximum. ***p < 0.001; **p < 0.01; Bonferroni-corrected.

3.3.State-sensitive representations of identical objects in the OFC were not explained by a change in subjective value

The results described above suggest that information about identical objects is not uniformly represented across brain regions; instead, each brain region represents different aspects of identical objects. Specifically, the representations of identical objects in the OFC are affected by metabolic state changes. Although identity refers to the physical and/or categorical aspect of a food item itself, sensory qualities refer to the uniqueness of the perceptual experience evoked by the item. This implies that a change in the metabolic state alters the inferred sensory qualities, but not the object identity. Thus, the state-sensitive representations of identical objects in the OFC, as described above, might refer to inferred sensory qualities other than object identity. However, considering the fact that previous studies have consistently demonstrated the involvement of the OFC in subjective value representation (Chikazoe et al., 2014; Gallagher et al., 1999; Gottfried et al., 2003; Kringelbach et al., 2003; Padoa-Schioppa & Assad, 2006, 2008; Stalnaker et al., 2014; Tremblay & Schultz, 1999), these results may be explained by a change in the palatability of food, accompanied by a metabolic state change (Supplementary Fig. 1). Moreover, the object specificity effect on neural representations might be explained by the difference in the value distance distribution between identical and non-identical object pairs (Fig. 3A). Indeed, the results of the two-way repeated-measures ANOVA

on the value distance, with object specificity and metabolic state as factors, revealed that the main effects of both factors were significant (object specificity: $F_{1,23}$ = 59.27, $p < 0.001$; metabolic state: $F_{1,23} = 46.35$, $p < 0.001$), as was their interaction ($F_{1,23} = 45.03$, $p < 0.001$) (Fig. 3B). The post hoc paired t-tests revealed significant object specificity effects in both the same and different metabolic state pairs (same state: $t_{23} = 9.064$, $p < 0.001$; different state: $t_{23} = 2.88$, $p = 0.025$) (Fig. 3B). These results indicate that the difference in the value distance distribution may explain the difference in the neural representational dissimilarities between identical and non-identical object pairs. To explore this possibility, I first investigated whether the OFC represented subjective values, in both the univariate and multivariate analyses. In a GLM approach, parametric modulation analysis revealed that subjective values were coded in multiple brain regions, including the lateral and medial OFC, superior frontal gyrus, and parietal lobules (Fig 3C; the whole brain map, see Supplementary Fig. 7A; the cluster list, see Supplementary Table 2). Next, I conducted a representational similarity analysis of the value representations. To ensure data independence from the analysis of identical object representations, I used the data in only the within-session pairs (Fig. 1C; blue). After the searchlight analysis, the relatedness of the neural and value distance model RDMs was estimated based on the correlation coefficients between these RDMs (see Supplementary Fig. 6). With the permutation procedure, I computed where the actual correlation fell among the simulated null distribution of correlations, which was used for the summary statistic. In the group analysis, under the null hypothesis of no relationship between the neural and model RDMs, these summary statistics were subjected to a one-sample t-test for each voxel.

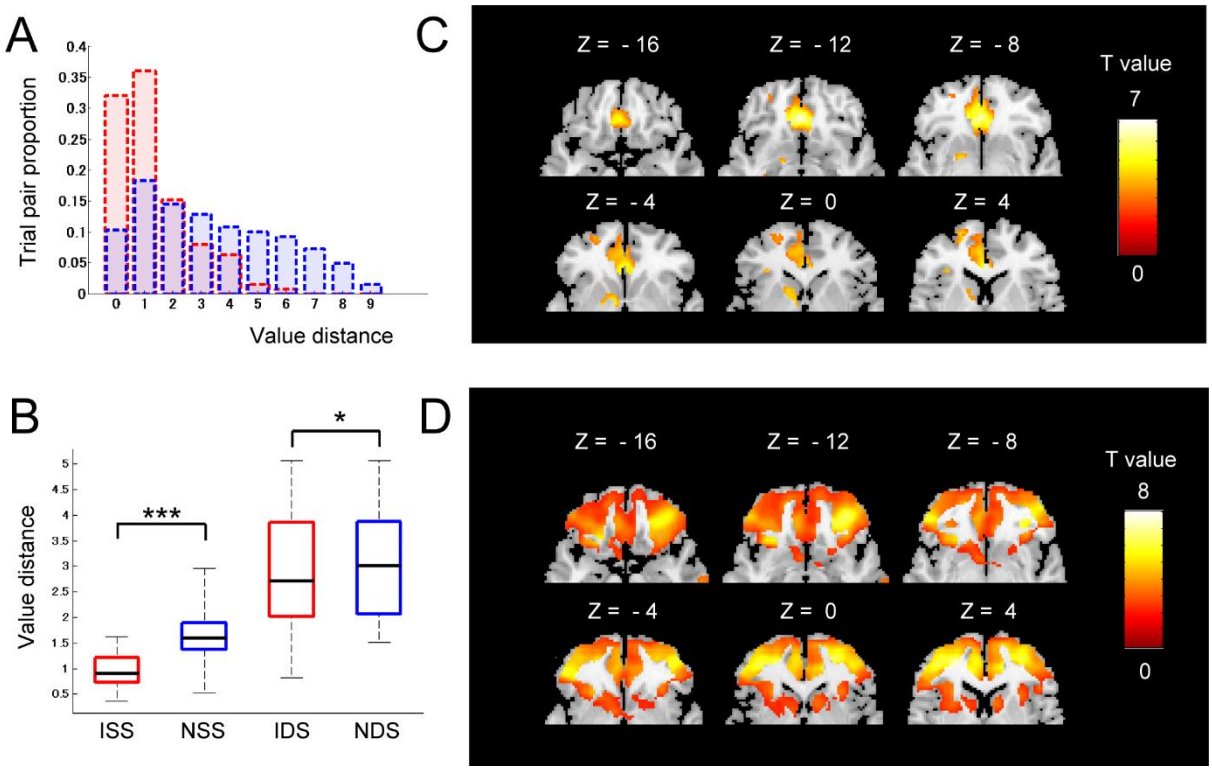


Fig. 3. Value representations in the OFC

(A) Proportion of value distance in identical (red) and non-identical (blue) object pairs in one representative session pair. Note that the value distance of the identical object pairs is smaller than that of the non-identical object pairs. Red: identical object pair; blue: non-identical object pair. (B) Boxplots show the value distance in each identical and non-identical object pairs and in each pair of the same and different metabolic states. Red: identical object pair; blue: non-identical object pair. ISS, identical object pair in the same metabolic state; NSS, non-identical object pair in the same metabolic state; IDS, identical object pair in different metabolic state; NDS, non-identical object pair in different metabolic state. N = 24. Boxes represent the median and $25^{th}/75^{th}$ percentiles, and whiskers represent the minimum and maximum. ***$p < 0.001$;

*p < 0.05; Bonferroni-corrected. (C) Univariate analysis revealed that subjective values are represented in the medial and lateral orbitofrontal cortices (OFCs). The color denotes the t-values thresholds at p < 0.05, corrected (FWE). N = 24. (D) Multivariate analysis revealed that subjective values are represented in the medial and lateral OFCs. The color denotes the t-values thresholds at p < 0.05, corrected (FWE). N = 24.

I found that subjective values were represented in multiple brain regions, including the lateral and medial OFC, precuneus, and parietal lobules (Fig 3D; for the whole brain map, see Supplementary Fig. 7B; for the cluster list, see Supplementary Table 5). To investigate whether subjective values were similarly represented independent of metabolic states, I conducted univariate and multivariate analyses separately for the hungry or satiated state. I found value representations in the OFC in both univariate and multivariate analyses, and in both hungry and satiated states (univariate analysis in the hungry state: for the whole brain map, see Supplementary Fig. 8A; for the cluster list, see Supplementary Table 3; univariate analysis in the satiated state: for the whole brain map, see Supplementary Fig. 9A; for the cluster list, see Supplementary Table 4; multivariate analysis in the hungry state: for the whole brain map, see Supplementary Fig. 8B; for the cluster list, see Supplementary Table 6; multivariate analysis in the satiated state: for the whole brain map, see Supplementary Fig. 9B; for the cluster list,

see Supplementary Table 7). These results suggest that to rigorously estimate the effect of identical object specificity, the effect of value distance must be analytically dissociated from that of the metabolic change, at least in value-related regions such as the OFC. To this end, I investigated the effect of metabolic states on the representations of identical objects while controlling for the effect of value distance differences in cross-session pairs. I matched the distribution of non-identical and identical object pairs, and vice versa. Fig 4A shows the details of our method; however, to avoid complexity, only one matching direction procedure is shown. After the adjustment procedure was conducted for each direction, two summed adjusted representational similarities were averaged. Then, these values in six cross-session pairs were averaged separately for the same metabolic state pairs (two session pairs, Fig 1C: red) and different metabolic states (four session pairs, Fig 1C: orange), as in the identical object analysis in Fig 2. To estimate the interaction of the object specificity and metabolic states, I calculated the difference of difference of the similarities [(ISS – NSS) – (IDS – NDS)] as the summary metric. In the group analysis, the summary metric was subjected to a one-sample t-test for each voxel.

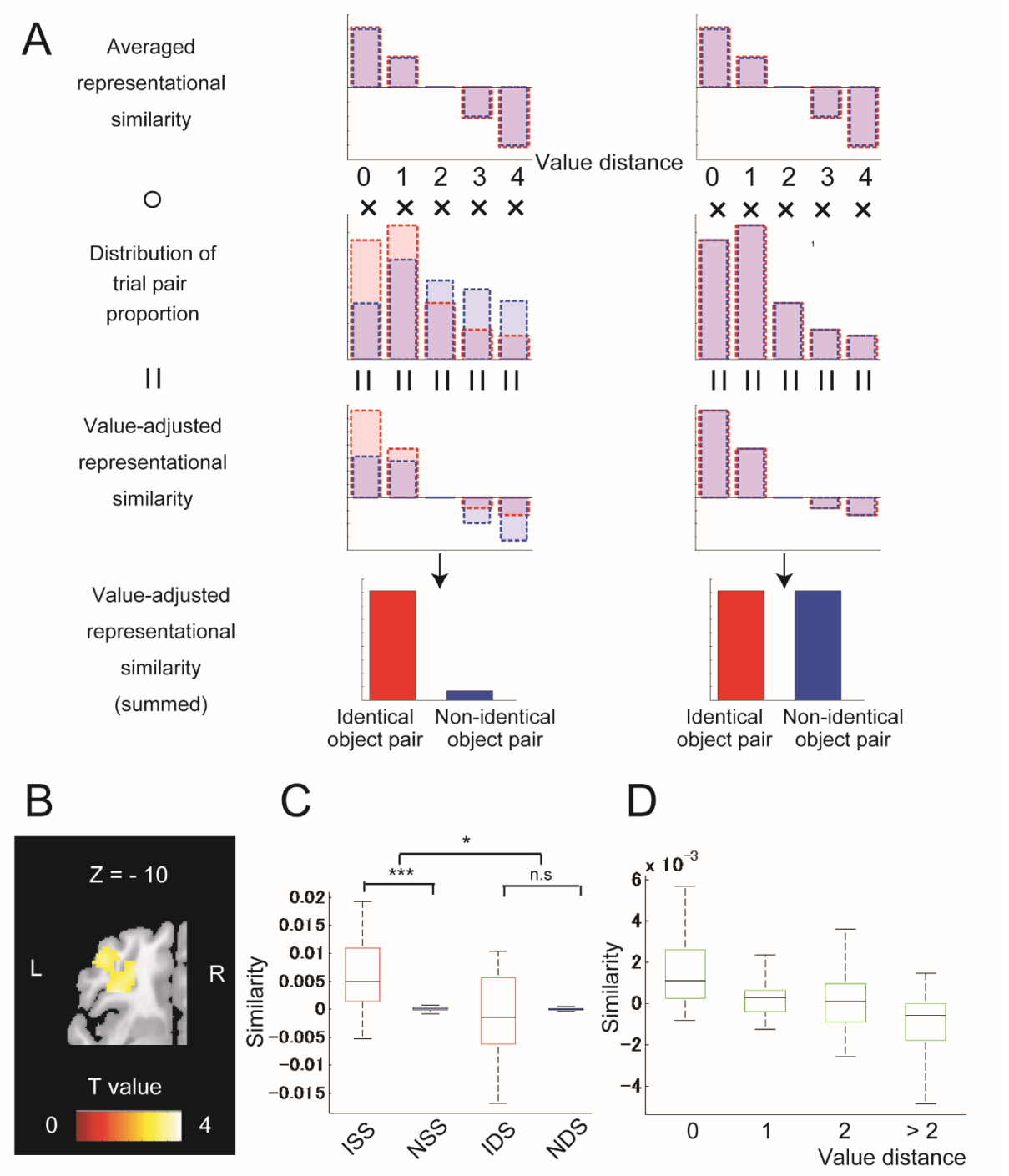


Fig. 4. Representations of the sensory qualities other than the subjective value in the OFC

(A) Schematic illustration of the distribution adjustment method for controlling for the effect of value distance. Because the value distance of identical object pairs was smaller than that of the non-identical pairs (See Fig. 3A, B), representational similarity differences associated with

identical objects, may just reflect the difference in value distance. To dissociate the effect of identical objects from that of value distance, I devised a method for controlling for the effect of value distance difference between identical and non-identical object pairs. This procedure is explained by a toy example in which the representational similarity difference between identical (red) and non-identical object pairs (blue) is fully explained by value distance. In fMRI studies, the summary metric for the effect of interest for each subject is subject to a group analysis. To calculate the summary metric, representational similarities were averaged in each identical and non-identical object pairs as shown in Fig 2. This procedure is mathematically equivalent to the way that averaged representational similarity (top row in the left column) and trial pair proportion (second row in the left column) are multiplied (third row in the left column) for each bin, and then summed across bins (bottom row in the left column). Thus, the difference between identical and non-identical object distributions in each bin (second row in the left column) might cause an artifactual effect of object specificity (bottom row in the left column). By contrast, the right column illustrates an analysis procedure that controls for the effect of value distance by distribution adjustment. I adjusted for the difference in the value distance distribution by replacing the value distance distribution of the non-identical object pairs with that of the identical object pairs (second row in the right column). With this procedure, the representational similarity difference between identical and non-identical object pairs diminished if the actual representational similarity difference did not exist (bottom row in the right column). This method provides two distribution adjustment directions: from the identical object pair distribution to the non-identical one, and vice versa. Hence, I averaged the summary metrics obtained by these procedures. After the distribution adjustment, I conducted a one-sample t-test on the difference of the object specificity between the same and different metabolic states. (B) Representational similarity analysis revealed that the left lateral OFC was associated with the difference of object specificity between the same and different metabolic

states. Colors represent the t values thresholds at $p < 0.05$, corrected (FWE). N = 24. (C) Neural representational similarity in l OFC. l OFC, lateral orbitofrontal cortex; ISS, identical object pair in the same metabolic state; NSS, non-identical object pair in the same metabolic state; IDS, identical object pair in different metabolic state; NDS, non-identical object pair in different metabolic state. N = 24. Boxes represent the median and 25^{th} /75^{th} percentiles and whiskers represent the minimum and maximum. *** $p < 0.001$; * $p < 0.05$; Bonferroni-corrected. (D) Neural representational similarity in value distance bin in the lateral OFC. N = 24. Boxes represent the median and 25^{th}/75^{th} percentiles and whiskers represent the minimum and maximum.

Even after applying the distribution adjustment, I found a significant difference in identical object specificity between the same and different metabolic states in the left OFC and lateral frontal area (Fig 4B; for the whole brain map, see Supplementary Fig. 10; for the cluster list, see Supplementary Table 8). To visualize the representations of identical objects and the effect of metabolic states on them under the value distribution control, I conducted ROI analysis in the lateral OFC (Fig. 4C). I adopted a leave-one-subject-out procedure (Chikazoe et al., 2019) in this analysis in the same way as in the identical object analyses in Fig. 2B (see details in Methods). This analysis revealed a significant effect of object specificity in the pairs of the same metabolic state ($t_{23} = 4.29$, $p = 5.46 \times 10^{-4}$) but not in pairs of different metabolic states ($t_{23} = -0.55$, $p > 0.05$). Furthermore, there was also a significant difference in object specificity between the same and different metabolic state pairs ($t_{23} = 2.64$, $p = 0.015$). Additionally, to confirm that object specificity effects were similar in both hungry–hungry and satiated–satiated pairs in the same metabolic state, I investigated whether the effect of object specificity was different between hungry–hungry and satiated–satiated pairs. Because no voxel

survived in this analysis, I averaged the two metrics in hungry–hungry and satiated–satiated pairs for the same metabolic state. To visualize the coexistence of value representations and the interaction between object specificity and metabolic states in this region, I plotted the neural representational similarity without the value distribution adjustment for each value distance bin (Fig. 4D). This revealed a monotonic decrease in representational similarity as the value distance increased, indicating that the same portion of the lateral OFC was associated with the interaction between object specificity and metabolic states, as well as the value. Additionally, the results of further analyses of the effect of food type on these object specificity representations are reported in Supplementary Results and Figures (Supplementary Fig. 11).

These results demonstrate that representations of identical objects in the lateral OFC were altered by metabolic state changes, which was not explained by the value representations. These results suggest that the lateral OFC integrates the visual features and infers the gustatory features, reflecting an inferred gustatory experience. I conclude that sensory qualities and subjective values of the expected outcomes co-exist in the lateral OFC.

4. Discussion and Conclusions

In this study, I explored the neural representation of object specificity and its relationship with the metabolic states, as well as the subjective values. During fMRI scanning, participants rated food palatability based on gustatory information inferred from the food photos. By applying representational similarity analysis, I revealed that object identity was represented in the primary visual cortex and FG. Although neuroimaging studies have rarely reported the involvement of the OFC in identity representations, I found that representations of identical objects were constantly maintained in the lateral OFC only under the same metabolic

conditions. Interestingly, representations of identical objects were altered when the metabolic state changed. I also found that the OFC represents the subjective value of food objects, which is consistent with previous studies (Chikazoe et al, 2014; Gallagher et al, 1999; Gottfried, O'Doherty, & Dolan, 2003; Kringelbach et al, 2003; Padoa-Schioppa & Assad, 2006, 2008; Stalnaker et al., 2014; Tremblay & Schultz, 1999). Importantly, the altered representation of identical objects due to a change in the metabolic state was not explained by the subjective value, which indicates the existence of integrated representations of object-specific information and metabolic states in addition to value representations in the OFC. These results suggest that the OFC does not represent pure visual features or object identity, but rather represents subjective value as well as sensory qualities of the expected outcome, which are adaptively altered, reflecting a change in the metabolic state.

4.1.Representations of the identical objects in the OFC

Animal studies have consistently demonstrated that reward identity is represented by the OFC neurons. For example, rodent electrophysiological studies have indicated that OFC neurons encode specific features, which constitute the odor identity (Schoenbaum & Eichenbaum, 1995; Stalnaker et al., 2014). Similarly, single-unit recordings in monkeys revealed that OFC neurons exhibited differential responses to different odors (Cichy et al, 2011) and tastes (Padoa-Schioppa, 2007). However, the human neuroimaging literature lacks evidence for identity representations in the OFC, although several studies have revealed that the interaction of reward and identity might reside in the OFC (Howard et al, 2015; Klein-flügge et al., 2013). Identity representations are generally assumed to be associated with specific features with which an animal or human subject can discriminate one stimulus from multiple stimuli (Cichy et al., 2011; Hung et al, 2005; Kriegeskorte et al, 2007). In this study, food objects could be identified by visual features, such as the color or shape, as well as categories. Consistent with

this, representations of identical objects were similarly represented in the visual cortices, including V1 and the FG (Fig. 2). Furthermore, these representations were stable across different metabolic states, which is a requisite characteristic for "object identity' (DiCarlo et al, 2012). By contrast, representations of identical objects in the OFC were sensitive to metabolic state changes (Fig. 2). This indicates that the OFC does not represent pure visual features, but instead represents inferred sensory qualities that are adaptively altered, reflecting changes in the metabolic states. These results suggest that the OFC appears to represent reward identity, depending on the experimental setting; however, our results clearly demonstrate that the OFC does not represent mere identity, but rather represents inferred sensory qualities as well as value.

4.2.Value representations of the food in the OFC

In the present study, I demonstrated that the lateral OFC was associated with the interaction between object specificity and metabolic states (Fig. 4B and 4C). This interaction was evident even after adjusting for the difference in palatability distribution between identical and non-identical object pairs. Considering that palatability is just one type of subjective value; this interaction effect may reflect other aspects of value beyond palatability. For example, several human neuroimaging studies have demonstrated that the OFC represents the subjective value of different types of stimuli, such as monetary, erotic, and social rewards (Levy and Glimcher, 2011, 2012; Sescousse et al., 2010). This raises the possibility that the lateral OFC may be involved in representing other value components such as retail value, or entirely different cognitive processes (e.g., less/more attention paid to food stimulus in the hungry state). However, an objective value such as retail value cannot explain this interaction, as the objective value is not affected by internal metabolic states. Furthermore, the task to be completed during

the fMRI scanning is the palatability ratings of food photos; thus, it is unlikely that erotic or social value is represented. Cognitive components, such as attention, cannot explain the interaction effect, as attention will be similarly recruited for identical and non-identical object pairs. Although I cannot exclude the possibility that the interaction effect in the lateral OFC may reflect a type of value that was not reported but represented, the current result suggests that the lateral OFC is associated with object representations that are affected by the metabolic states.

4.3.Relationship with cognitive map theory in the OFC

Recently, the cognitive map theory was revisited in order to interpret the complex roles of the OFC (Behrens et al., 2018; DiCarlo et al., 2012; Wikenheiser & Schoenbaum, 2016; Wilson et al, 2014). In 1948, that animals construct map-like representations of causal associations, posited by Tolman, was named the cognitive map (Tolman, 1948). In this map, the objects represent not only the physical properties but also their causal outcomes, which must be inferred rather than directly observed. Such abstract representations enable the derivation of novel means of achieving the outcomes. In this study, participants were asked to infer the palatability of food from a photo. I revealed that representations of identical objects in the OFC were constantly maintained only when the metabolic states were matched. This result is in line with the cognitive map theory. It has been posited that the cognitive map encodes rich representations of outcomes that are associated with internal/external states (Wikenheiser & Schoenbaum, 2016; Wilson et al., 2014). Although such representations are distributed in multiple brain regions besides the OFC (Wikenheiser & Schoenbaum, 2016), recent studies have indicated that the OFC encodes integrated representations of all relevant internal and external stimuli or features that define a particular situation in the world (Bradfield et al, 2015; Schuck et al, 2016, Wimmer & Büchel, 2019). Consistent with this, our results revealed such

integrated representations in highly cognitive areas, including the lateral orbitofrontal and prefrontal cortices (Fig 4B; for the whole brain map, see Supplementary Fig. 10; for the cluster list, see Supplementary Table 8), suggesting that representations in these regions may constitute a cognitive map. Assuming that a cognitive map may represent all possible combinations for specific objects in specific conditions, in such a map, even the same food objects can be placed distantly, if the metabolic states are different. Alternatively, changes in the metabolic state may alter the cognitive map. In this case, the same food objects in different states may be represented in an incomparable manner. In either case, the representations of the same food objects in the different metabolic states would be dissimilar, whereas those in the same metabolic states would be similar (Fig. 4B and 4C).

4.4.The OFC as the gustatory cortex

The gustatory pathway is presumed to be similar in humans and nonhuman primates. Gustatory information is conveyed from taste receptors to the ventral posterior medial nucleus of the thalamus (Beckstead et al, 1980). Although the primary projection terminates in the anterior insula and frontal operculum, it is thought that the secondary or tertiary projections terminate in the OFC (Petrides, M. & Pandya, D.N., 1994). Consistent with this, monkey electrophysiology and human neuroimaging studies have demonstrated the involvement of the OFC in gustatory processing (Padoa-Schioppa, 2007; Rolls & Baylis, 1994; Veldhuizen et al., 2011). Based on this evidence, the OFC is thought to play an important role in gustatory processing (Small et al., 2007). In this study, I revealed that representations of identical objects in the OFC were altered by a change in the metabolic state. Considering that even the same food may taste differently under different metabolic states, this interaction effect in the OFC represents gustatory experiences inferred from the food photo. Although it remains unclear whether this effect is specific to gustatory experiences or can be extended to other modalities,

such as visual or auditory experiences; our results illustrate that the role of the OFC is not limited to value representations.

4.5.Implications for functional subdivisions in the OFC

Functional distinctions between medial and lateral OFCs have been reported in many studies (Noonan et al, 2012; Noonan et al., 2010; Rudebeck & Murray, 2014; Wallis, 2012). Connectivity studies have revealed distinct medial and lateral orbitofrontal networks in rats (Ongür & Price, 2000), monkeys (Carmichael & Price, 1996), and humans (Kahnt et al, 2014), indicating a functional dissociation between the medial and lateral OFCs across species. Consistent with this, recent studies have demonstrated that although the lateral OFC is required for evaluating options (Rudebeck & Murray, 2011; Walton et al., 2010), the medial OFC is associated with choices among objects based on value comparisons (Boorman et al, 2009; FitzGerald et al, 2009; Hunt et al., 2012; Noonan et al., 2010; Rudebeck & Murray, 2011; Strait et al, 2014). Indeed, a recent multivariate pattern analysis of human fMRI revealed identity-specific value representations of food in the lateral OFC and identity-general value representations of food in the ventromedial prefrontal cortex (Howard et al., 2015). Moreover, another multivariate analysis of fMRI data demonstrated that although food value is represented in neural activity patterns in both the medial and lateral OFC, only the lateral OFC represents the elemental nutrient content (Suzuki et al, 2017). These results suggest that the value of objects is represented in both the medial and lateral OFCs, whereas more complicated associations that cannot be explained by value alone are represented in the lateral OFC. Consistent with this, I replicated the value representations in the medial and lateral OFCs (Fig. 3C, D). I further demonstrated that object-specific features represented in the lateral OFC were not invariant, but rather, were flexibly altered, reflecting the internal metabolic states. These results are in line with previous connectivity studies showing sensory orbitofrontal network in

the primate lateral OFC, which is characterized by connections with the gustatory, olfactory, and visual cortices (Carmichael & Price, 1996). In this network, the visual information of objects converges with the olfactory, gustatory, and visceral inputs (Rudebeck & Murray, 2011). This complex visual–chemovisceral convergence might enable the lateral OFC to play an important role in inferring specific outcomes based on integrated multidimensional information.

In summary, this study revealed that the OFC represents the inferred perceptual experience linked with the current internal state, which supports flexible decision-making. These results support the idea that representations in the OFC have unique flexible functions required for updating an individual's moment-to-moment values based on the current internal state.

References

Alink , A., Walther, A., Krugliak, A., van den Bosch, J. J. F., Kriegeskorte, N., 2015. Mind the drift - improving sensitivity to fMRI pattern information by accounting for temporal pattern drift. BioRxiv, 032391. https://doi.org/10.1101/032391.

Beckstead, R. M., Morse, J. R., Norgren, R., 1980. The nucleus of the solitary tract in the monkey: Projections to the thalamus and brain stem nuclei. J. Comp. Neurol. 190(2), 259–282. https://doi.org/10.1002/cne.901900205.

Behrens, T. E. J., Muller, T. H., Whittington, J. C. R., Mark, S., Baram, A. B., Stachenfeld, K. L., Kurth-Nelson, Z., 2018. What is a cognitive map? Organizing knowledge for flexible behavior. Neuron. 100(2), 490–509. https://doi.org/10.1016/j.neuron.2018.10.002.

Boorman, E. D., Behrens, T. E. J., Woolrich, M. W., Rushworth, M. F. S., 2009. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of

alternative courses of action. Neuron. 62(5), 733–743. https://doi.org/10.1016/j.neuron.2009.05.014.

Bradfield, L. A., Dezfouli, A., Van Holstein, M., Chieng, B., Balleine, B. W., 2015. Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. Neuron. 88(6), 1268–1280. https://doi.org/10.1016/j.neuron.2015.10.044.

Burke, K. A., Franz, Theresa M., Miller, Danielle N., Schoenbaum, G., 2008. The role of the orbitofrontal cortex in the pursuit of happiness and more specific rewards. Nature. 454(7202), 340–344. https://doi.org/10.1038/nature06993.

Carmichael, S. T., Price, J. L., 1996. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. J. Comp. Neurol. 371(2), 179–207.

Chikazoe, J., Lee, D. H., Kriegeskorte, N., Anderson, A. K., 2014. Population coding of affect across stimuli, modalities and individuals. Nat. Neurosci. 17(8), 1114–1122. https://doi.org/10.1038/nn.3749.

Chikazoe, J., Lee, D. H., Kriegeskorte, N., Anderson, A. K., 2019. Distinct representations of basic taste qualities in human gustatory cortex. Nat. Commun. 10(1), 1–8. https://doi.org/10.1038/s41467-019-08857-z.

Cichy, R. M., Chen, Y., Haynes, J. D., 2011. Encoding the identity and location of objects in human LOC. NeuroImage. 54(3), 2297–2307. https://doi.org/10.1016/j.neuroimage.2010.09.044.

Curtin, L. R., Klein, R. J., 1995. Direct standardization (age-adjusted death rates). Healthy People 2000 Stat. Notes. (6), 1–10.

DiCarlo, J. J., Zoccolan, D., Rust, N. C., 2012. How does the brain solve visual object recognition? Neuron. 73(3), 415–434. https://doi.org/10.1016/j.neuron.2012.01.010.

FitzGerald, T. H. B., Seymour, B., Dolan, R. J., 2009. The role of human orbitofrontal cortex in value comparison for incommensurable objects. J. Neurosci. 29(26), 8388–8395. https://doi.org/10.1523/JNEUROSCI.0717-09.2009.

Gallagher, M., McMahan, R. W., Schoenbaum, G., 1999. Orbitofrontal cortex and representation of incentive value in associative learning. J. Neurosci. 19(15), 6610–6614. https://doi.org/10.1523/JNEUROSCI.19-15-06610.1999.

Gottfried, J. A., O'Doherty, J., Dolan, R. J., 2003. Encoding predictive reward value in human amygdala and orbitofrontal cortex. Science, 301(5636), 1104–1107. https://doi.org/10.1126/science.1087919.

Gremel, C. M., Costa, R. M., 2013. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nat. Commun. 4, 1–12. https://doi.org/10.1038/ncomms3264.

Grill-Spector, K., Malach, R., 2004. The human visual cortex. Annu. Rev. Neurosci. 27(1), 649–677. https://doi.org/10.1146/annurev.neuro.27.070203.144220.

Haxby, J. V, Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., Hanke, M., Ramadge, P. J., 2011. A common, high-dimensional model of the representational space in human ventral temporal cortex. Neuron. 72(2), 404–416. https://doi.org/10.1016/j.neuron.2011.08.026.

Howard, J. D., Gottfried, J. A., Tobler, P. N., Kahnt, T., 2015. Identity-specific coding of future rewards in the human orbitofrontal cortex. Proc. Natl. Acad. Sci. 112(16), 5195–5200. https://doi.org/10.1073/pnas.1503550112.

Hung, C. P., Kreiman, G., Poggio, T., Dicarlo, J. J., 2005. Fast readout of object identity from macaque inferior temporal cortex. Science. 310(5749), 863–866. https://doi.org/10.1126/science.1117593.

Hunt, L. T., Kolling, N., Soltani, A., Woolrich, M. W., Rushworth, M. F. S., Behrens, T. E. J., 2012. Mechanisms underlying cortical activity during value-guided choice. Nat. Neurosci. 15(3), 470–476. https://doi.org/10.1038/nn.3017.

Izquierdo, A., 2004. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. J. Neurosci. 24(34), 7540–7548. https://doi.org/10.1523/JNEUROSCI.1921-04.2004.

Klein-flügge, M. C., Barron, H. C., Brodersen, K. H., Dolan, R. J., Edward, T., & Behrens, J. (2013). Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. J. Neurosci., 33(7), 3202–3211. https://doi.org/10.1523/JNEUROSCI.2532-12.2013.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. Proc. Natl. Acad. Sci., 103, 3863–3868. https://doi.org/10.1073/pnas.0600244103.

Kriegeskorte, N., Formisano, E., Sorger, B., Goebel, R., 2007. Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc. Natl. Acad. Sci., 104(51), 20600–20605. https://doi.org/10.1073/pnas.0705654104.

Kriegeskorte, N., 2008. Representational similarity analysis – connecting the branches of systems neuroscience. Front. Syst. Neurosci., 2, 1–28. https://doi.org/10.3389/neuro.06.004.2008.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., Baker, C. I., 2009. Circular analysis in systems neuroscience: The dangers of double dipping. Nat. Neurosci., 12(5), 535–540. https://doi.org/10.1038/nn.2303.

Kringelbach, M. L., O'Doherty, J., Rolls, E. T., Andrews, C., 2003. Activation of the human orbitofrontal cortex to a liquid food stimulus is correlated with its subjective

pleasantness. Cereb. Cortex, 13, 1064–1071. https://doi.org/10.1093/cercor/13.10.1064.

Kringelbach, M. L., 2005. The human orbitofrontal cortex: linking reward to hedonic experience. Nat. Rev. Neurosci., 6, 691-702. https://doi.org/10.1038/nrn1747.

Lawson, E. A., Marengi, D. A., Desanti, R. L., Holmes, T. M., Schoenfeld, D. A., Tolley, C. J., 2015. Oxytocin reduces caloric intake in men. Obesity, 23(5), 950–956. https://doi.org/10.1002/oby.21069.

Levy, D. J., & Glimcher, P. W., 2011. Comparing apples and oranges: Using reward-specific and reward-general subjective value representation in the brain. J. Neurosci., 31(41), 14693–14707. https://doi.org/10.1523/JNEUROSCI.2218-11.2011.

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

McCandliss, B. D., Cohen, L., Dehaene, S., 2003. The visual word form area: Expertise for reading in the fusiform gyrus. Trends Cogn. Sci., 7(7), 293–299. https://doi.org/10.1016/S1364-6613(03)00134-7.

McNamee, D., Rangel, A. O'Doherty, J.P., 2013. Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. Nat. Neurosci. 16, 479–485. https://doi.org/10.1038/nn.3337.

Mummery, C. J., Patterson, K., Hodges, J. R., Price, C. J. 1998. Functional neuroanatomy of the semantic system: Divisible by what? J. Cogn. Neurosci, 10(6), 766–777. https://doi.org/10.1162/089892998563059.

Neison F. G. P., 1844. On a method recently proposed for conducting inquiries into the comparative sanatory condition of various districts. J. R. Stat. Soc. (the Royal Statistical Society), 7, 40–68. https://doi.org/10.2307/2337745.

Noonan, M. P., Walton, M. E., Behrens, T. E. J., Sallet, J., Buckley, M. J., Rushworth, M. F. S., 2010. Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proc. Natl. Acad. Sci., 107(47), 20547–20552. https://doi.org/10.1073/pnas.1012246107.

Noonan, M. P., Kolling, N., Walton, M. E., Rushworth, M. F. S., 2012. Re-evaluating the role of the orbitofrontal cortex in reward and reinforcement. European J. Neurosci., 35(7), 997–1010. https://doi.org/10.1111/j.1460-9568.2012.08023.x.

Ongür, D., Price, J. L., 2000. The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. Cereb Cortex, 10(3), 206–219. https://doi.org/10.1093/cercor/10.3.206.

Ott, V., Finlayson, G., Lehnert, H., Heitmann, B., Heinrichs, M., Born, J., Hallschmid, M., 2013. Oxytocin reduces reward-driven food intake in humans. Diabetes, 62(10), 3418–3425. https://doi.org/10.2337/db13-0663.

Padoa-Schioppa, C., Assad, J. A., 2006. Neurons in the orbitofrontal cortex encode economic value. Nature, 441(7090), 223–226. https://doi.org/10.1038/nature04676.

Padoa-Schioppa, C., 2007. Orbitofrontal cortex and the computation of economic value. Ann. N. Y. Acad. Sci., 1121, 232–253. https://doi.org/10.1196/annals.1401.011.

Padoa-Schioppa, C., Assad, J. A., 2008. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. Nat. Neurosci., 11(1), 95–102. https://doi.org/10.1038/nn2020.

Petrides, M., Pandya, D.N., 1994. Comparative architectonic analysis of the human and the macaque frontal cortex, in: Boller, F., Grafman, J. (Eds.), Handbook of Neuropsychology. Elsevier Science, Amsterdam, pp. 17-58.

Pickens, C. L., Saddoris, M. P., Gallagher, M., Holland, P. C., 2005. Orbitofrontal lesions impair use of cue-outcome associations in a devaluation task. Behav. Neurosci., 119(1), 317–322. https://doi.org/10.1037/0735-7044.119.1.317.

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

Rudebeck, P. H., Murray, E. A., 2011. Balkanizing the primate orbitofrontal cortex: Distinct subregions for comparing and contrasting values. Ann. N. Y. Acad. Sci., 1239, 1–13. https://doi.org/10.1111/j.1749-6632.2011.06267.x.

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

Rudebeck, P. H., Murray, E. A., 2014. The orbitofrontal oracle: Cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. Neuron, 84(6), 1143–1156. https://doi.org/10.1016/j.neuron.2014.10.049.

Rolls, E. T., Baylis, L. L., 1994. Gustatory, olfactory, and visual convergence within the primate orbitofrontal cortex. J. Neurosci., 14(9), 5437–5452. https://doi.org/10.1523/JNEUROSCI.14-09-05437.1994.

Schoenbaum, G., Eichenbaum, H., 1995. Information coding in the rodent prefrontal cortex. I. Single-neuron activity in orbitofrontal cortex compared with that in pyriform cortex. J. Neurophysiol, 74, 733–750. https://doi.org/10.1152/jn.1995.74.2.733.

Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., Takahashi, Y. K., Yuji, K., 2009. A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. Nat. Rev. Neurosci., 10(12), 885–892. https://doi.org/10.1038/nrn2753.

Schoenbaum, G., Takahashi, Y., Liu, T. L., Mcdannald, M. A., 2011. Does the orbitofrontal cortex signal value? Ann. N. Y. Acad. Sci., 1239, 87–99. https://doi.org/10.1111/j.1749-6632.2011.06210.x.

Schuck, N. W., Cai, M. B., Wilson, R. C., Niv, Y., 2016. Human orbitofrontal cortex represents a cognitive map of state space. Neuron, 91, 1402–1412. https://doi.org/10.1016/j.neuron.2016.08.019.

Sescousse, G., Redoute, J., Dreher, J.-C., 2010. The architecture of reward value coding in the human orbitofrontal cortex. J. Neurosci., 30, 13095–13104. https://doi.org/10.1523/JNEUROSCI.3501-10.2010.

Small, D. M., Bender, G., Veldhuizen, M. G., Rudenga, K., Nachtigal, D., Felsted, J., 2007. The role of the human orbitofrontal cortex in taste and flavor processing. Ann. N. Y. Acad. Sci., 1121, 136–151. https://doi.org/10.1196/annals.1401.002.

Smith, S. M., Nichols, T. E., 2009. Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage, 44, 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061.

Stalnaker, T. A., Cooch, N. K., McDannald, M. A., Liu, T. L., Wied, H., Schoenbaum, G., 2014. Orbitofrontal neurons infer the value and identity of predicted outcomes. Nat. Commun., 5, 1–13. https://doi.org/10.1038/ncomms4926.

Stalnaker, T. A., Cooch, N. K., Schoenbaum, G., 2015. What the orbitofrontal cortex does not do. Nat. Neurosci., 18(5), 620–627. https://doi.org/10.1038/nn.3982.

Strait, C. E., Blanchard, T. C., Hayden, B. Y., 2014. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. Neuron, 82, 1357–1366. https://doi.org/10.1016/j.neuron.2014.04.032.

Suzuki, S., Cross, L., O'Doherty, J. P., 2017. Elucidating the underlying components of food valuation in the human orbitofrontal cortex. Nat. Neurosci., 20, 1780–1786. https://doi.org/10.1038/s41593-017-0008-x.

Thorsten Kahnt, T., Chang, L.J., Park, S.Q. Heinzle, J., Haynes J-D., 2014. Connectivity-based parcellation of the human posteromedial cortex. J. Neurosci. 32, 6240-6250. https://doi.org/10.1523/JNEUROSCI.0257-12.2012.

Tolman, E. C., 1948. Cognitive maps in rats and men. Psychol. Rev. 55, 189–208. https://doi.org/10.1037/h0061626.

Tremblay, L., Schultz, W., 1999. Relative reward preference in primate orbitofrontal cortex. Nature, 398(6729), 704–708. https://doi.org/10.1038/19525.

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

Veldhuizen, M. G., Albrecht, J., Zelano, C., Boesveldt, S., Breslin, P., Lundström, J. N., 2011. Identification of human gustatory cortex by activation likelihood estimation. Hum Brain Mapp, 32, 2256–2266. https://doi.org/10.1002/hbm.21188.

Wallis, J. D., 2012. Cross-species studies of orbitofrontal cortex and value-based decision-making. Nat. Neurosci., 15, 13–19. https://doi.org/10.1038/nn.2956.

Walton, M. E., Behrens, T. E. J., Buckley, M. J., Rudebeck, P. H., Rushworth, M. F. S., 2010. Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. Neuron, 65(6), 927–939. https://doi.org/10.1016/j.neuron.2010.02.027.

West, E. A., Forcelli, P. A., McCue, D. L., Malkova, L., 2013. Differential effects of serotonin-specific and excitotoxic lesions of OFC on conditioned reinforcer

devaluation and extinction in rats. Behav Brain Res, 246, 10–14. https://doi.org/10.1016/j.bbr.2013.02.027.

Wikenheiser, A. M., Schoenbaum, G., 2016. Over the river, through the woods: Cognitive maps in the hippocampus and orbitofrontal cortex. Nature Publishing Group, 17, 513–523. https://doi.org/10.1038/nrn.2016.56.

Wilson, R. C., Takahashi, Y. K., Schoenbaum, G., Niv, Y., 2014. Orbitofrontal cortex as a cognitive map of task space. Neuron, 81, 267–278. http://dx.doi.org/10.1016/j.neuron.2013.11.005.

Wimmer, E.G., Büchel, C., 2019. Learning of distant state predictions by the orbitofrontal cortex in humans. Nat. Commun., 10, 1–11. https://doi.org/10.1038/s41467-019-10597-z.

6. Supplementary Methods

Comparison between ordinary least square (OLS) and ridge regression

A general linear model (GLM) is commonly used to estimate neural activation in the neuroimaging field. However, when trials are temporally adjacent, explanatory variables coding for respective trials might be highly correlated. In this situation, the regression coefficient estimation might be unstable in the OLS regression model. Ridge regression is well known to alleviate this problem by adding a penalty term to the loss function (but see Mumford et al., 2012). To confirm that the regression coefficients are more stably estimated by the ridge than the OLS regression model using the data of the present study, a comparison analysis was conducted. Because the same stimuli were repeatedly presented (once per session, four times in total), the regression coefficients for the respective stimuli were assumed to be similar across the experimental sessions. Based on this assumption, by splitting the data, the regression coefficients, estimated using one dataset, were used to predict activity in another independent dataset. I adopted the mean squared error (MSE) as the loss function, in this procedure. I conducted a four-fold cross-validation, which splits the data into a training dataset from the three experimental sessions and the test dataset from the remaining one (Supplementary Fig. 3A, left and 3B, upper-left). In the OLS regression model, model fitting and prediction were conducted, as shown in Supplementary Fig. 3A (right). For each voxel, the regression

coefficient for each stimulus was calculated using the training dataset. This procedure resulted in a vector with 128 elements, each of which encoded the brain activity evoked by each stimulus. Using these coefficients, the brain activity vectors in the test data were predicted. The MSE was calculated between the predicted and actual activities in the test dataset (Supplementary Fig. 3A, right). In the ridge regression model, the training and test datasets were split similarly as in the OLS regression model. However, because the ridge parameter should be empirically selected, I adopted a nested cross-validation technique (Supplementary Fig. 3B, upper panel). The training dataset in the outer fold was split into the inner training set from two experimental sessions and the validation set from one session. The range for the parameter search was 100 logarithmically spaced points ranging from 1.0E-10 to 1.0E+10. I computed the regression coefficients by ridge regression with each ridge parameter in the inner training dataset and then applied them to the validation dataset at each round of cross-validation. I selected the optimal ridge parameter to minimize the MSE in the validation dataset in each round of inner cross-validation. I calculated the geometric mean of these three ridge parameters as the optimal ridge parameter in the outer fold. Regression coefficients for the respective stimuli were estimated using the optimized ridge parameter in the outer training set (Supplementary Fig. 3B, upper left). The activity vector in the test dataset was predicted using the regression coefficients. The MSEs from four rounds of outer cross-validations were averaged for each participant. The mean values from the 24 participants are plotted in

Supplementary Fig. 4. The difference in the mean values between OLS and ridge regression was examined using a paired t-test.

Ridge parameter selection in the main analysis

The ridge parameter selection procedure used in the main analysis is shown in Supplementary Fig. 3C. The full dataset was split into a training dataset from three experimental sessions and a validation dataset from the remaining one. The range of the parameter search was the same (from 1.0E-10 to 1.0E+10) as in the confirmation analysis. With the regression coefficients extracted from the training dataset, the MSE in the validation dataset was calculated. The optimal ridge parameter was selected to minimize the MSE in the validation dataset at each round of cross-validation, and then, the optimal ridge parameter for a voxel was calculated by the geometrical mean of the four optimal parameters from the four rounds. The median of these parameters from all voxels was regarded as the optimal ridge parameter for each participant. This optimal ridge parameter was used to estimate the brain activity evoked by each food photo on a trial-by-trial basis, which was later subjected to the representational similarity analysis.

7. Supplementary Results

Comparison analysis of OLS and ridge regression

The MSE of the OLS regression was significantly higher than that of the ridge regression (t_{23} = 15.53, $p < 0.001$) (Supplementary Fig. 4). This analysis indicated that the regression coefficients for the respective stimuli were more stably estimated in the ridge regression than in the OLS regression (Supplementary Fig. 3A and 3B). Thus, I adopted the ridge regression to estimate the regression coefficients for each food photo.

Effect of the food type (main dish/dessert) in the visual stimuli

The visual stimuli in the fMRI session consisted of the main dishes (salty and savoury) and desserts (sweet) in equal proportions. I analyzed the effect of the food type on both the behavioral and neural representations in the OFC.

Supplementary Fig. 2 shows the palatability of the food that was sorted based on meal (before and after the meal), experimental day (days 1 and 2), and food type (main dish and dessert). A three-way repeated measures ANOVA was conducted with these three factors. While the interaction effect between the meal and food type was significant, no other interaction effects were significant (meal × experimental day, $F_{1,23} = 2.487$, $p > 0.05$; meal × food type, $F_{1,23} = 54.460$, $p < 0.001$; experimental day × food type, $F_{1,23} = 2.830$, $p > 0.05$; meal

× experimental day × food type, $F_{1,23} = 0.273$, $p > 0.05$). While the main effect of the meal was significant ($F_{23} = 81.286$, $p < 0.001$), the main effects of other factors were not significant (experimental day, $F_{23} = 4.042$, $p > 0.05$; food type, $F_{23} = 0.261$, $p > 0.05$). To further interpret the significant interaction effect between meal and food type, simple effect tests were conducted. Both palatability of the main dish and dessert decreased significantly after eating a meal (main dish, $t_{23} = 13.24$, $p < 0.001$; desert $t_{23} = 9.11$, $p < 0.001$). While the palatability of the main dish was significantly higher than that of the dessert before meals ($t_{23} = 3.07$, $p = 0.021$), the former was significantly lower than the latter after meals ($t_{23} = -3.51$, $p = 0.0074$). According to a proverb, "there is always room for dessert," the palatability of the desserts remains relatively high even after satiation (Supplementary Fig. 2).

Participants consumed three types of snacks (chocolate cookie, salt crackers, and rice crackers) ad libitum, immediately after eating the lunch box. I examined whether the amount of snack consumption influenced the difference in palatability of food photos before and after a meal. I used the calorie consumption of each snack as an explanatory variable. To estimate the effect of each food type individually, I used the difference in the palatability of each food type before and after the meal as the response variable. Model fitting was performed separately for each food type. As the experiments were conducted over two days, these variables were averaged across days. A GLM was applied to these data using a MATLAB function, 'glmfit'. These analyses revealed that the amount of consumption of all three snacks did not influence

palatability changes of the main dishes and desserts (chocolate cookie to the main dish, t_{23} = 0.213, $p > 0.05$; salt crackers to the main dish, $t_{23} = 0.143$, $p > 0.05$; rice crackers to the main dish, $t_{23} = 0.299$, $p > 0.05$; chocolate cookie to the dessert, $t_{23} = 0.057$, $p > 0.05$; salt crackers to the dessert, $t_{23} = 1.221$, $p > 0.05$; rice crackers to the dessert, $t_{23} = 1.741$, $p > 0.05$).

To confirm the consistency of the representations of the identical object in the OFC among different food types, representational similarity analysis was conducted using the same procedure as the main analysis, but using two distinct representational similarity matrices (RSMs) for the main dish or dessert trial pairs. I conducted region of interest (ROI) analyses using the same ROI, as shown in Fig. 4B. These analyses were repeated twice, with and without the value distribution adjustment technique. Using the same procedures as in the main analysis, I obtained four representational similarity metrics, categorized as identical object pair in the same metabolic state (ISS), non-identical object pair in the same metabolic state (NSS), identical object pair in the different metabolic state (IDS), and non-identical object pair in the different metabolic state (NDS) in the searchlight voxels used for the analysis reported in Fig. 4C. I found similar representational similarity patterns in the OFC (Supplementary Fig. 11), as found in all the food photos (Fig. 4C). Supplementary Fig. 11 shows that this similarity was consistent across four conditions: the main dish without value adjustment (A), dessert without value adjustment (B), main dish with value adjustment (C), and dessert with value adjustment (D). One-sample t-tests revealed significant effects of identical object specificity in the same

metabolic state (ISS – NSS) in all conditions (main dish without value distribution adjustment, $t_{23} = 5.03$, $p < 0.001$; dessert without value distribution adjustment, $t_{23} = 3.54$, $p = 0.0035$; main dish with value distribution adjustment, $t_{23} = 4.93$, $p < 0.001$; and desert with value distribution adjustment, $t_{23} = 3.54$, $p = 0.0035$). In contrast, one-sample t-tests revealed no significant effects of identical object specificity in the different metabolic states (IDS – NDS) in all conditions (main dish without value distribution adjustment, $t_{23} = -0.29$, $p > 0.05$; desert without value distribution adjustment, $t_{23} = -0.72$, $p > 0.05$; main dish with value distribution adjustment, $t_{23} = -0.32$, $p > 0.05$; and desert with value distribution adjustment, $t_{23} = -0.76$, $p > 0.05$). Furthermore, the differences of the differences in the similarities were significant [(ISS - NSS) - (IDS - NDS)] in all conditions (main dish without value distribution adjustment, $t_{23} = 3.34$, $p = 0.0029$; desert without value distribution adjustment, $t_{23} = 2.63$, $p = 0.015$; main dish with value distribution adjustment, $t_{23} = 3.28$, $p = 0.0033$; and desert with value distribution adjustment, $t_{23} = 2.65$, $p = 0.014$). These results indicate that the difference in the object specificity effects between metabolic state pairs in the OFC was independent of the food type.

Nasal oxytocin administration

All data in this study were derived from a project aimed at investigating the effect of nasal oxytocin on neural responses to changes in palatability. Before all experimental procedures, participants took nasal oxytocin on an experimental day and placebo on the other day, in a

randomized crossover, double-blind manner. Behavioral, physiological, and ROI analyses were conducted to investigate the effects of nasal oxytocin administration.

The feeling of hunger was measured using the visual analogue scale (VAS) before and after scanning. Supplementary Fig. 13 shows these metrics averaged across the scanning, sorted on meal (before and after meal) and drug (placebo and active drug). To investigate the effect of nasal oxytocin administration on the feeling of hunger, I conducted a two-way repeated measures ANOVA with the drug (placebo and active drug) and meal (before and after a meal). Although neither of the interaction and main effects of the drug administration were significant (drug × meal: $F_{1,23} = 0.487$, $p = 0.492$; drug: $F_{1,23} = 0.043$, $p = 0.838$), the main effect of the meal was highly significant (meal: $F_{1,23} = 326.24$, $p < 0.001$). Post-hoc simple t-tests revealed that while eating a meal significantly decreased the feeling of hunger ($t_{23} = 18.06$, $p < 0.001$), drug administration did not have a significant effect ($t_{23} = 1.21$, $p > 0.05$). Blood glucose levels were measured using a glucometer before and after the meal (Supplementary Fig. 14). To investigate the effect of nasal oxytocin administration on blood glucose levels, I conducted a two-way repeated measures ANOVA with drug (placebo and active drug) and metabolic state (before and after a meal). Both the main and interaction effects of the drug administration were not significant, while the main effect of the meal was highly significant (drug: $F_{1,23} = 1.72 \times 10^{-4}$, $p = 0.990$; meal: $F_{1,23} = 208.19$, $p < 0.001$; drug × meal: $F_{1,23} = 0.880$, $p = 0.358$). Post-hoc simple t-tests revealed that while eating a meal significantly increased the blood glucose level ($t_{23} = 14.43$, $p < 0.001$), drug administration did not have a significant effect ($t_{23} = 1.98$, $p > 0.05$). Supplementary Fig. 15 shows the palatability of the food that was sorted based on drug (placebo and active drug) and meal (before and after a meal). To investigate the effect of nasal oxytocin administration on the palatability of food, I conducted a two-way repeated measures ANOVA with these two factors. The interaction and main effects of the drug administration were not significant; however, the main effect of the meal was highly significant

(drug: $F_{1,23} = 0.038$, $p = 0.85$; meal; $F_{1,23} = 81.37$, $p < 0.001$; drug × meal; $F_{1,23} = 2.47$, $p = 0.13$). Post-hoc simple t-tests revealed that while eating a meal significantly decreased the palatability of the food ($t_{23} = 9.02$, $p < 0.001$), drug administration did not have a significant effect on palatability ($t_{23} = 0.19$, $p > 0.05$).

Next, I investigated the effects of nasal oxytocin on the neural representations of both object specificity and palatability. Supplementary Fig. 16A shows a schematic illustration of the session pair structure that was used to estimate the effect of nasal oxytocin administration on representations of object specificity. In this analysis, the experimental sessions were sorted based on drug administration (placebo or active drug). I used the different metabolic state pairs only, so that the experimental conditions could be matched by these pairs. These four pairs were labeled as placebo–placebo ('drug-free data,' shown in purple), drug–drug, or drug–placebo ('drug-treated data,' shown in green). I calculated the object specificity metrics (i.e., the difference in the representational similarity between identical and non-identical object pairs) in each session pair, as shown in Supplementary Fig. 5. These metrics, from the three-session pairs of the drug-treated data, were averaged. I conducted a one-sample t-test on the differences between these metrics using the drug-free and drug-treated data. The test statistic did not reach significance level for any voxel of the whole cerebral cortex, at $p < 0.05$ threshold, corrected (family wise error [FWE]). Supplementary Fig. 16B shows a schematic illustration of the session pair structure that was used to estimate the effect of nasal oxytocin administration on representations of the subjective value. The experimental sessions were sorted based on drug administration (placebo or active drug), as shown in Supplementary Fig. 16A. The four within-session pairs were divided into two placebo-placebo pairs (shown in purple) and two drug-drug pairs (shown in green). I calculated the value-associated metrics for each session pair, as shown in Supplementary Fig. 6, and then averaged them separately for the placebo and active drug pairs. I conducted a one-sample t-test on the difference in the value-associated

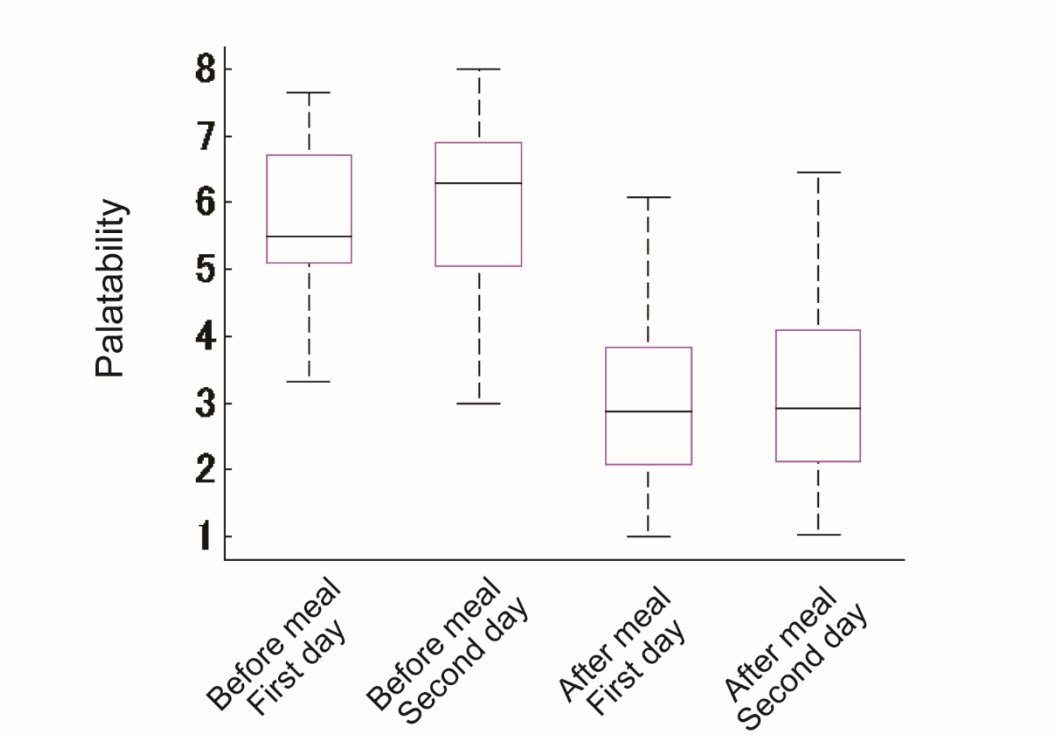
metrics between the placebo and active drug pairs. The test statistic did not reach significance level for any voxel of the whole cerebral cortex, at $p < 0.05$ threshold, corrected (FWE). Thus, the administration of this drug and metabolic state change were not associated with neural representations of object specificity and value in any brain region. I concluded that nasal oxytocin administration had no effect on the results of our main analyses.

8. References for supplementary information

Mumford, J. A., Turner, B. O., Ashby, F. G., & Poldrack, R. A., 2012. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. NeuroImage, 59, 2636–2643. doi: 10.1016/j.neuroimage.2011.08.076.

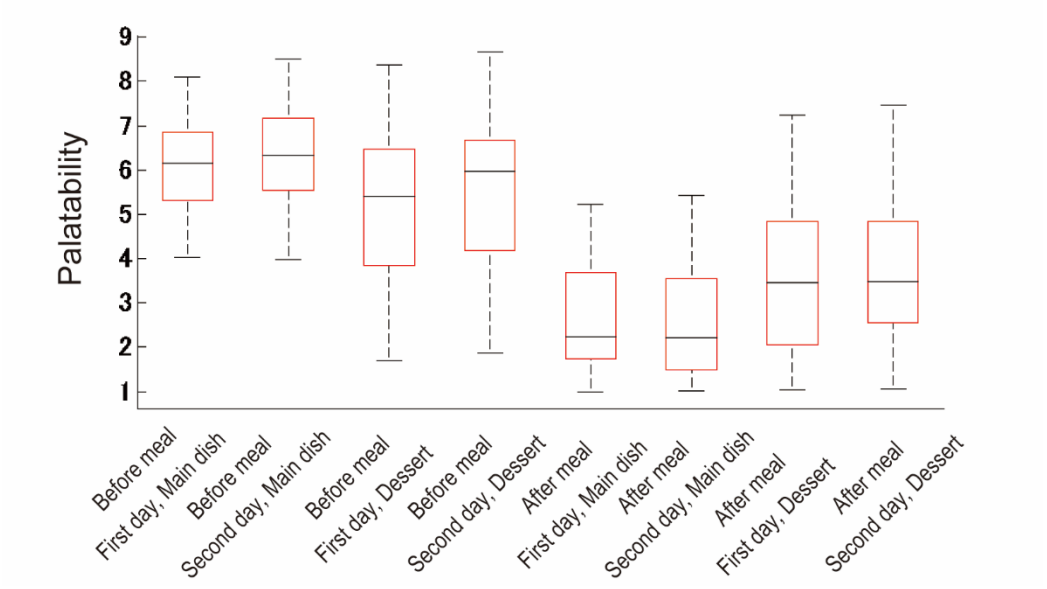
Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage, 15, 273–289. https://doi.org/10.1006/nimg.2001.0978.

9. Supplementary figures



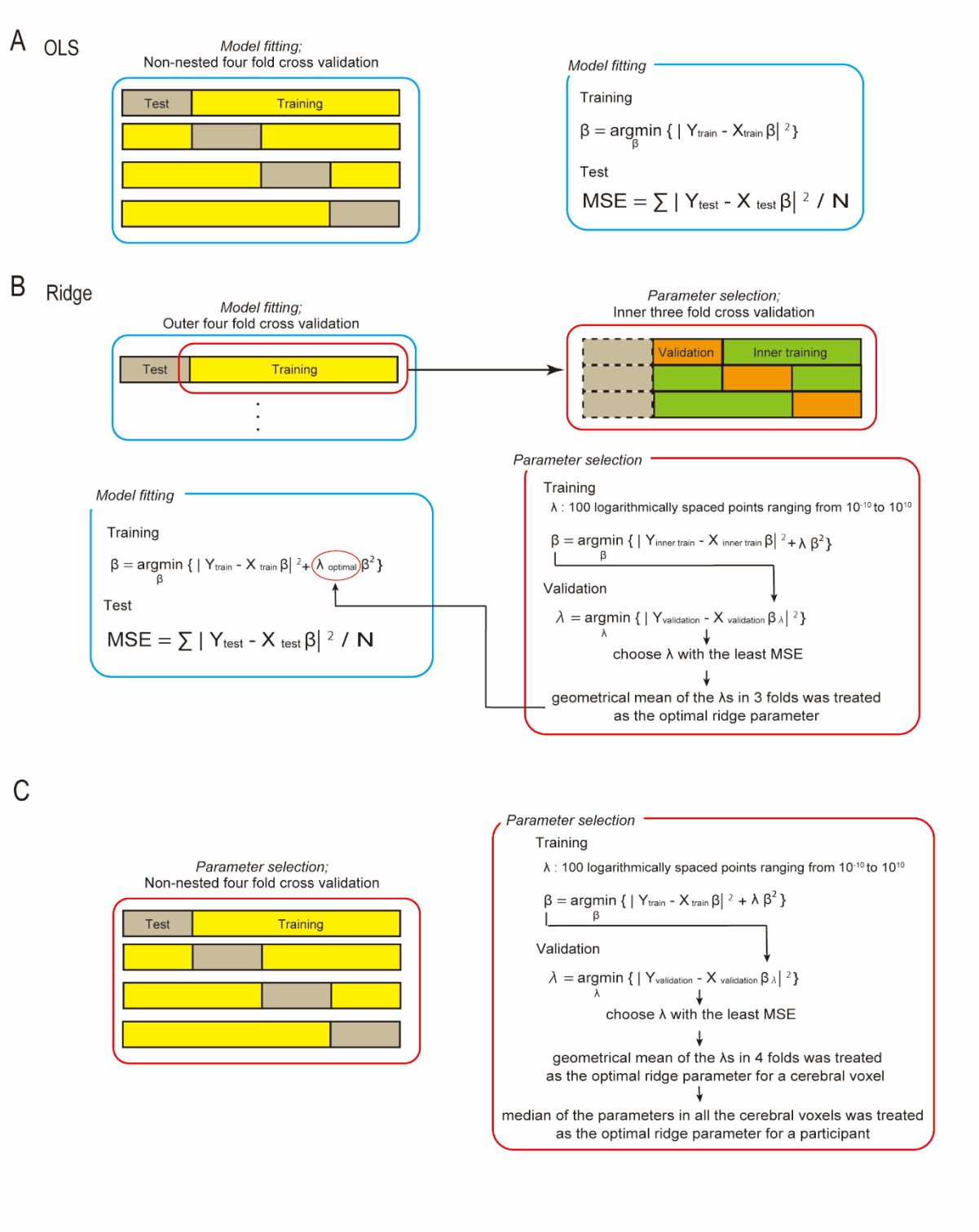
Supplementary Fig. 1. Rating scores of the palatability of food

The palatability of food was plotted based on meal (before and after meals) and experimental day (days 1 and 2). N=24. Boxes represent the median and 25^{th}/75^{th} percentiles, and whiskers represent the minimum and maximum values.



Supplementary Fig. 2. Rating scores of the palatablitiy of two food types (main dish and dessert)

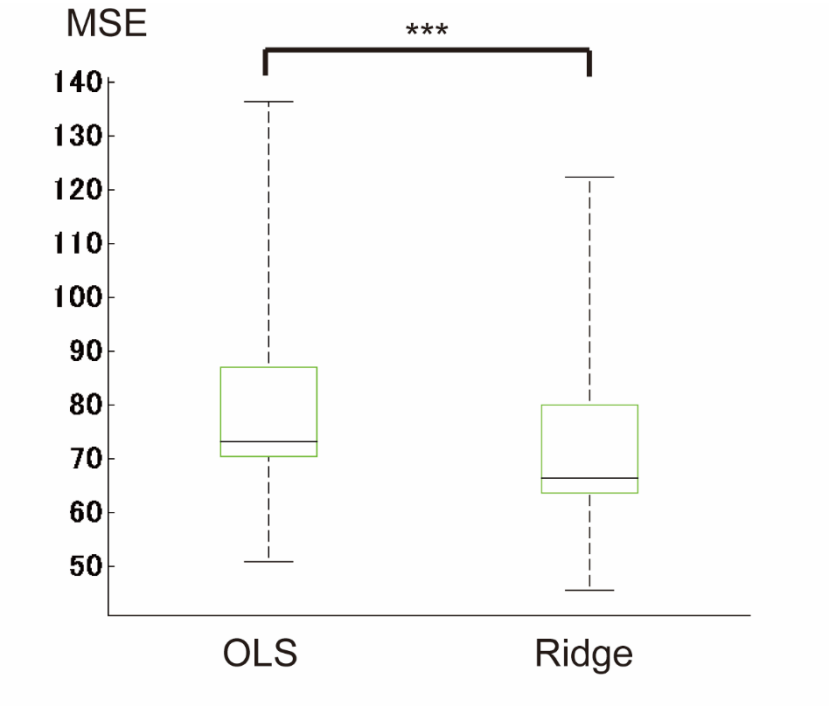
The palatability of food was plotted based on meal (before and after meals), experimental day (days 1 and 2), and food type (main dish and dessert). N=24. Boxes represent the median and 25^{th}/75^{th} percentiles, and whiskers represent the minimum and maximum values.



Supplementary Fig. 3. Comparison between the ordinary least square (OLS) and ridge regression methods and parameter selection in ridge regression

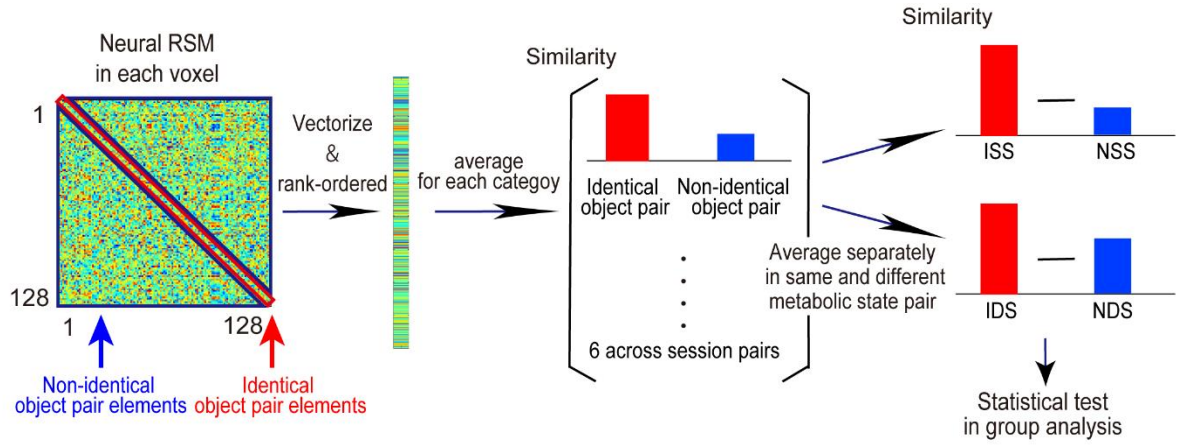
Confirmation analysis was conducted to confirm the efficiency of ridge regression in the estimation of the activity evoked by each trial. (A) Model fitting and prediction of voxel activity evoked by the respective stimuli. The mean squared error (MSE) in the test dataset was calculated using non-nested four-fold cross-validation (left). Note that the data in one fold correspond to one experimental session. Regression coefficients associated with the respective stimuli, which were calculated using the training dataset, were applied to predict the activity in the test dataset (right). The MSE in the test dataset was calculated to evaluate the model fitting. (B) In the ridge regression, the nested cross-validation technique was used. MSE was calculated in the outer four-fold cross-validation, as in the OLS regression (upper-left). Within the outer training dataset, parameter selection was conducted using the inner three-fold cross-validation (upper-right). Regression coefficients, which were calculated for each ridge parameter in the training dataset, were applied to estimate the MSE in the test dataset (lower right). The optimal parameter selected to minimize the MSE was chosen from each fold, and then, the geometrical mean of the three optimal parameters was used as the optimal ridge parameter after the inner cross-validation. Using this optimal parameter, the MSE was estimated from the test dataset (lower-left). (C) The procedure of parameter selection for ridge regression in the main analysis is illustrated. The full data consisting of four experimental sessions was used for the four-fold cross-validation. Regression coefficients calculated with each ridge parameter in the training dataset were applied to estimate the MSE in the validation dataset. The optimal parameter was

selected to minimize the MSE at each round, and then, the geometrical mean of the four optimal ridge parameters from the four rounds was treated as the optimal ridge parameter for one voxel. Finally, the median of the optimal ridge parameters from all voxels was adopted as the optimal ridge parameter for each participant.



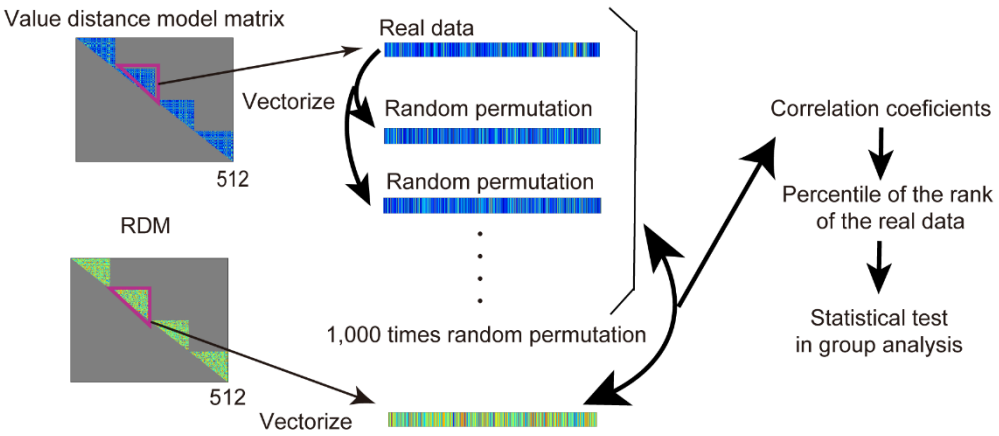
Supplementary Fig. 4. Comparison of the mean squared error (MSE) between ordinary least square (OLS) and Ridge regressions

The MSE of the OLS regression was significantly higher than that of the ridge regression. N = 24. Boxes represent the median and 25^{th}/75^{th} percentiles, and whiskers represent the minimum and maximum values. ***$p < 0.001$.



Supplementary Fig. 5 Schematic illustrations in the analytical procedures of the identical object representation

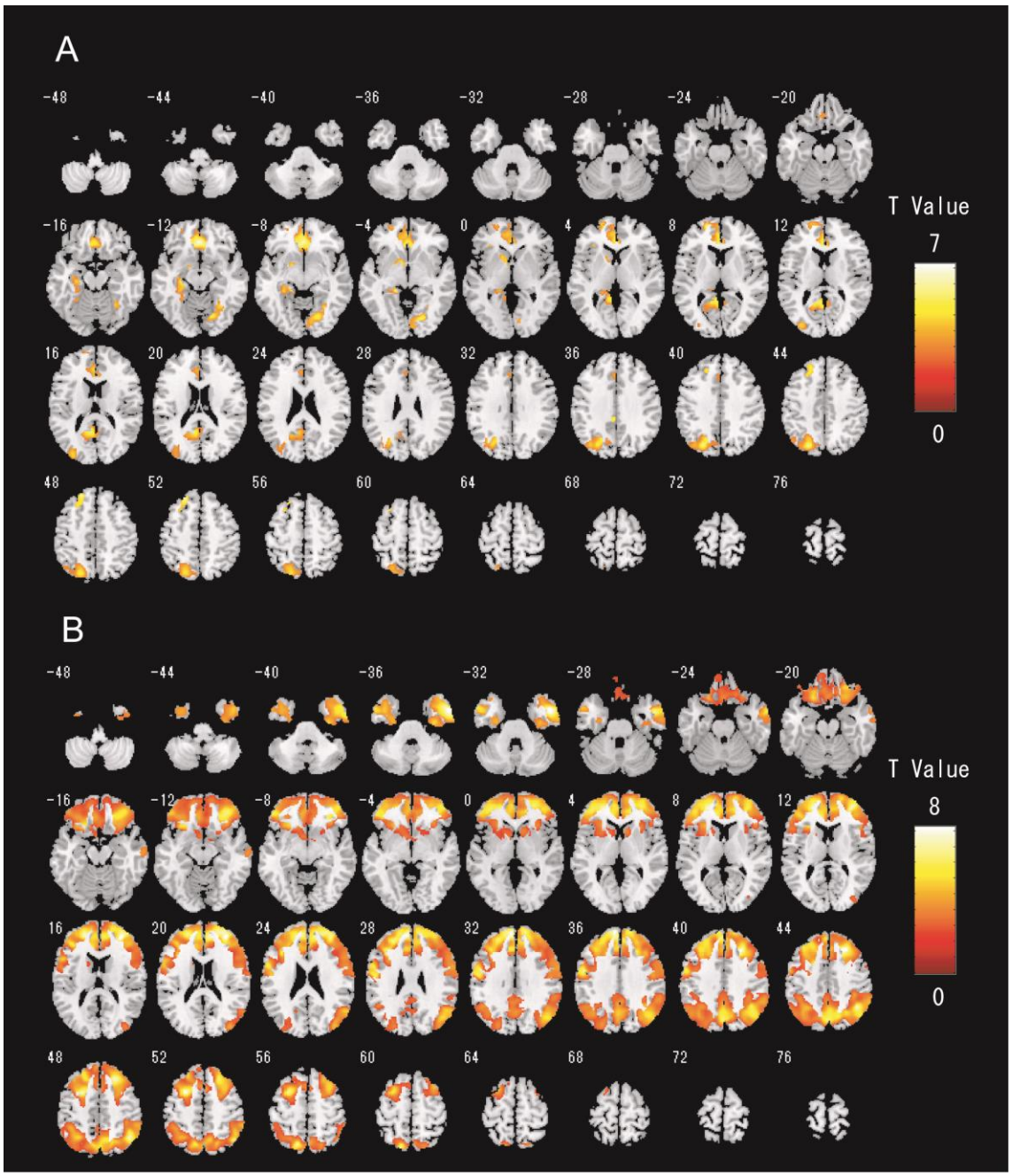
The neural representational similarity matrix (RSM) was converted to a vector and then rank-ordered. I calculated the average of the percentiles of the rank-ordered correlations between identical object pairs (red; diagonal elements in RSM) and average of the percentiles of the rank-ordered correlations between non-identical object pairs (blue; off-diagonal elements in RSM) in each cross-session pair (Fig. 1C, red and orange). To investigate the effect of the metabolic state change, I averaged these metrics separately for the same metabolic state experimental session pairs (Fig. 1C, red) and for different metabolic state pairs (Fig. 1C: orange). These representational similarity metrics were subjected to a one-sample t-test for group analysis. ISS, identical object pair in the same metabolic states; NSS, non-identical object pair in the same metabolic states; IDS, identical object pair in the different metabolic states; NDS, non-identical object pair in the different metabolic states.



Supplementary Fig. 6 Schematic illustrations in the analytical procedures of the multivariate analysis for the subjective value representations

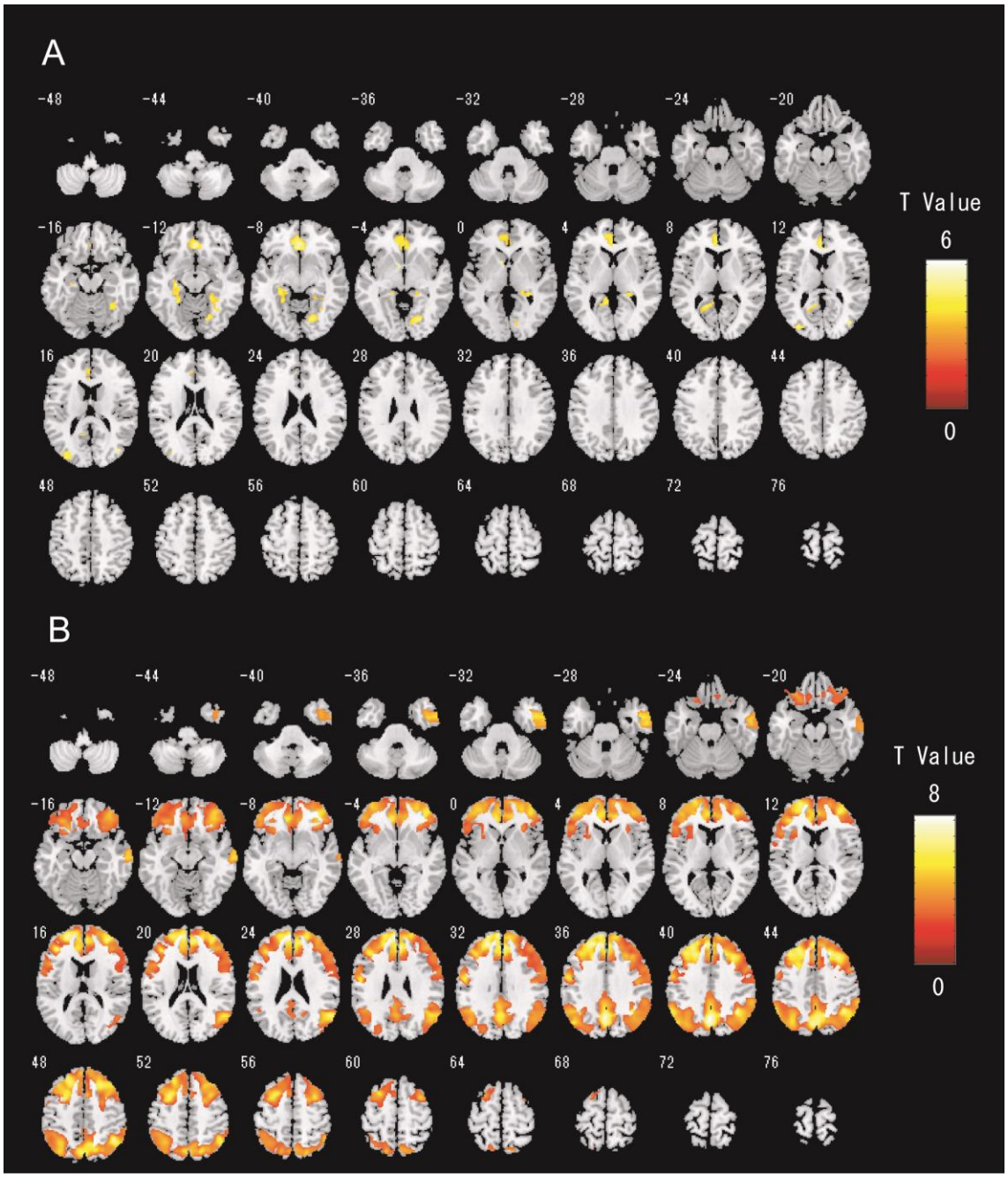
To maintain data independence, I used data from within-experimental session pairs (Fig. 1C, blue) only, in this analysis. I constructed a value distance model matrix (hereafter referred to as representational dissimilarity matrix, RDM) whose elements were used to code the absolute value of the subjective value difference in each food object pair. Neural RDM was also constructed by searchlight analysis (Fig. 1D). The RSM was converted to RDM by calculating $(1 - r)$ for each element. Because of the symmetry, only the data in the upper triangle were used. Both the value distance model RDM and neural RDM were converted into vectors. As in the identical object, white matter noise signals were regressed out from the neural RDMs. Then, the random permutation of the vector from the value distance model RDM was repeated 1,000 times. Correlation coefficients were calculated between the vector from the neural RDM and each permuted vector from the model RDM 1,000 times. The percentiles of the ranks of the

real data in a total of 1,000 correlation coefficients were averaged across the four experimental sessions. This RDM was submitted to a one-sample t-test for group analysis.



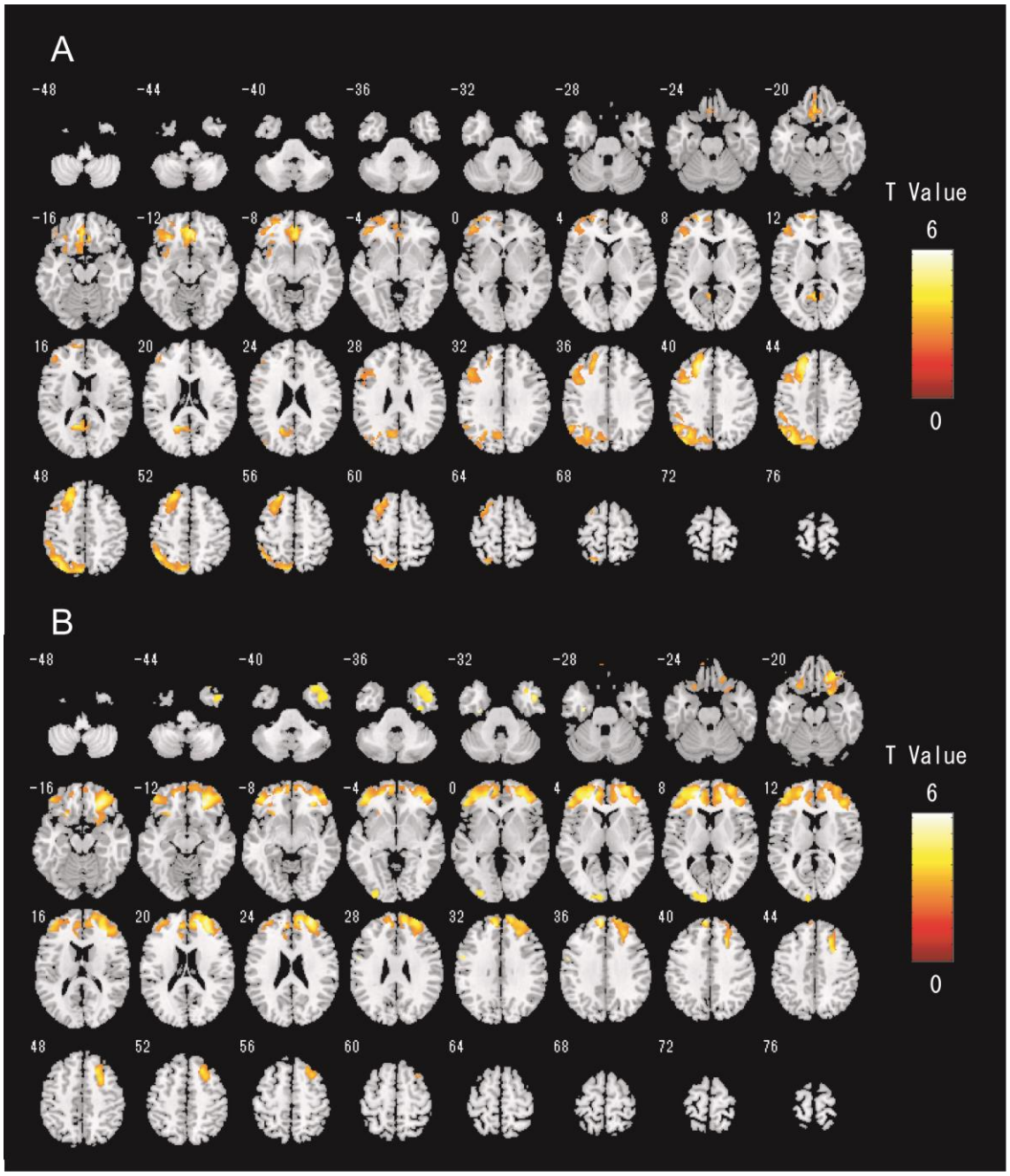
Supplementary Fig. 7. Value map in whole brain

Value map revealed at univariate (A) and multivariate analyses (B). The color represents the t-value threshold at $p < 0.05$, corrected (family wise error [FWE]). N=24.



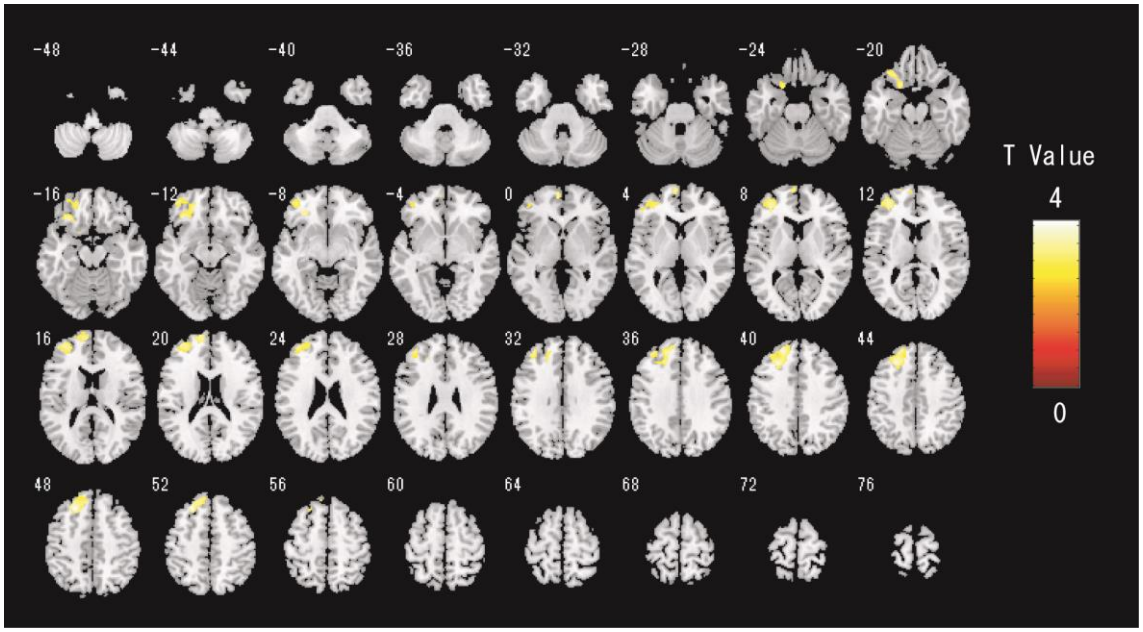
Supplementary Fig. 8. Value map in the hunger state in the whole brain

Value map in the hunger state revealed at univariate (A) and multivariate analyses (B). The color represents the t-value threshold at $p < 0.05$, corrected (family wise error [FWE]). N=24.



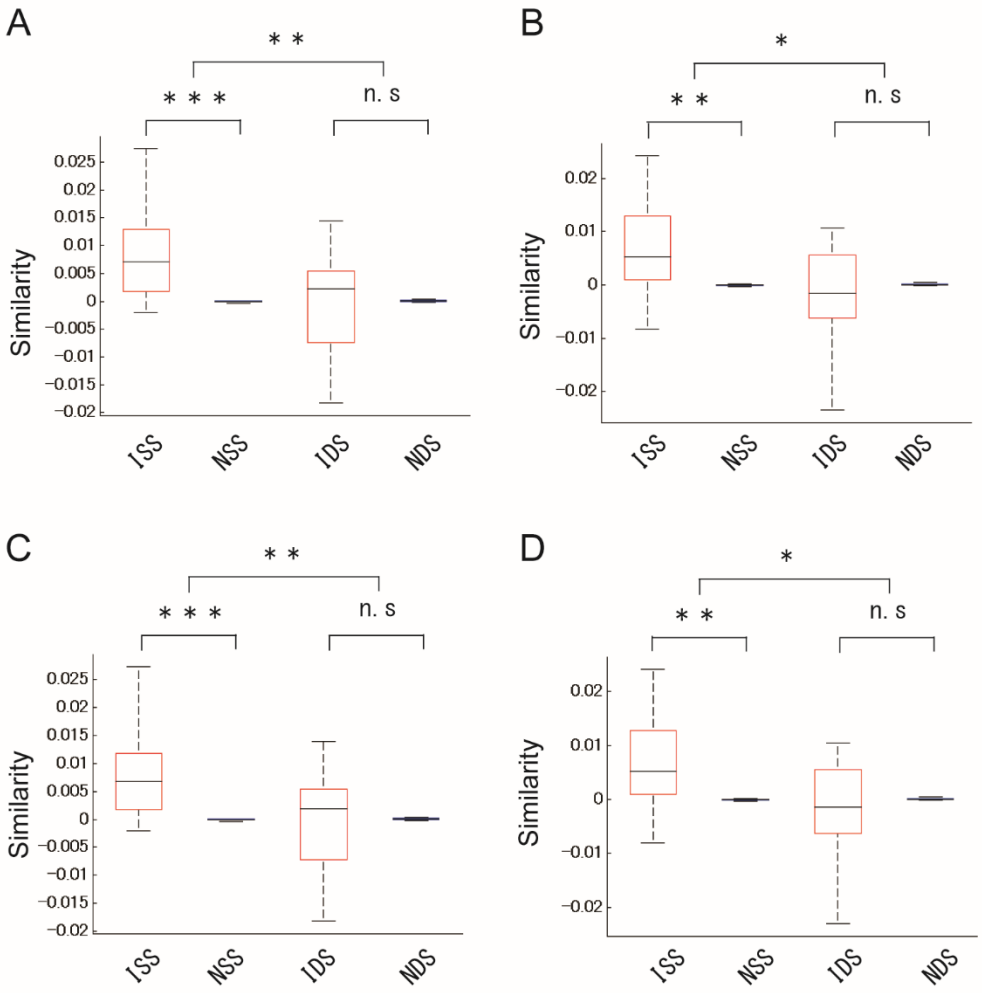
Supplementary Fig. 9. Value map in the satiated state in the whole brain

Value map in the hunger state revealed at univariate (A) and multivariate analyses (B). The color represents the t-value threshold at $p < 0.05$, corrected (family wise error [FWE]). N=24.



Supplementary Fig. 10. Whole-brain maps revealed at multivariate analysis of the difference in the identical object specificity effects between same and different state pairs while controlling for the effect of differences in value distance

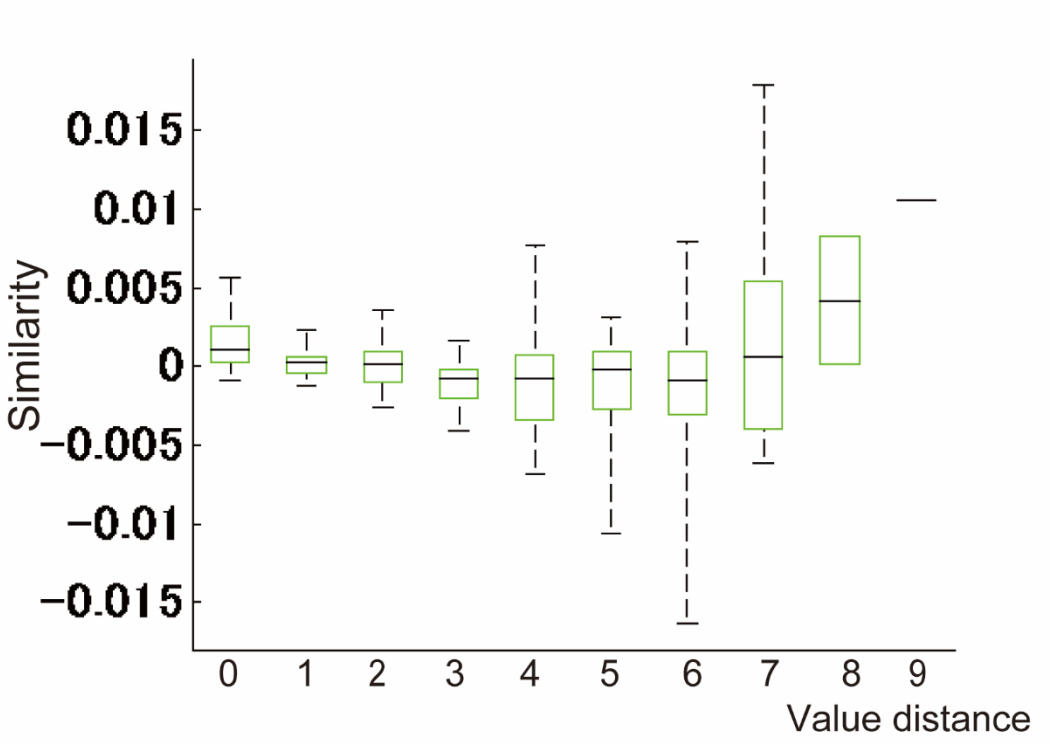
Multivariate analysis revealed that the left lateral orbitofrontal cortex (OFC) was associated with differences in identical object specificity effects across metabolic states. Colors represent t values in regions at $p < 0.05$ threshold, corrected (family wise error [FWE]). N=24.



Supplementary Fig. 11 Representational similarities in the lateral OFC for the main dish and dessert separately

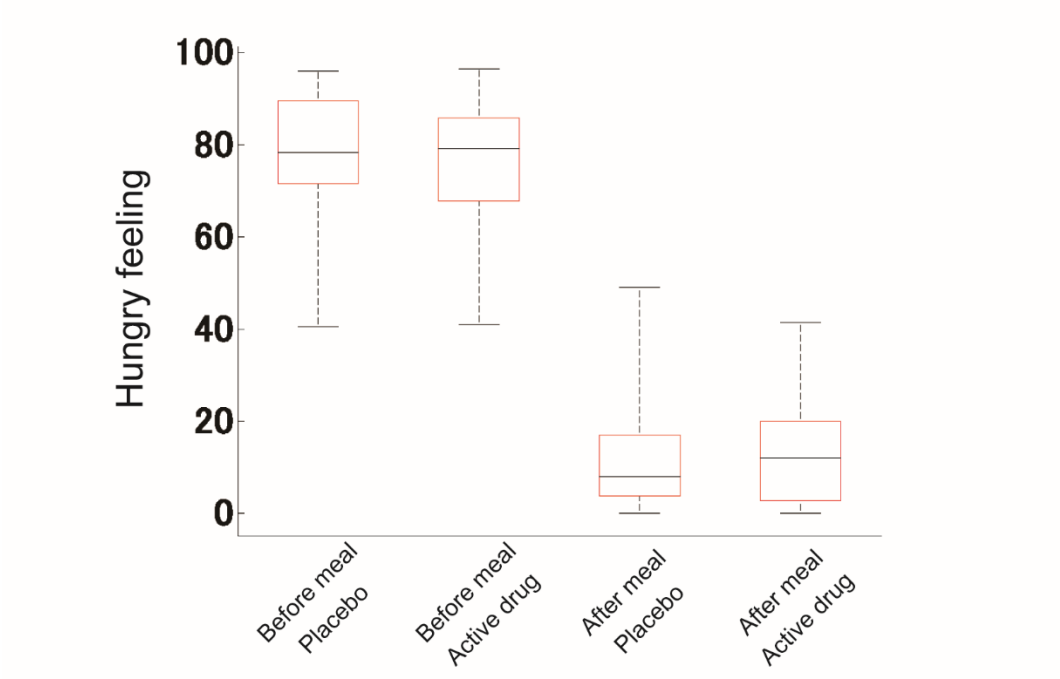
Representational similarities in the lateral orbitofrontal cortex (OFC) are shown separately by food type (main dish or dessert) and the analytic technique (with or without value distribution adjustment). (A) Main dish without adjustment for value distribution adjustment. (B) Dessert without value distribution adjustment. (C) Main dish with value distribution adjustment. (D)

Dessert with value distribution adjustment. ISS, identical object pair in the same metabolic state; NSS, non-identical object pair in the same metabolic state; IDS, identical object pair in different metabolic states; NDS, non-identical object pair in different metabolic states. n.s. not significant. Boxes represent the median and 25th/75th percentiles, and whiskers represent the minimum and maximum values. N=24. ***$p < 0.001$. **$p < 0.01$. *$p < 0.05$. Bonferroni corrected.



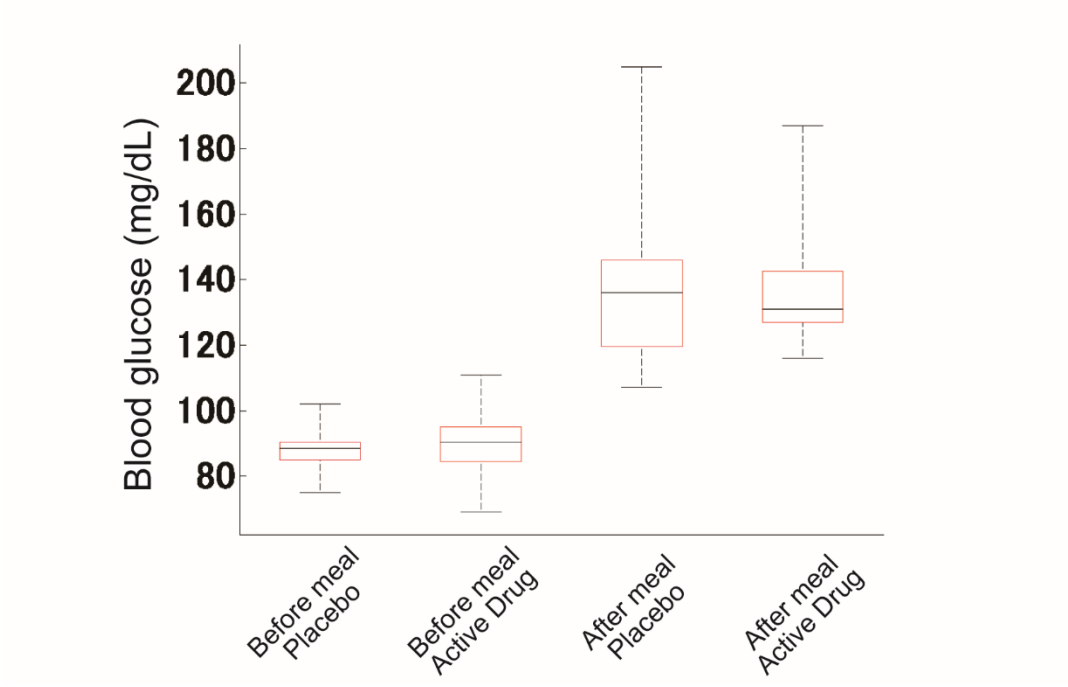
Supplementary Fig. 12 Neural representational similarity in all value distance bins in the lateral OFC

The neural representational similarities shown in Figure 4D were plotted for all value distance bins (from zero to nine). From left to right, the number of participants included in each bin was 24, 24, 23, 21, 16, 12, 9, 6, 2, and 1. Boxes represent the median and 25^{th}/75^{th} percentiles, and whiskers represent the minimum and maximum values.



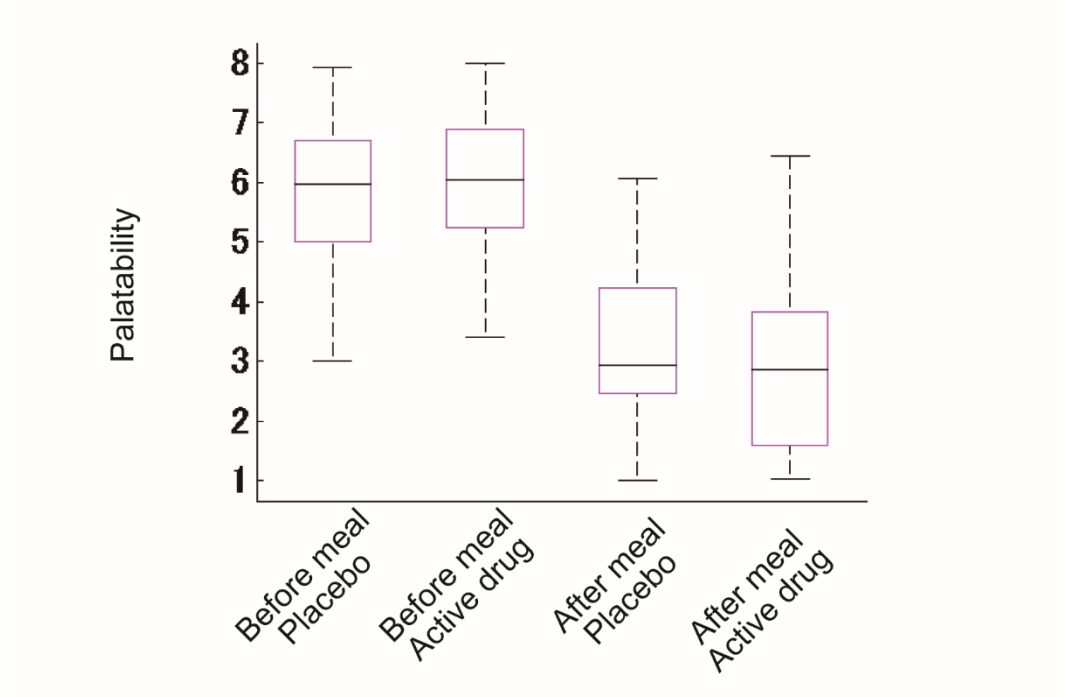
Supplementary Fig. 13 Rating scores of the hunger feeling by visual analogue scale (VAS)

The hunger feeling was plotted based on meal (before and after meals) and drug (placebo and active drug). N=24. Boxes represent the median and 25^{th}/75^{th} percentiles, and whiskers represent the minimum and maximum values.



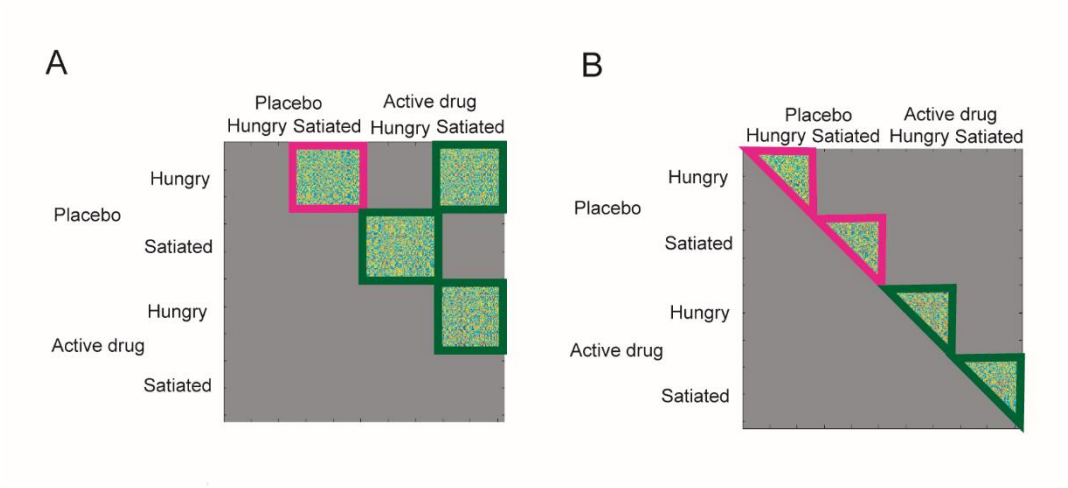
Supplementary Fig. 14 Blood glucose levels

Blood glucose levels were plotted based on meal (before and after meals) and drug (placebo and active drug). N=24. Boxes represent the median and 25^{th}/75^{th} percentiles, and whiskers represent the minimum and maximum values.



Supplementary Fig. 15. Rating scores of the palatability of food and nasal oxytocin administration

The palatability of the visual food stimuli was plotted based on meal (before and after meals) and drug (placebo and active drug). N=24. Boxes represent the median and 25^{th}/75^{th} percentiles, and whiskers represent the minimum and maximum values.



Supplementary Fig. 16. Effect of the nasal oxytocin administration on the representations of identical object and subjective value

(A) Schematic illustration of the session pair structure that was used to estimate the effect of nasal oxytocin administration on representations of identical objects. The experimental sessions were sorted based on drug (placebo or active drug). These four pairs were labeld as placebo–placebo (‘drug-free data,’ shown in purple), drug–drug, or drug–placebo (‘drug-treated data,’ shown in green). (B) Schematic illustration of the session pair structure that was used to estimate the effect of nasal oxytocin administration on value representation. The experimental sessions were sorted based on drug (placebo or active drug). Four within-session pairs were labeled as placebo–placebo (shown in purple) and drug–drug (shown in green).

10. Supplementary tables

Supplementary Table 1. Description of the visual food stimuli

Run 1		
No.	Main dish	Dessert
1	radish sprouts salad	plane scones
2	taco rice	coffee cream cake
3	Dried squid, fish roe and seaweed marinated with soy sauce	custard pudding with caramel
4	cut open and dried Atka mackerel	butter cookie
5	miso soup	waffle
6	hashed beef rice	fruit tart
7	sandwich with ham and vegitables	mont blanc
8	fried potato	fruit parfait
9	curry and naan	donut
10	a side dish of boiled seasoned cucumber and chinese cabbage	choco scones
11	salada with cabbage, tomatoes, and radish	choco cookie
12	pork curry and rice	Japanese bean-jam-filled wafers
13	meatballs with parsley	vanilla ice cream with blueberry sauce
14	boiled pumpkin	pound cake
15	salada with tomatoes, cucumber and cabbage	vanilla ice cream
16	cutlet sandwich	mile crape
Run 2		
No.	Main dish	Dessert
1	chirashi shushi, rice topped with sashimi and other ingredients	chocolate coated wheat puffs
2	udon, thick Japanese noodle with bonito flakes	"takenokonosato", a Japanese brand of chocolate cookies shaped like bamboo sprouts
3	chilled Chinese noodle topped with boiled egg, ham	sesame flavor cookie
4	fried chikin	almond ice cream
5	caesar salad	"M&M's", American brand cholate snacks
6	Chinese noodle with grilled pork, spring onion, bean sprout, wakame seaweed	crepe cookies
7	coleslaw salad	choco pie
8	curry rice with deep-fried oysters	blueberry tart cake
9	stewed hamburger steaks with demi-glace saurce	sponge cake
10	spaghetti with tomato sauce	bean-paste rice-dumpling
11	grilled tofu with slices of the garlic	cream puff
12	spaghetti with green onion and meat	honey churros
13	pot sticker	almond jelly
14	steamed pork dumplings	unbaked cheesecake
15	bread roll with cabbage, cucumber and ham	Aerated cholate snack with cone
16	rolled omlet	"mamedaifuku", a sweet rice cake stuffed with a sweet bean paste
Run 3		
No.	Main dish	Dessert
1	Udon noodles in a hot soup with shrimp tempura	cocoa biscuits cookie
2	sliced tangle boiled in sweetened soy sauce	thick "pocky", a Japanese brand of cholate-covered bisucuit sticks
3	chinese noodle in miso soup with cone, ham, sesame, seaweed and bean sprouts	eclair
4	pizza with salami, parma, cone, broccoli	muffin
5	beefsteak	gummy candy, cola taste
6	deep fried tofu with bean paste	milk chocolate topped with strawberry flavored white chocolate
7	hijiki seaweed salad	choco tip cookie
8	fried chikien and cabbage	caramel candy
9	omelet rice with starchy sauce	bun filled with sweet bean paste
10	omelet, cabbage and parsley	cup cone ice cream with strawbbery and bread
11	pepper stuffed with meat	cream puff
12	mapo tofu	honey dip donut
13	sandwich with bacon and lettuce	baumkuchen
14	baked sweet green pepper	baked cheesecake
15	a piece of the pizza with potato, cone, tuna, moyonnaise saurce	thin "Pocky", a Japanese brand of cholate-covered bisucuit sticks
16	baked salmon	butter crunch donut
Run 4		
No.	Main dish	Dessert
1	grilled squid	chocolate icecream
2	hard boiled egg	"kinokonoyama", a Japanese brand of chocolate cookies shaped like mushrooms
3	"katsudon", a bowl of rice topped with slices of deep-fried pork	chocolate cake
4	Szechuan sesame spicy hot noodles	crunky chocolate
5	"kimuchi", Korean side dish of salted and fermented vegetables	tiramisu
6	thick pizza with minced meat, tomato,arugura	chocolate chunk cookie
7	raw egg on rice	sweet bean paste in thin dough wrapping
8	salt-grilled saury	crepe with ice cream, cream, blueberry and strawberry
9	thin crust pizza with tomato,arugura	baumkuchen with cream
10	Ebi katsu (Japanese shrimp cutlet) burger	strawberry shortcake
11	omlet topped with demi-glace sauce	mille-feuille
12	pot-au-feu	almond cream puff
13	deep-fried horse mackerel	tart with coffee cream
14	steamed white rice	French cruller
15	spaghetti with scallop and dried mullet roe	chocolate bar
16	bacon egg and cheese English muffin	cookie with strawberry jam

Supplementary Table 2. Clusters in the value map revealed at univariate analysis

Cluster No	Cluster size (voxel)	MNI coordinate (mm)			t Value(24)	Location	
		x	y	z		Side	Area
1	1004	-2	36	-10	7.12	Medial orbitofrontal cortex	L
		-2	38	16	4.24	Antegior cingulate	L
		-4	58	8	3.69	Superior medial frontal gyrus	L
2	780	-22	-72	44	5.88	Superior parietal lobule	L
		-36	-64	32	4.10	Middle occipital gyrus	L
3	406	-8	-54	16	5.47	Precuneus	L
4	350	22	-76	-4	5.37	Lingual gyrus	R
		30	-56	-14	4.45	Fusiform gyrus	R
		6	-88	-6	3.80	Lingual gyrus	R
5	278	-20	36	52	6.13	Superior frontal gyrus	L
6	156	-16	-34	-6	5.72	ParaHippocampus	L
		-32	-16	-14	3.74	Hippocampus	L
7	109	-12	10	-6	3.44	Putamen	L
8	97	-32	-88	14	4.59	Middle occipital gyrus	L
9	68	-18	58	8	6.06	Superior frontal gyrus	L
10	25	-4	-34	36	4.84	Middle cingulate	L
11	11	-30	30	4	3.98	Insula	L

Listed are the cluster size, coordinates in MNI peak, maximum t-value at local peak, and anatomical label from the automated anatomical labeling (AAL) (Tzourio-Mazoyer et al., 2002).

Supplementary Table 3. Clusters in the value map revealed at univariate analysis in hunger state

Cluster No	Cluster size (voxel)	MNI coordinate (mm)			t Value(24)	Location	
		x	y	z		Side	Area
1	736	0	36	-10	5.94	Medial orbitofrontal cortex	R
		-4	40	10	4.64	Anterior cingulate	L
2	242	30	-36	0	4.60	Hippocampus	R
		32	-56	-14	4.47	Fusiform gyrus	R
3	211	-26	-22	-12	5.04	Hippocampus	L
		-24	-44	-8	4.68	Lingual gyrus	L
4	206	18	-72	-8	5.15	Lingual gyrus	R
5	146	-16	-52	8	4.82	Calcarine gyrus	L
6	135	-34	-82	14	5.19	Middle occipital gyrus	L
7	31	40	-78	14	5.53	Middle occipital gyrus	R
8	19	-12	10	0	5.68	Caudate	L

Listed are the cluster size, coordinates in MNI peak, maximum t-value at local peak, and anatomical label from the automated anatomical labeling (AAL) (Tzourio-Mazoyer et al., 2002).

Supplementary Table 4. Clusters in value map revealed at univariate analysis in the satiated state

Cluster No	Cluster size (voxel)	MNI coordinate (mm)			t Value(24)	Location	
		x	y	z		Side	Area
1	486	-34	-72	46	5.27	Parietal lobule	L
		-8	-74	56	4.32	Precuneus	L
		-52	-60	40	3.95	Angular gyrus	L
2	456	-22	32	42	5.76	Middle frontal gyrus	L
		-28	12	46	3.66	Middle frontal gyrus	L
3	273	4	40	-10	5.46	Medial orbitofrontal cortex	R
4	85	-12	-56	18	4.62	Precuneus	L
5	18	-30	32	-10	4.32	Inferior orbitofrontal cortex	L
6	16	-20	58	8	4.98	Superior frontal gyrus	L
7	15	-40	12	40	3.94	Middle frontal gyrus	L
8	13	-28	12	56	3.71	Middle frontal gyrus	L
9	8	-8	22	-18	4.52	Rectus	L

Listed are the cluster size, coordinates in MNI peak, maximum t-value at local peak, and anatomical label from the automated anatomical labeling (AAL) (Tzourio-Mazoyer et al., 2002).

Supplementary Table 5. Clusters in the value map revealed at multivariate analysis

Cluster No	Cluster size (voxel)	MNI coordinate (mm)			t Value(24)	Location	
		x	y	z		Side	Area
1	14273	32	48	18	9.69	Middle frontal gyrus	R
		24	28	42	8.47	Middle frontal gyrus	R
		-8	50	28	8.23	Superior medial frontal gyrus	L
		-26	8	50	8.08	Middle frontal gyrus	L
		-22	50	10	7.79	Superior frontal gyrus	L
		12	48	20	7.69	Anterior cingulate	R
		-28	42	30	6.90	Middle frontal gyrus	L
		34	44	-10	6.78	Middle orbitofrontal cortex	R
		-48	40	-6	6.08	Inferior orbitofrontal cortex	L
		-22	32	48	5.53	Superior frontal gyrus	L
		-8	30	32	5.48	Anterior cingulate	L
		-30	22	-12	5.33	Insula	L
		-30	22	12	5.16	Insula	L
		32	16	56	5.04	Middle frontal gyrus	R
		20	26	-18	4.50	Inferior orbitofrontal cortex	R
		8	36	0	4.43	Anterior cingulate	R
		-56	30	10	3.74	inferior frontal gyrus triangular part	L
2	3455	44	-56	46	7.71	Inferior parietal lobule	R
		-16	-68	58	6.99	Supeior parietal lobule	L
		-2	-72	42	6.55	Precuneus	L
		46	-68	30	5.52	Middle occipital gyrus	R
		56	-40	50	4.38	Inferior parietal lobule	R
		2	-48	38	3.72	Precuneus	R
3	1477	50	6	-32	8.99	Middle temporal pole	R
		32	-12	-36	4.97	Fusiform gyrus	R
4	732	-52	-2	36	6.55	Precentral gyrus	L
		-54	16	18	4.80	inferior frontal gyrus opercular part	L
5	697	-52	6	-32	5.53	Middle temporal gyrus	L
		-30	-12	-32	4.37	Fusiform gyrus	L
6	404	-30	-56	48	5.01	Inferior parietal lobule	L
		-46	-64	38	3.69	Angular gyrus	L
7	160	60	4	30	4.07	Precentral gyrus	R
8	34	62	-18	-14	3.91	Middle temporal gyrus	R
9	33	48	-6	36	4.03	Precentral gyrus	R
10	22	38	-36	46	3.99	Supramarginal gyrus	R
11	21	-46	-52	30	3.74	Angular gyrus	L
12	14	48	22	12	4.00	inferior frontal gyrus opercular part	R
13	11	-26	-80	34	3.63	Middle occipital gyrus	L
14	11	-2	16	52	3.73	Supplementary motor area	L
15	10	4	12	52	4.05	Supplementary motor area	R
16	5	62	-8	32	3.58	Postcentral gyrus	R

Listed are the cluster size, coordinates in MNI peak, maximum t-value at local peak, and anatomical label from the automated anatomical labeling (AAL) (Tzourio-Mazoyer et al., 2002).

Supplementary Table 6. Clusters in the value map revealed at multivariate analysis in the hunger state

Cluster No	Cluster size (voxel)	MNI coordinate (mm)			t Value(24)	Location	
		x	y	z		Side	Area
1	12898	-6	54	24	8.28	Superior medial frontal gyrus	L
		-22	44	34	8.07	Middle frontal gyrus	L
		20	30	44	7.90	Superior frontal gyrus	R
		6	38	30	7.24	Anterior cingulate	R
		-38	36	16	7.04	inferior frontal gyrus triangular part	L
		34	50	16	7.01	Middle frontal gyrus	R
		-28	10	52	6.94	Middle frontal gyrus	L
		-34	30	42	6.70	Middle frontal gyrus	L
		-8	40	-6	6.15	Medial orbitofrontal cortex	L
		42	48	-4	6.01	Inferior orbitofrontal cortex	R
		14	42	4	5.70	Anterior cingulate	R
		32	8	56	5.54	Middle frontal gyrus	R
		-34	54	0	5.17	Middle orbitofrontal cortex	L
		12	62	2	5.06	Superior medial frontal gyrus	R
		-10	58	4	4.66	Superior medial frontal gyrus	L
		-48	40	-6	4.65	Inferior orbitofrontal cortex	L
		6	12	48	4.56	Supplementary motor area	R
		-8	28	46	4.22	Superior medial frontal gyrus	L
		42	38	32	3.86	Middle frontal gyrus	R
2	4542	-2	-64	38	8.29	Precuneus	L
		38	-54	40	8.10	Angular gyrus	R
		54	-60	24	6.84	Angular gyrus	R
		16	-66	48	5.80	Precuneus	R
		58	-40	42	4.84	Supramarginal gyrus	R
		-18	-70	58	4.42	Superior parietal gyrus	L
3	1588	-56	-56	42	5.96	Inferior parietal gyrus	L
		-34	-68	46	5.48	Inferior parietal gyrus	L
4	1149	50	2	-28	5.90	Middle temporal gyrus	R
		64	-18	-14	5.43	Middle temporal gyrus	R
5	464	-54	-8	30	6.02	Postcentral gyrus	L
		-48	16	16	4.27	inferior frontal gyrus opercular part	L
6	227	-26	22	-12	4.51	Insula gyrus	L
7	89	52	8	44	4.27	Precentral gyrus	R
8	48	46	-12	34	4.55	Postcentral gyrus	R
9	37	-50	26	6	3.92	inferior frontal gyrus tricular part	L

Listed are the cluster size, coordinates in MNI peak, maximum t-value at local peak, and anatomical label from the automated anatomical labeling (AAL) (Tzourio-Mazoyer et al., 2002).

Supplementary Table 7. Clusters in value map revealed at multivariate analysis in the satiated state

Cluster No	Cluster size (voxel)	MNI coordinate (mm)			t Value(24)	Location	
		x	y	z		Side	Area
1	1664	24	38	-14	5.87	Superior orbitofrontal cortex	R
		26	54	22	5.75	Middle frontal gyrus	R
		48	50	-6	4.27	Middle orbitofrontal cortex	R
		40	38	12	4.09	inferior frontal gyrus triangular part	R
2	886	-22	54	8	5.71	Superior frontal gyrus	L
		-44	44	-4	4.64	Inferior orbitofrontal cortex	L
3	473	48	6	-32	5.23	Middle temporal pole	R
		30	18	-40	4.39	Middle temporal pole	R
4	213	20	24	42	4.84	Superior frontal gyrus	R
5	206	-10	-98	8	4.48	Superior occipital gyrus	L
6	155	-8	56	38	5.70	Superior medial fronal gyrus	L
7	101	-14	50	6	4.59	Superior medial fronal gyrus	L
8	82	-30	-92	-4	4.32	Middle occipital gyrus	L
9	37	34	-10	-36	4.20	Fusiform gyrus	R
10	36	14	48	8	4.04	Superior medial frontal gyrus	R
11	33	-54	0	34	4.93	Precentral gyrus	L
12	32	26	38	34	3.85	Middle frontal gyrus	R
13	18	-32	-14	-30	5.78	Fusiform gyrus	L
14	10	22	10	-20	4.11	Superior temporal pole	R
15	7	-26	22	-12	3.79	Insula	L

Listed are the cluster size, coordinates in MNI peak, maximum t-value at local peak, and anatomical label from the automated anatomical labeling (AAL) (Tzourio-Mazoyer et al., 2002).

Supplementary Table 8. Clusters in the whole-brain maps revealed at multivariate analysis of the difference in identical object specificity effects between same and different pairs states while controlling for the effect of differences in value distance

Cluster No	Cluster size (voxel)	MNI coordinate (mm)			t Value(24)	Location	
		x	y	z		Side	Area
1	121	-24	30	50	4.05	Middle frontal gyrus	L
2	37	-20	44	40	3.84	Supperior frontal gyrus	L
3	31	-34	44	8	3.65	Middle frontal gyrus	L
4	9	-34	30	-12	3.79	Inferior orbitofrontal cortex	L
5	7	-10	62	18	3.61	Superior medial frontal gyrus	L
6	2	-38	42	34	3.54	Middle frontal gyrus	L
7	1	-20	32	32	3.52	Supperior frontal gyrus	L

Listed are the cluster size, coordinates in MNI peak, maximum t-value at local peak, and anatomical label from the automated anatomical labeling (AAL) (Tzourio-Mazoyer et al., 2002).