PNAS
www.pnas.org

Supplementary Information for

Spatiotemporal dynamics of odor representations in the human brain revealed by EEG decoding

Mugihiko Kato, Toshiki Okumura, Yasuhiro Tsubo, Junya Honda, Masashi Sugiyama, Kazushige Touhara and Masako Okamoto

*Masako Okamoto and Kazushige Touhara
Email: masakookamoto3@gmail.com , ktouhara@g.ecc.u-tokyo.ac.jp

This PDF file includes:

SI Method

Subjects. Twenty-six subjects (age: 19–28 years; 13 females) participated in the study with monetary compensation for their participation. Among them, two subjects who did not respond to a task or sense odor in more than half of the trials (15 trials) for at least one odor and two subjects who did not complete the experiment due to physical conditions were excluded from the analysis. Consequently, 22 subjects remained (19–28 years; 11 females). The inclusion criteria for participation were being right-handed (Edinburgh Handedness Inventory (1), score higher than 60) and native Japanese speakers (self-report), not having either olfactory, respiratory, psychiatric, or neurological disorders, traumatic brain injury, or metals in the body, and not currently smoking, pregnant, or using medication. Subjects were asked to refrain from consuming any food or drink except for water within 2 h before the experiment. Written informed consent was obtained from the subjects before conducting the experiments. The study was approved by the ethics committee of the University of Tokyo and in accordance with the Declaration of Helsinki.

Odor selection. Ten odors were selected from 138 monomolecular odors listed on a database compiled by Dravnieks, in which the applicability of 146 semantic descriptors based on ratings of 120–140 panelists was reported for each odor (2). To select odors with a variety of pleasantness and quality, we conducted principal component analysis on the applicability and selected odors that were widely distributed in 2-dimensional principal component space. To minimize the influence of the trigeminal system, we excluded odors that induced apparent trigeminal sensations from candidates (see also subsection, “Trigeminal test”). Consequently, the following odors were selected: allyl caproate (Ally; 10%; TCI), fructone (Fru; 10%; TCI), citral (Cit; 1%; TCI), linalool (Lin; 1%; Santa), vanillin (Van; 8%; TCI), acetophenone (Ace; 1%; TCI), hexanal (Hex; 1%; TCI), alpha-pinene (Pin; 10%; Sigma), cyclodithalfarol (Cyc; 1%; TCI), and 4-pentanoic acid (4Pe; 10%; TCI). The odors were diluted in propylene glycol (kindly provided by T. Hasegawa Company, Tokyo, Japan). The dilution ratio was determined based on pilot studies using a separate group of subjects (n = 8), so that the subjective intensity of odors became approximately equal.

Odor delivery. Odors were presented in a computer-controlled setup using an olfactometer (OL022; Burghart Messtechnik GmbH), which was supplemented with a custom-made heating system and upgraded to allow a high flow rate. The olfactometer provided a constant flow of humidified and heated air via a nosepiece inserted in the right nostril and allowed switching of the base airflow to stimulus airflow without mechanical stimulation so that odor presentation was possible without tactile stimulation (Fig. S1A) (3). While OL022 originally had only four odor channels, we modified the system to allow presentation of 10 odors without a loss of temporal precision (Fig. S1B and C) (4). Odor onset time (time = 0) was defined as the time point at which the PID value was above 10% of the maximum value (60 ms after valve switching; Fig. S1C), and all analyses were conducted after correcting this delay. The flow rate was 7.5 L/min for both the baseline and stimulus airflows, which were humidified and heated to approximately 36 °C. To avoid time lags due to differences of odor accessibility between right and left nostrils, caused by factors such as differences of nasal cavity structures, air was presented monorhinally, as in previous studies on olfactory event-related potentials (OERPs) (5–10). Subjects performed the velopharyngeal closure breathing technique to minimize nasal airflow during the experiment (11). The olfactometer was located outside an electrically shielded EEG chamber. The subjects were seated in a relaxed position and listened to white noise to mask the valve switching click sounds from the olfactometer.

Experimental procedure. The experiment consisted of five sessions conducted over four days (Fig. 1A): an EEG session, an intensity rating session, a pairwise similarity rating session, a trigeminal test session, and a quality and unipolar (un)pleasantness rating session (details are as follows). Obtaining perceptual ratings on a separate day from the EEG session was necessary for reducing subjects’ fatigue, as both EEG recordings and ratings required a long time. The order of the sessions was as shown in Fig. 1A, and the same across subjects. The inter-subject mean ± SD of interval between Day 1 and Day 4 was 18.0 ± 7.0 days.

EEG recording

Before the EEG recordings, the subjects were provided explanations regarding the experimental task. Next, they were trained for velopharyngeal closure breathing, which was continued until airflow via the nostril detected by the breathing sensor became flat. They were also trained in the experimental task for 5 to 15 min, which was continued until they were able to respond to tasks within the time limit and tolerate without eye blinking during the fixation phase (see below). This training was conducted in almost the same way as during the EEG recording, but no odors were presented.

After training, each subject underwent 10 EEG recording sub-sessions. Each recording sub-session consisted of 30 trials, three for each odor, which were presented in a pseudorandom order. Thus, data from 30 trials for each of the odors were acquired across all sub-sessions. At the beginning of each recording sub-session, the subjects inserted a nosepiece to their right nostril, through which the olfactometer airflow was presented. Trials started when the base airflow stabilized to the specified flow rate. Each trial consisted of three phases: fixation, response, and blinking (Fig. 1B). During the fixation phase, subjects were instructed to focus on a white cross and avoid blinking. The olfactometer valve switched the airflow from base air to stimulus air for 1 s within this fixation phase. The stimulus onset time was jittered to minimize the odor-related expectancy. The fixation phase lasted for 3 s after stimulus offset, and subsequently, the response phase began. During the response phase, subjects were asked to rate odor pleasantness on a scale of 1 (very unpleasant) to 5 (very pleasant). They were instructed that if they did not sense any odor, they should report 0. To avoid motor-related confounding factors in the decoding analysis, ratings were provided as follows (12). First, a random number from 0 to 5 was displayed at the beginning of the response phase. The numbers were changed by pressing two keys on a keyboard (left arrow key labeled with the “up” mark for increment, and down arrow key labeled with “down” mark for decrement) using the index and middle fingers of their right hands. When a desired number was displayed, subjects finalized the rating by pressing another key (up arrow key labeled with a circle mark) using either the index or middle fingers of their right hands. The delayed-response paradigm was used to minimize motor-related confounds in the ERP (13). The response phase lasted for a maximum of 7 s or until response, and if the response phase was less than 7 s, the blinking phase followed, where a white circle was displayed. Subjects were allowed to blink during the response and blinking phases. The total duration of the response and blinking phases was 7 s. The inter-stimulus intervals (ISIs) were 16 ± 1 s. Each recording sub-session lasted ~9 min, with a 1 to 3 min break between them.

Intensity rating

In the intensity rating session, subjects underwent 10 trials, one for each odor, which were presented in random order. The odor delivery condition (e.g., flow rate), and the timeline of each trial were the same as those of the EEG session, except that subjects rated odor intensity on a scale of 1 (odorless) to 5 (extremely strong) instead of pleasantness, by changing a random number from 1 to 5, as described previously.

Pairwise similarity rating

In the similarity rating session, the subjects evaluated the similarity of two odors presented consecutively. Each subject rated all 45 possible odor pairs twice, once in the forward and once in the reverse order. The odor pairs were presented in pseudorandom order, with trials separated into six sub-sessions, each consisting of 15 trials, and with 90 trials in total. Odor delivery conditions were the same as those of the EEG session, except for the timeline of the trials. At the beginning of each trial, a visual cue, ‘Odor A,’ was displayed for 5 s. The first odor was presented during the last 1 s of the visual cue presentation. Then, an 11 s fixation period followed, and a cue, ‘Odor B’ was presented for 5 seconds, with the second odor being presented in the last 1 s. The subjects then rated the similarity of the two odors using a visual analog scale (0, dissimilar; 100, similar). The next trial started 11 s after the offset of odor B in the previous trial. Each sub-session lasted for approximately 8 min, with a 1 to 3 min break between them.

Trigeminal test

While we selected odors that did not induce apparent trigeminal sensations at the concentration used here, considering that individual differences in trigeminal sensitivities are known (14, 15), we tested the trigeminal sensations of each odor for each subject.

The trigeminal sensitivity was measured using the lateralization test (we refer to this test as trigeminal test in the rest of the manuscript) (16, 17). This test is based on experimental evidence that when odors are administered monorhinally, it is difficult for humans to identify the stimulated side of the nostril, unless the odor stimuli simultaneously excite the trigeminal system (18). The test was administered using an odor presentation condition as similar as possible to that used in the EEG session (using the olfactometer with a 7.5 L/min flow rate for each nostril; each airline was humidified and heated to 36 °C; odor duration, 1000 ms; ISI, 17 s; velopharyngeal closure breathing), with the major exception that the air was presented birhinally as described below. Each subject underwent 10 sub-sessions—one sub-session for one odor—comprising 20 trials. At the start of each trial, subjects inserted two nosepieces, one for the right and the other for the left nostril, each connected to a separate olfactometer channel calibrated to an identical flow rate. During the baseline period, odorless air was delivered to both nostrils. During the odor period, the odorless air was switched to odorous air for one nostril, while for the other nostril, it was switched to another odorless airline. Odorless to odorless switching was performed to control for any possible tactile stimulation evoked by switching. After odor presentation, subjects handed the nosepieces to the experimenter and stated the side in which they felt the odor in a forced-choice manner. The subjects were then handed two nosepieces again for the next trial. For each subject for each odor, a trigeminal test score was calculated as the percentage of correct trials. In the positive control experiment using an independent subject group, we found that the inter-subject mean trigeminal test score of the eucalyptol (2.5%; TCI), which is known to stimulate the trigeminal system (19), was 96 (20 trials, 5 subjects). The subjects were blindfolded during the trigeminal test. The order of the odors was randomized between the subjects.

Quality, unipolar pleasantness and unpleasantness ratings

Subjects rated the perceived quality of odors using 44 descriptors on a scale of 1 (not at all applicable) to 5 (very well applicable). These descriptors were selected from a translated version of 146 descriptors used in the odor profile database of Dravnieks (2) based on pilot studies using a separate subject group (n = 9). Moreover, to evaluate the odor unipolar pleasantness and unpleasantness, subjects also rated the applicability for two additional descriptors, ‘pleasant’ and ‘unpleasant’, in the same 5-step scale. Consequently, there were 46 descriptors in total. The odor delivery condition was the same as that used in the EEG session, except that subjects provided ratings for all 46 descriptors after every odor presentation, and each trial lasted until the ratings were completed (approximately 3 min). Each odor was presented three times, for a total of 30 trials. The order of odors was randomized across all trials, with the descriptors presented in random order in each trial.

EEG acquisition. The EEG signals were obtained using 64 active Ag-AgCl electrodes placed according to the international 10–20 system (BioSemi Active Two, BioSemi GmbH, Amsterdam, NL). A standard BioSemi cap was used for electrode placement. In addition, Biosemi active electrodes were placed at the following four locations: the right and left mastoids for later offline referencing, and below and above the left eye to obtain the electrooculogram (EOG). The signals were recorded at a sampling rate of 2048 Hz. Online low-pass filtering was performed in the analog-to-digital converter’s decimation filter (hardware bandwidth limit) using the default settings of BioSemi, which has a fifth order cascaded integrator-comb filter response with a -3 dB point at 1/5th of the sample rate (i.e., 410 Hz in the current study).

EEG preprocessing. EEG data were analyzed using EEGLAB (version 13.6.5b)–an open-source toolbox for EEG data analysis (20)– and custom-written MATLAB scripts unless otherwise stated. All offline filters were carried out using the default settings of the EEGLAB function pop_eegfiltnew, which implements linear, non-causal, zero-phase Hamming windowed sinc finite impulse response (FIR) filters (21). The continuous time series data were first processed by PREP (22) to remove line noise and detect bad electrodes. After line noise removal, re-referencing to a robust mean reference, and interpolation of bad electrodes, the data, which we

will refer to as “post-PREP,” were saved for later analyses. Independent component analysis (ICA) was performed using the runica function in the EEGLAB (23). To optimize ICA, the post-PREP data were bandpass filtered between 1 (passband edge, 1 Hz; cutoff frequency (-6 dB), 0.5 Hz; transition band width, 1 Hz; filter order, 6760) and 249 Hz (passband edge, 249 Hz; cutoff frequency (-6 dB), 280.125 Hz; transition band width, 62.25 Hz; filter order, 110), divided into 5400 ms epochs for each trial, beginning 2200 ms before the stimulus onset. Note that the resulting data were only used for ICA and not for later analyses. The rank of the data was adjusted for the number of interpolated electrodes. Independent components (ICs) were probabilistically classified using MARA (24), and ICs with artifact probability scores above 90% were considered artifacts. Next, the post-PREP data (not the preprocessed data for ICA) were high pass filtered at 0.2 Hz (passband edge, 0.2 Hz; cutoff frequency (-6 dB), 0.1 Hz; transition band width, 0.2 Hz; filter order, 33792) and divided into 5400 ms epochs. Then, ICs regarded as artifacts were removed, and bad electrodes detected by PREP were interpolated by the eeg_interp function in EEGLAB. The inter-subject mean and SD of the number of removed ICs was 32.1±6.0, and the number of interpolated electrodes was 8.9±5.8. Data were baseline corrected using a pre-stimulus interval of −500 to 0 ms and low-pass filtered at 30 Hz (passband edge, 30 Hz; cutoff frequency (-6 dB), 33.75 Hz; transition band width, 7.5 Hz; filter order, 902). After filtering, the data were resampled to 500 Hz to reduce computation time and re-referenced to a mean reference. Trials in which the absolute amplitude exceeded 100 μV in the vertical EOG (top minus bottom EOG) or 80 μV at any scalp electrodes were rejected to remove any remaining artifacts (7.4±19.4 trials were rejected). In addition, trials in which the subjects did not sense odors were also removed. Consequently, 283.0 ± 24.1 trials were used for the following analysis, and the minimum number of trials for one odor across subjects was 18.

Visualization of OERPs. To visually inspect odor-evoked potentials, we plotted the time courses of the grand mean global field powers (GFPs; Fig. 1G) and OERP waveforms (Fig. S2). GFP is a measure of electric field strength and is defined as the standard deviation of ERP amplitudes across electrodes (25). GFP was computed for each subject, each odor, and each time point and was baseline corrected by subtracting the mean amplitude during the pre-stimulus interval of −500 to 0 ms. An increase in GFP against the baseline was tested for each of the post-stimulus time points using one-sided, one-sample Student’s t-test ($P < 0.05$, Bonferroni–Holm corrected for number of time points). In addition, classical OERP waveforms were obtained by averaging the ERP amplitudes across trials for each condition. An average of the mastoid electrodes was used as a reference, as in previous OERP studies (26).

Source reconstruction. To examine the spatiotemporal dynamics of odor representations, we estimated the electrical activity in the brain using the standardized low-resolution brain electromagnetic tomography (sLORETA) software available at http://www.uzh.ch/keyinst/loreta.htm. Source reconstruction was performed using the trial-averaged ERPs for each subject and each odor, and the 3D coordinates of the EEG electrode positions were obtained from the BioSemi official website (https://www.biosemi.com/headcap.htm). The forward problem was addressed using a realistic head model (27) derived from the Montreal Neurological Institute (MNI) 152 template (28). The inverse problem was addressed using the sLORETA method (29), which has been shown to correctly localize EEG signals to deep structures, including those known to be associated with olfaction, such as the medial temporal regions (30–32). Brain activity was estimated as the standardized current density power for each of the 6239 isotropic (5 × 5 × 5 mm) voxels in the cortical gray matter in the MNI space for every time point after odor onset. The source orientations were unconstrained (29). For each voxel in the source space, the standardized current density was estimated for three Cartesian components, and the standardized current density power of each voxel was computed as the sum of their squires. The data were then down-sampled to 20 Hz, the same temporal resolution as in the sensor-level decoding analysis and used for source-level tRSA. The results of the source-level tRSA were visualized by overlaying on a brain using a sLORETA software. Because the anatomical atlas provided by sLORETA software does not define the POA, voxels containing the POA (piriform cortex, anterior

olfactory nucleus and olfactory tubercle) were determined based on their MNI coordinates and a previous study that reported the MNI coordinates of the POA (33).

Statistical hypothesis testing. All statistical tests were performed using MATLAB functions (Statistics and Machine Learning Toolbox, R2021a), with the number of observations equal to the number of subjects (n = 22), unless otherwise stated. We report the specific method for statistical testing, multiple testing correction, and alpha level used for each analysis in the legends of figures where corresponding results are presented. Whenever we used Student's t-test, we confirmed that the distributions did not significantly deviate from the normal distribution using the Shapiro-Wilk test (alpha level = 0.05) in advance.

Time-resolved multivariate pattern analysis (tMVPA). To assess the temporal dynamics of neural representations of odor information, we performed tMVPA decoding of individual odors from EEG time series data using pairwise decoding models (Fig. 2A) and multi-class decoding models (Fig. S3). For both pairwise and multi-class decoding, we employed two approaches, one based on a subject-wise model, and the other based on a cross-subject model. In the former approach, the decoding models were built separately for each subject, and then group-level inference was performed by combining the subject-wise decoding performances of all the subjects. This is the most commonly used approach in cognitive neuroscience (34, 35) and has the advantage of allowing flexibility to conduct individually tailored analysis. In contrast, in the cross-subject approach, the decoding models were generalized across subjects. Therefore, whether and when brain signatures of odors common across subjects exist can be addressed using this approach.

For all the analyses, EEG data were first temporally down-sampled to 100 Hz. Then, decoding models were built every 50 ms. For each model, single-trial ERP amplitudes at all 64 scalp electrodes within a preceding time interval of 200 ms were used as features. This yielded 1280 features per model (64 electrodes × 20 time points).

Pairwise decoding models

For both the subject-wise and cross-subject approaches, pairwise decoding models were built for each of the 45 possible odor pairs using an ℓ^2-regularized linear least-squares classifier. We used a nested cross-validation (CV) procedure, where an inner loop CV responsible for tuning the cost parameter was nested in an outer loop CV, which was responsible for model evaluation. The cost parameter, λ, was chosen from 11 possible values (log-spaced, between 2^{-15} and 2^{15}) based on 10-fold CV in the inner loop of the nested CV at every iteration of the outer loop CV.

Subject-wise pairwise decoding model

In the subject-wise approach, decoding models were independently built for each subject. For each iteration of the outer loop CV, EEG trials of a given subject for a given odor pair were randomly split into a training and test set at a ratio of 9:1 (the number of trials used is shown in Table S4). To avoid classification bias due to an unbalanced number of trials, the number of trials was adjusted using random undersampling for both the training and test sets. The range of each of the 1280 features was scaled across trials using a scaling factor estimated based on a training set using the minimum-maximum (min-max) scaling method. The training set was further split into 10-folds for the inner loop CV to choose the best value for the cost parameter λ. Finally, the performance of the decoding model built using the selected λ and training set was evaluated using the test set. This entire procedure was repeated 60 times (i.e., number of outer loops, CV = 60).

For each subject and each odor pair, the decoding accuracy was measured as the mean percentage of correct trials across the outer CVs. The grand average across-odor-pair decoding accuracy shown in Fig. 2B was calculated by first obtaining the across-odor-pair mean accuracy for each subject and then averaging them across subjects. To test whether the decoding performance was significantly above the chance level (50%), a one-sided, one-sample Student's t-test was used, with the number of subjects (n = 22) being the number of observations. The alpha level was set at 0.05, after applying the Bonferroni–Holm correction for the number of time

points. The grand average decoding accuracy for each odor pair shown in Fig. 2D (upper triangle) was calculated by averaging the decoding accuracies of each odor pair across the subjects.

Cross-subject pairwise decoding model

In the cross-subject approach, decoding models were generalized across subjects using leave-one-subject-out CV. For each iteration of the outer loop CV, EEG trials for a given odor pair from 21 subjects were used for training, and those of the remaining one subject were used for testing. The range of each of the 1280 features was scaled across trials, which were pooled from subjects, using a scaling factor estimated based on the training set using the min-max scaling method. The training set was further split into 10-folds for the inner loop CV to choose the best value for the cost parameter λ. Finally, the performance of the decoding model built using the selected λ and training set was evaluated using the test set. The entire procedure was repeated until all subjects were used as a test set (i.e., the number of outer loops CV = 22).

For each of the outer loop CVs, for each odor pair, the decoding accuracy was measured as the mean percentage of correct trials. The grand average decoding accuracy across odor pairs shown in Fig. 2C was calculated by first obtaining the across-odor-pair mean accuracy for each of the outer loop CVs and averaging them across CVs. To test whether the decoding performance was significantly above the chance level (50%), a one-sided, one-sample Student's t-test was used, with the number of outer loop CVs (n = 22) being the number of observations. After applying the Bonferroni–Holm correction, the alpha level was set at 0.05. The grand average decoding accuracy for each odor pair shown in Fig. 2D (lower triangle) was calculated by averaging the decoding accuracies of each odor pair across the outer loop CVs.

Multi-class decoding

In both the subject-wise and cross-subject approaches, a ten-class classification model was built for each time point. An ℓ^2-regularized multinomial logistic regression classifier, provided by scikit-learn python module (version 0.23.2), was used. As in the pairwise models, a nested cross-validation (CV) procedure was used, where an inner loop CV responsible for tuning the cost parameter was nested in an outer loop CV, which was responsible for model evaluation. The training algorithm used the cross-entropy loss, and the cost parameter, λ, was chosen from 6 possible values (log-spaced, between 2^{-3} and 2^{12}) based on the 10-fold CV in the inner loop of the nested CV at every iteration of the outer loop CV. The temporal resolution of EEG data and features used for each model were the same as those of the pairwise models.

Subject-wise multi-class model

In the subject-wise approach, decoding models were independently built for each subject. For each iteration of the outer loop CV, EEG trials of a given subject for each of the 10 odors were split into training and test sets at a ratio of 9:1. To avoid classification bias, stratified random sampling was performed to create training and test sets with the same class ratio. The remaining procedures (feature scaling, selection of λ and number of outer CVs) were the same as those of subject-wise pairwise decoding.

For each subject, the decoding accuracy was measured as a mean percentage of correct trials across the outer CVs (CV = 60). The grand average decoding accuracy shown in Fig. S3A was calculated by first obtaining the across-odor mean accuracy for each subject and then averaging them across subjects. To test whether the decoding performance was significantly above the chance level (10%), a one-sided, one-sample Student's t-test was used, with the number of subjects (n = 22) being the number of observations. The alpha level was set at 0.05, after applying the Bonferroni–Holm correction for the number of time points. The grand average confusion matrix shown in Fig. S3B was calculated by averaging the confusion matrices across subjects.

Cross-subject multi-class model

In the cross-subject approach, decoding models were generalized across subjects using leave-one-subject-out CV. For each iteration of the outer loop CV, EEG trials for all 10 odors from 21 subjects were used for training, and those of the remaining one subject were used for testing.

Remaining procedures (feature scaling, selection of λ and number of outer CVs) were the same as those of cross-subject pairwise decoding.

For each of the outer loop CVs (CV = 22), the decoding accuracy was measured as a mean percentage of correct trials. The grand average decoding accuracy across odors in Fig. S3C was calculated by first obtaining the across-odor mean accuracy for each of the outer loop CVs and then averaging them across CVs. To test whether the decoding performance was significantly above the chance level (10%), a one-sided, one-sample Student's t-test was used, with the number of outer loop CVs (n = 22) being the number of observations. After applying the Bonferroni–Holm correction, the alpha level was set at 0.05. The grand average confusion matrix shown in Fig. S3D was calculated by averaging the confusion matrices across the outer loop CVs.

Time-resolved representational similarity analysis (tRSA). To examine how the sensor-level, source-level, and perceptual representations of odors relate to each other, we conducted a representational similarity analysis (RSA) (36, 37) in a time-resolved manner (Fig. 1F). In the RSA framework, a representational structure of a given concept (e.g., odor) in a given space (e.g., neural space) can be defined as a representational dissimilarity vector (RDV), which consists of pairwise distances of all pairs of samples of that concept (e.g., odors used in the experiment) in that space. Then, similarities between different representational structures, for example, odors in perceptual space vs. those in neural space, are examined as correlations between RDVs.

In the current study, representational structures were defined based on the data for all the 45 possible odor pairs, and thus, each RDV comprised 45 elements (specific variables used are explained in the following subsections). When RDVs were constructed based on decoding accuracies or trigeminal test scores, we replaced the values under the chance level (50%) with 50%. As recommended for RSA (36, 38), relationships between representational structures were examined using Spearman's rank correlation, and the statistical significance of correlation coefficients was tested using a resampling-based method. In particular, null distributions of the correlation coefficients were obtained by randomly shuffling the odor labels of the real data and computing rank correlations between RDVs constructed using the shuffled data. This procedure was repeated 10,000 times. The *p* values were computed as a proportion of simulated correlation coefficients that were larger than the actual correlation in all simulated correlation coefficients (one-sided test) to test whether the representational dissimilarity vectors were positively correlated or not (e.g., odor pairs with more similar perception evoke more similar brain activities). Multiple testing corrections were applied, as described in the following subsections. tRSA was conducted every 50 ms at the same temporal resolution as tMVPA decoding.

Sensor-level tRSA based on the subject-wise model

To examine whether and when the neural representational structure revealed by tMVPA decoding relates to the perceptual characteristics of odors, we performed sensor-level tRSA, where correlations between RDVs based on decoding accuracies and perceptual scores were examined (Fig. 3). Given the large individual differences in odor perception (Fig. 1 C–E), we primarily focused on tRSA based on subject-wise models (Fig. 3, Fig. S4A). In this analysis, RDVs were constructed by first computing individual-level RDVs, each consisting of 45 odor pairs, and averaging them across subjects (Fig. 3A). To better reflect individual differences, the order of the odor pairs in individual-level RDVs was sorted according to the decoding accuracies of a given subject for a given time point before averaging across subjects.

In particular, for each time point for each subject, an individual-level brain RDV was constructed using pairwise decoding accuracies based on the decoding models of the subject. An individual-level perceptual RDV was constructed using a subject's own perceptual score(s) for each of the following perceptual aspects: quality, unipolar pleasantness, unipolar unpleasantness, intensity, similarity, pleasantness rated during EEG recordings, and trigeminal stimulation. For similarity, pairwise odor dissimilarities were quantified using reversed pairwise similarity rating scores. For the other perceptual aspects, pairwise odor dissimilarities were quantified as pairwise Euclidean distances of corresponding rating scores (for trigeminal and trigeminal test scores). Then, these

individual-level RDVs were averaged across subjects to construct brain and perceptual RDVs, as described in the previous paragraph, and were used for correlation analysis. In the case of rank-transformed RSA, every RDV was rank-transformed before averaging (Fig. S4B). This made the RDV of each subject a vector consisting of numbers 1 to 45. In this way, all subjects contributed equally to averaged RDVs. To examine the time course of neural representations associated with each of the perceptual aspects, zero-order correlation analysis was performed between the brain RDVs and each of the perceptual RDVs described above. To identify the time course of neural representations associated with specific aspects of perception, partial correlation analysis was performed between the brain RDVs and each of the following perceptual RDVs: unipolar pleasantness, unipolar unpleasantness, and quality. For unipolar pleasantness and unipolar unpleasantness, intensity was partialled out. For quality, unipolar pleasantness, unipolar unpleasantness, and intensity were partialled out. Significance thresholds were corrected for the number of time points using the Bonferroni–Holm method.

Sensor-level tRSA based on cross-subject model

Although large individual differences existed, when looking at the group level, sensory characteristics significantly differed across odors (Fig. 1 C–E). To examine whether such perceptual differences of odors are associated with the neural representational structures common across subjects (Fig. 2 D, lower triangle), we also conducted the tRSA based on inter-subject mean sensory scores and the cross-subject decoding models (Fig. S4E). In this analysis, brain RDVs were constructed based on pairwise cross-subject decoding accuracies, and perceptual RDVs were constructed by computing the inter-subject mean pairwise Euclidean distances for each of the following perceptual aspects: quality, unipolar pleasantness, unipolar unpleasantness, and intensity. Zero-order correlation analysis was performed between the brain RDVs and each of the following perceptual RDVs: quality, unipolar pleasantness, and unipolar unpleasantness. Partial correlation analysis was performed between the brain RDVs and each of the following perceptual RDVs: unipolar pleasantness while partialling out intensity, unipolar unpleasantness while partialling out intensity, and quality while partialling out unipolar pleasantness, unipolar unpleasantness, and intensity. Significance thresholds were corrected for the number of time points using the Bonferroni–Holm method.

tRSA between source-level activity and sensor-level decoding accuracy (source-sensor tRSA)

To gain insights into the cortical generators underlying the neural representational structure revealed by sensor-level tMVPA decoding, we performed source-sensor tRSA, where correlations between RDVs based on source-level brain activities and sensor-level tMVPA decoding accuracies were examined (Fig. 4). In this analysis, we used decoding accuracies based on the subject-wise models. RDVs were constructed by first computing the individual-level RDVs, each consisting of 45 odor pairs, and averaging them across subjects. In particular, for each time point for each subject, an individual-level, source-level brain RDV was constructed for each of the 6239 voxels by computing the pairwise differences in the standardized current density powers for all possible 45 odor pairs. An individual-level, sensor-level brain RDV was constructed based on the decoding accuracies obtained by the subject's own decoding models for the given time point. For both source-level and sensor-level brain RDVs, odor pairs of individual-level RDVs were sorted according to the source-level dissimilarities of the given time point and were then averaged across subjects. Significance thresholds were corrected for the number of time points and number of voxels using the Benjamini–Hochberg false discovery rate (FDR) controlling procedure.

tRSA between source-level activity and perceptual ratings

To examine when and where different aspects of odor perception are represented in the brain, we performed tRSA between source-level activity and perceptual ratings. The procedures for constructing RDVs were the same as those for the source-sensor tRSA, except that the RDVs based on perceptual scores were used instead of sensor-level brain RDVs. In particular, for each time point for each subject, an individual-level, source-level brain RDV was constructed for each voxel, as described in the previous subsection. An individual-level perceptual RDV was constructed using the subject's own perceptual score(s) for each of the following perceptual

aspects: quality, unipolar pleasantness, unipolar unpleasantness, and intensity. The odor pairs of individual-level RDVs were sorted according to the source-level dissimilarities of the given time point and then averaged across subjects. To reveal the source activities associated with holistic perception, zero-order correlations between source-level brain RDVs and RDVs based on quality ratings were examined (source-quality tRSA). To identify source activities associated with specific aspects of perception, partial correlation analysis was performed between the source-level brain RDVs and each of the following perceptual RDVs: unipolar pleasantness, unipolar unpleasantness, and quality. For unipolar pleasantness and unipolar unpleasantness, the intensity was partialled out. For quality, unipolar pleasantness, unipolar unpleasantness, and intensity were partialled out. The analysis was limited to the significant time points in the corresponding sensor-level tRSA (Fig. 3B, C), that is, for the zero-order correlation analysis using quality, 400–1500 ms; for partial correlation analysis focusing on unipolar pleasantness, 500–1500 ms; unipolar unpleasantness, 300–1500 ms; and quality, 500–600 ms. Significance thresholds were corrected for the number of time points and number of voxels using the Benjamini–Hochberg FDR procedure.

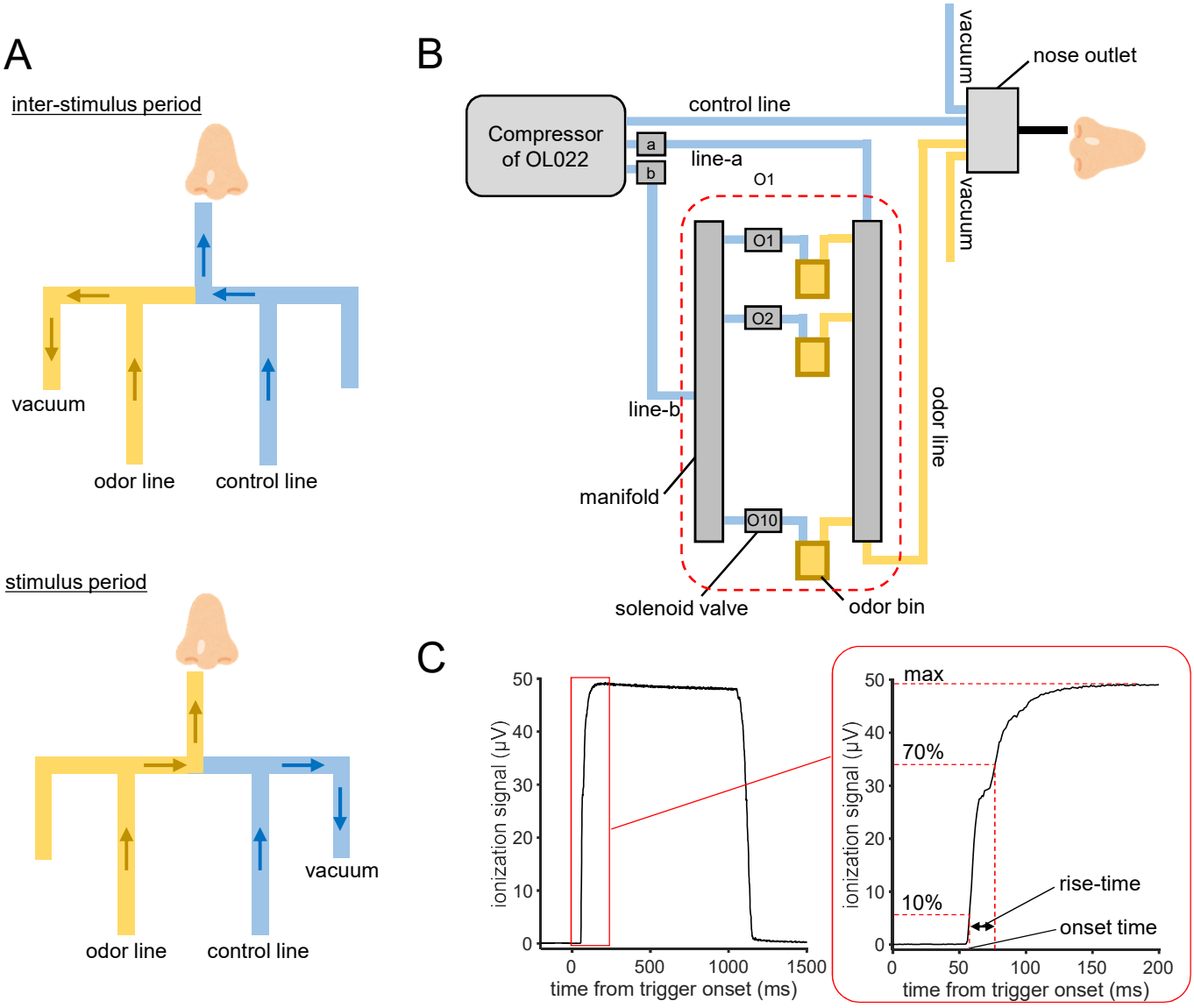


Fig. S1. Olfactometer setup (A) Schematic representation of Kobal's switching principle used in the olfactometer (OL022; Burghart Messtechnik GmbH). This switching principle allows odor presentation without tactile stimulation, which has been recommended for OERP measurement. In this switching principle, both odor and control lines were active throughout the experimental session, and the switching between lines occurred at the outlet of the olfactometer by vacuuming either odorized air (during inter-stimulus period) or control air (during the stimulus period). (B) Expansion of OL022. A set of manifolds and valves, shown in the red dotted box, were inserted into the upper stream of the outlet to represent 10 different odors without affecting the Kobal's switching principle. Presentation of different odors was achieved by opening one of the valves (valves O1-O10) placed downstream of line-b. Five seconds before odor onset, valve-b was opened to allow airflow in line-b, which joined the odor line after passing one of the odor bins. After the odor presentation period, valve-b was closed, while valve-a was opened, and airflow from line-a joined the odor line after passing the manifold and rinsing it. Importantly, the olfactometer's original function, which is based on Kobal's switching principle, was used for switching between the control and odor lines at the nose outlet. (C) The performance of the modified olfactometer was examined by measuring the odor onset time and odor rise time and their variability across odor channels and trials using a fast-response photo-ionization detector (mini-PID Model 200B, Aurora Scientific Inc., Ontario, Canada), as described previously (39). A time course of PID signals from a representative trial is shown. Timing of valve switching was defined as the onset of the transistor-transistor logic (TTL) signal that triggered valve switching. The odor onset time was defined as the time point at which the PID value was above 10% of the maximum value, and the rise-time was defined as the time elapsed from 10% to 70% of the maximum value. The TTL and PID signals were recorded at a sampling rate of 1000 Hz using a data acquisition system (MP150, BIOPAC Systems, Inc., Goleta, Calif, USA). The setting of the olfactometer was the same as that in the EEG session (flow rates, 7.5 L/min for both baseline and odor air flows; odor duration, 1000 ms, ISIs; 17 s), and the number of trials for each channel was 10. The mean odor onset time across channels was 59.6±0.4 ms, and the rise-time was 38.6±2.3 ms. The mean of the S.D. of odor onset time across trials was 1.1±0.3 ms and that of rise-time was 2.6±0.8 ms. These results were compatible with those of the original olfactometer and met the recommendations for OERP measurements (40).

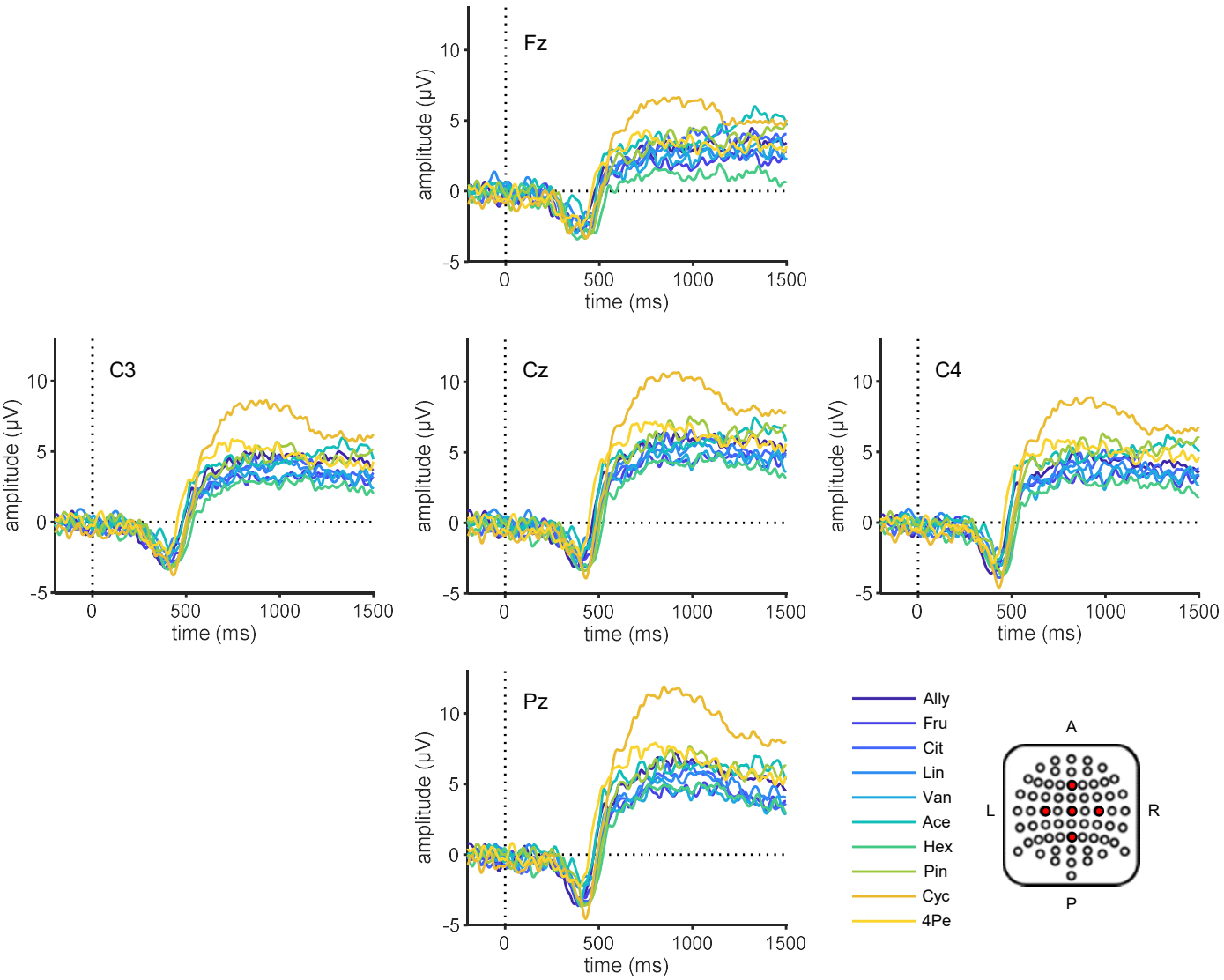


Fig. S2. Grand average OERP waveforms for each of the 10 odors at five representative electrodes. The red circles in the inset indicate the locations of the selected electrodes. Each waveform was referenced to the average of the mastoid electrodes. A, anterior; P, posterior; L, left hemisphere; R, right hemisphere.

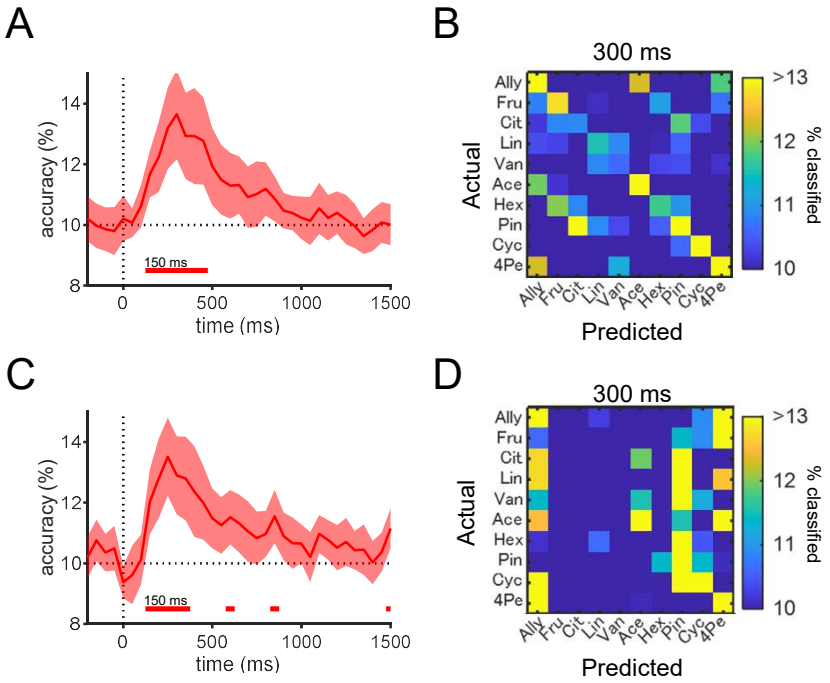


Fig. S3. Multi-class (10-class) decoding with multinomial logistic regression classifier. The number of cross-validations and features used for decoding analysis were the same as those of pairwise decoding (Fig. 2; see SI method for details). The grand mean accuracy of subject-wise decoding (A) and cross-subject decoding (C) averaged across 10 odors. Bottom horizontal lines indicate statistical significance (one-sided, one-sample Student's t-test, $p < 0.05$, Bonferroni–Holm corrected), with numbers indicating their onsets. Shaded areas indicate 95% confidential intervals across subjects (A) or across outer loop CV (C). Confusion matrices for subject-wise decoding (B) and cross-subject decoding (D) at 300 ms after odor onset, when the grand mean accuracies reached maxima. Rows represent actual class, and columns represent predicted class. The diagonal and off-diagonal cells correspond to correct and incorrect classifications, respectively. For each row, the color of each cell indicates the percentage of trials that were classified to the given predicted class over the total number of trials for the corresponding actual class, so that sum of cells for each row equals 100%. Abbreviations of odors are shown at the left and bottom (see Methods for full names).

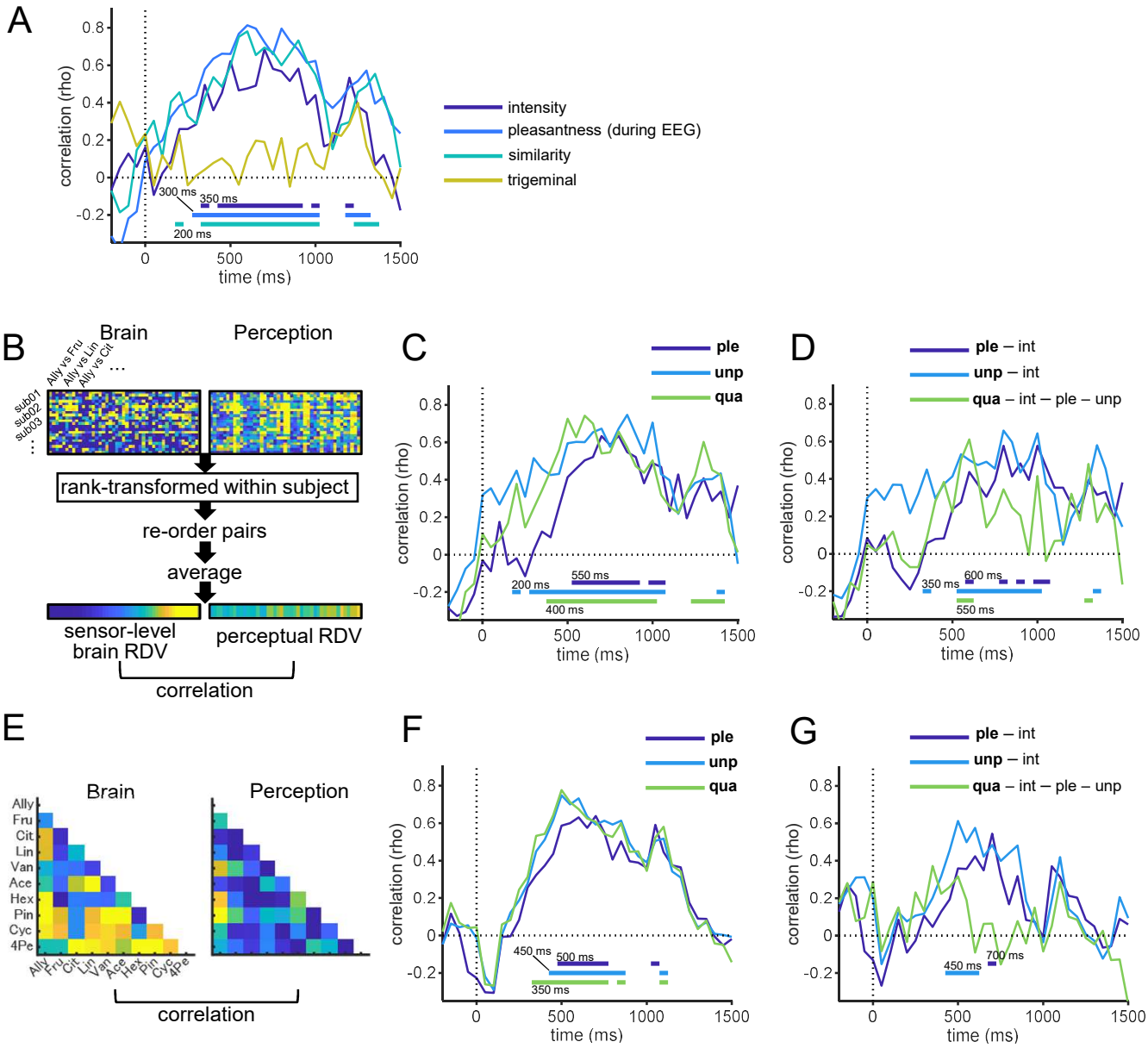


Fig. S4. Sensor-level tRSA based on subject-wise (A–D), and cross-subject decoding (E–G). (A) The remaining results of sensor-level tRSA presented in Fig. 3B. Sensor-level brain representational dissimilarity vectors (RDVs) computed as shown in Fig. 3A were rank correlated with each of the 4 RDVs computed based on perceptual scores: intensity ratings, pleasantness ratings obtained during EEG recordings, similarity ratings, and performances on the trigeminal test. The bottom horizontal lines indicate statistical significance (permutation test, 10,000 permutations, $p < 0.05$, Bonferroni-Holm corrected), with numbers indicating their onsets. (B–D) To confirm that a small number of subjects did not bias the RDVs, every RDV was rank-transformed before averaging across subjects. This made the RDV of each subject a vector consisting of numbers 1 to 45. In this way, all subjects contributed equally to averaged RDVs. The remaining processes used for (C) (zero-order correlation) and (D) (partial correlation) were the same as those used for Fig. 3B and Fig. 3C, respectively. (E-G) To examine whether neural representations of odor perception that are common across subjects exist, we also conducted tRSA based on the decoding accuracies obtained by the cross-subject decoding models and inter-subject mean perceptual scores. (E) Procedures for conducing tRSA based on cross-subject decoding. The brain RDVs were constructed based on pairwise cross-subject decoding accuracies, and the perceptual RDVs were constructed by computing the inter-subject mean pairwise Euclidean distances of the given perceptual score for all odor pairs. Then, the Spearman's rank correlation between the brain and the perceptual RDVs were computed. Results of tRSA with zero-order correlation (F) and with partial correlation (G). For unipolar pleasantness and unpleasantness, the intensity was partialled out, and for quality, intensity, unipolar pleasantness and unpleasantness were partialled out. The bottom horizontal lines indicate statistical significance (permutation test, 10,000 permutations, $p < 0.05$, Bonferroni-Holm corrected), with numbers indicating their onsets. Ple, unipolar pleasantness; unp, unipolar unpleasantness; qua, quality; int, intensity.

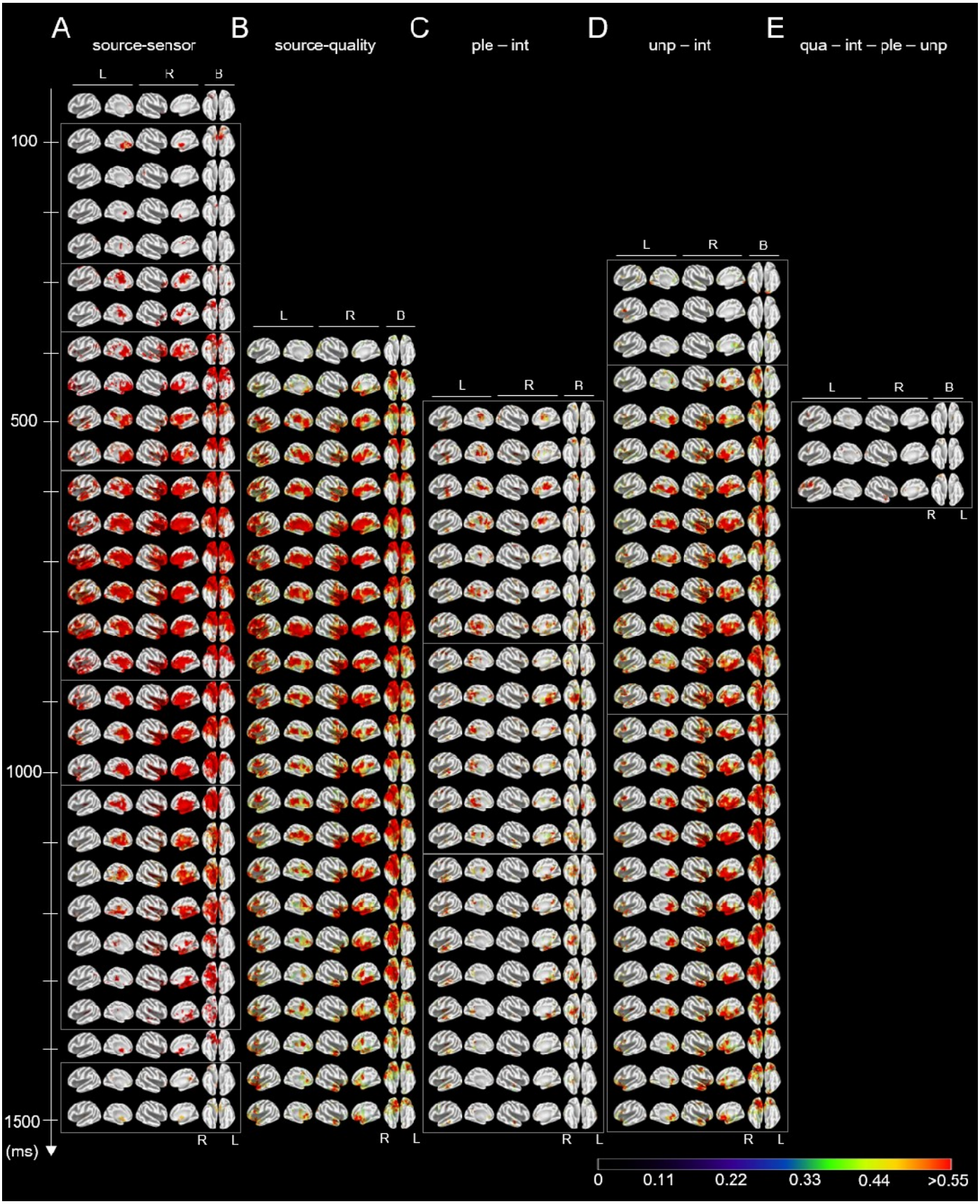
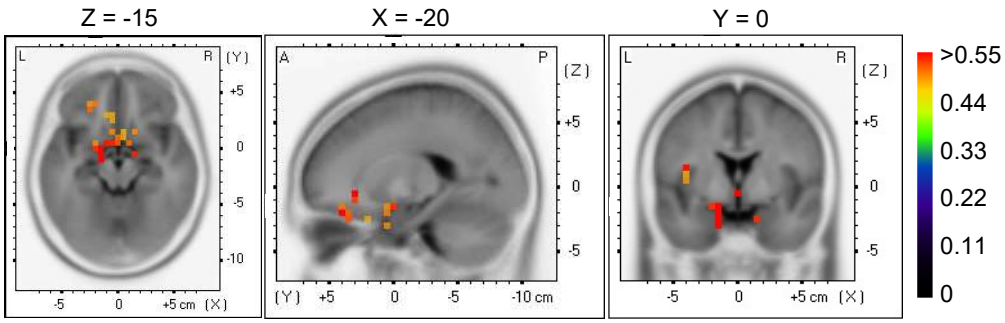


Fig. S5. Results of source-level tRSA for all the time-points. Results for tRSA relating to representational structure based on source activities with that of sensor-level decoding accuracies (source-sensor tRSA; (A)) and quality ratings as a metric of perceptual dissimilarities in the whole olfactory perceptual space (source-quality tRSA; (B)), unipolar pleasantness while partialling out intensity (C), unipolar unpleasantness while partialling out intensity (D), and quality while partialling out intensity, unipolar pleasantness, and unipolar unpleasantness (E). tRSA was conducted as described in Fig. 4A. Significant voxels were colored according to the color bar (permutation test, 10,000 permutations, $p < 0.05$, FDR corrected for all voxels and time-points). Colors indicate correlation coefficients (A–B) or partial correlation coefficients (C–E). Ple, unipolar pleasantness; unp, unipolar unpleasantness; qua, quality; int, intensity; L, left hemisphere; R, right hemisphere; B, bottom view.

A

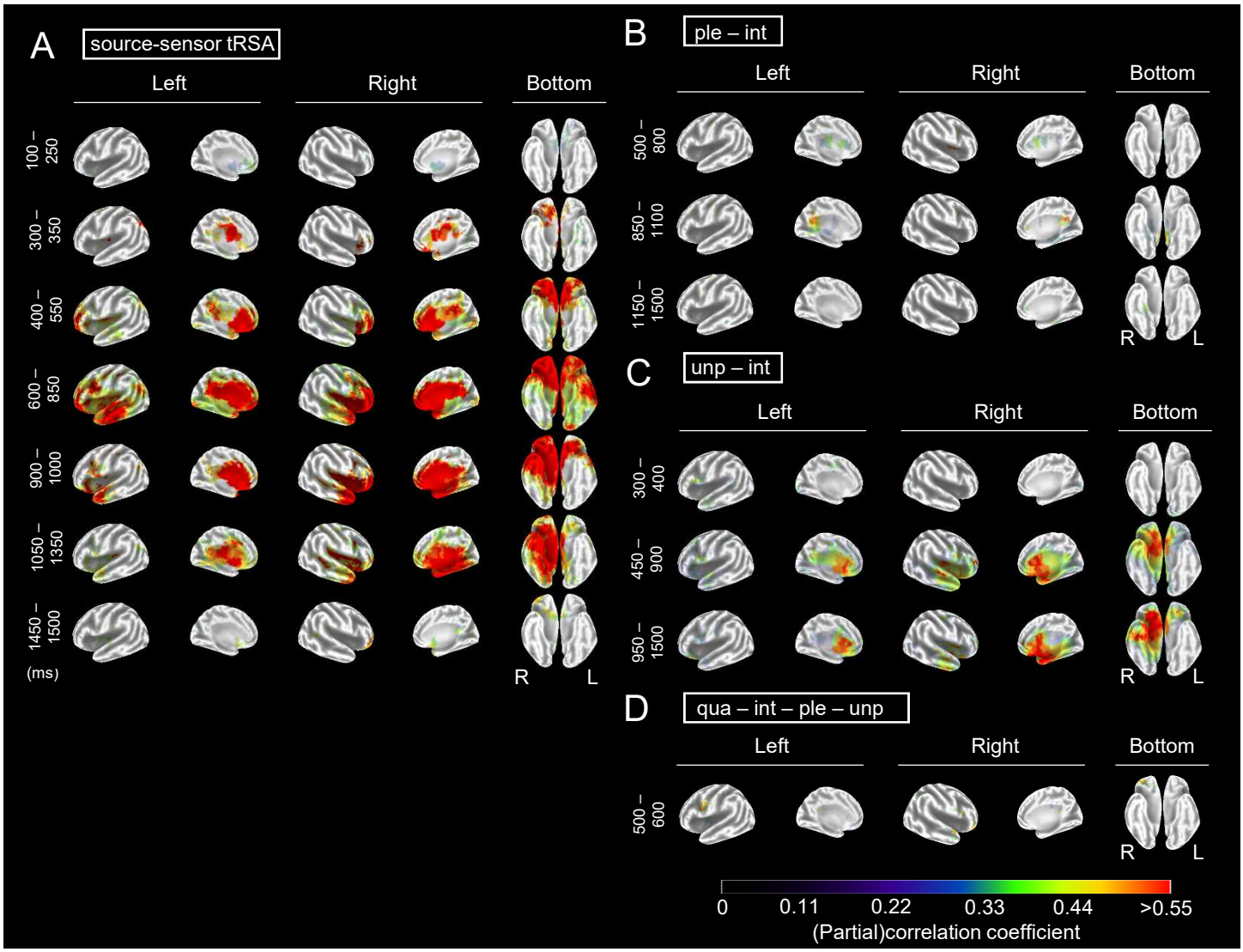


B

X	Y	Z	rho	p	HS	Lobe	BA	Structure
-10	40	-10	0.48	0.042	L	F	10	MedFG
-10	40	-5	0.54	0.010	L	F	10	MedFG
-10	50	-5	0.58	0.001	L	F	10	MedFG
-5	45	-10	0.52	0.019	L	F	10	MedFG
-5	50	-5	0.51	0.022	L	F	10	MedFG
-5	55	0	0.48	0.042	L	F	10	MedFG
-25	35	-25	0.50	0.030	L	F	11	IFG
-25	35	-20	0.58	0.002	L	F	11	IFG
-20	35	-20	0.51	0.020	L	F	11	IFG
-15	30	-25	0.50	0.028	L	F	11	IFG
-15	35	-25	0.54	0.012	L	F	11	IFG
-15	35	-20	0.58	0.001	L	F	11	IFG
-15	40	-20	0.52	0.019	L	F	11	IFG
-5	25	-15	0.48	0.042	L	F	11	MedFG
-5	30	-15	0.48	0.045	L	F	11	MedFG
-5	50	-10	0.55	0.005	L	F	11	MedFG
-25	35	-15	0.51	0.022	L	F	11	MidFG
-25	40	-20	0.49	0.032	L	F	11	MidFG
-25	40	-15	0.49	0.037	L	F	11	MidFG
-25	40	-10	0.51	0.022	L	F	11	MidFG
-25	40	-5	0.52	0.018	L	F	11	MidFG
-20	40	-15	0.51	0.026	L	F	11	MidFG
-15	40	-25	0.50	0.028	L	F	11	OrbitalG
-15	45	-25	0.51	0.024	L	F	11	OrbitalG
-10	25	-30	0.48	0.042	L	F	11	RectalG
-10	30	-25	0.56	0.004	L	F	11	RectalG
-10	35	-25	0.59	0.000	L	F	11	RectalG
-10	35	-20	0.49	0.033	L	F	11	RectalG
-5	15	-25	0.48	0.047	L	F	11	RectalG
-5	20	-25	0.52	0.019	L	F	11	RectalG
-5	25	-25	0.58	0.001	L	F	11	RectalG
-5	30	-25	0.55	0.007	L	F	11	RectalG
-20	40	-20	0.56	0.004	L	F	11	SFG
-15	60	-10	0.51	0.022	L	F	11	SFG
-10	10	-20	0.62	0.000	L	F	25	MedFG
-10	25	-20	0.53	0.013	L	F	25	MedFG
-10	30	-15	0.48	0.042	L	F	25	MedFG
-5	10	-20	0.51	0.021	L	F	25	MedFG
-5	25	-20	0.51	0.025	L	F	25	MedFG
-5	30	-20	0.49	0.032	L	F	25	MedFG
-5	5	-15	0.54	0.012	L	F	25	Subcallosal
-5	15	-15	0.50	0.032	L	F	25	Subcallosal
-20	5	-15	0.50	0.032	L	F	34	Subcallosal
-10	5	-15	0.54	0.011	L	F	34	Subcallosal
-35	20	-5	0.48	0.046	L	F	47	IFG
-30	20	-25	0.51	0.023	L	F	47	IFG
-25	30	-10	0.49	0.039	L	F	47	IFG
-25	35	-10	0.55	0.007	L	F	47	IFG
-25	35	-5	0.53	0.015	L	F	47	IFG
-20	20	-25	0.48	0.043	L	F	47	IFG
-20	30	-10	0.53	0.015	L	F	47	IFG
-20	30	-5	0.64	0.000	L	F	47	IFG
-20	35	-25	0.52	0.017	L	F	47	OrbitalG
-5	25	15	0.56	0.002	L	Lim	24	ACC
-5	25	20	0.54	0.013	L	Lim	24	ACC
-5	30	-5	0.52	0.016	L	Lim	24	ACC
-5	30	0	0.57	0.002	L	Lim	24	ACC
-5	30	15	0.59	0.000	L	Lim	24	ACC
-5	35	5	0.59	0.000	L	Lim	24	ACC
-5	35	10	0.53	0.015	L	Lim	24	ACC
-5	15	-10	0.50	0.028	L	Lim	25	ACC
-5	20	-5	0.52	0.019	L	Lim	25	ACC
-15	-10	-15	0.62	0.000	L	Lim	28	PHG
-15	-5	-15	0.58	0.001	L	Lim	28	PHG

X	Y	Z	rho	p	HS	Lobe	BA	Structure
-20	5	-30	0.50	0.032	L	Lim	28	Uncus
-15	0	-30	0.63	0.000	L	Lim	28	Uncus
-10	35	-10	0.56	0.002	L	Lim	32	ACC
-10	35	-5	0.62	0.000	L	Lim	32	ACC
-10	45	-5	0.53	0.015	L	Lim	32	ACC
-10	45	0	0.53	0.013	L	Lim	32	ACC
-10	45	5	0.61	0.000	L	Lim	32	ACC
-5	25	-10	0.55	0.008	L	Lim	32	ACC
-5	30	-10	0.56	0.002	L	Lim	32	ACC
-5	35	-5	0.60	0.000	L	Lim	32	ACC
-5	35	0	0.62	0.000	L	Lim	32	ACC
-5	40	-10	0.61	0.000	L	Lim	32	ACC
-5	40	-5	0.67	0.000	L	Lim	32	ACC
-5	40	0	0.61	0.000	L	Lim	32	ACC
-5	40	5	0.48	0.042	L	Lim	32	ACC
-5	45	-5	0.60	0.000	L	Lim	32	ACC
-5	45	0	0.54	0.011	L	Lim	32	ACC
-5	10	25	0.52	0.016	L	Lim	33	ACC
-5	20	20	0.50	0.027	L	Lim	33	ACC
-20	0	-15	0.53	0.013	L	Lim	34	PHG
-20	5	-20	0.50	0.028	L	Lim	34	PHG
-15	-5	-20	0.56	0.004	L	Lim	34	PHG
-15	0	-20	0.57	0.002	L	Lim	34	PHG
-15	0	-15	0.62	0.000	L	Lim	34	PHG
-15	-5	-25	0.53	0.015	L	Lim	34	Uncus
-15	0	-25	0.61	0.000	L	Lim	34	Uncus
-15	5	-25	0.52	0.017	L	Lim	34	Uncus
-25	-50	50	0.50	0.032	L	P	7	Precuneus
-35	-55	45	0.50	0.031	L	P	40	IPL
-40	-45	20	0.56	0.002	L	S	13	Insula
-40	0	5	0.49	0.034	L	S	13	Insula
-40	0	10	0.48	0.044	L	S	13	Insula
-40	0	15	0.55	0.008	L	S	13	Insula
-40	5	15	0.54	0.013	L	S	13	Insula
0	5	-15	0.52	0.019	C	F	25	Subcallosal
0	10	-15	0.51	0.025	C	F	25	Subcallosal
0	0	-5	0.58	0.000	C	Lim	25	ACC
0	5	-10	0.51	0.022	C	Lim	25	ACC
0	10	-10	0.48	0.046	C	Lim	25	ACC
0	10	-5	0.51	0.025	C	Lim	25	ACC
0	15	-10	0.48	0.045	C	Lim	25	ACC
0	45	0	0.49	0.037	C	Lim	32	ACC
30	-30	60	0.49	0.040	R	F	4	Precentral
15	35	-25	0.56	0.002	R	F	11	IFG
15	35	-20	0.51	0.022	R	F	11	IFG
15	40	-25	0.48	0.042	R	F	11	OrbitalG
10	15	-20	0.48	0.046	R	F	11	RectalG
20	45	-20	0.49	0.034	R	F	11	SFG
15	15	-15	0.49	0.032	R	F	13	Subcallosal
10	10	-20	0.48	0.042	R	F	25	MedFG
15	15	-20	0.53	0.015	R	F	25	MedFG
5	10	-15	0.48	0.045	R	F	25	Subcallosal
5	15	-15	0.47	0.048	R	F	25	Subcallosal
10	5	-15	0.51	0.024	R	F	34	Subcallosal
5	5	-5	0.55	0.005	R	Lim	25	ACC
15	-5	-15	0.54	0.012	R	Lim	28	PHG
5	45	0	0.52	0.016	R	Lim	32	ACC
15	-10	-20	0.53	0.015	R	Lim	34	PHG
15	-5	-20	0.50	0.028	R	Lim	34	PHG
20	-10	-20	0.51	0.021	R	Lim	34	PHG
15	-5	-25	0.53	0.015	R	Lim	34	Uncus
15	0	-25	0.54	0.013	R	Lim	34	Uncus
20	-5	-35	0.53	0.015	R	Lim	36	Uncus
25	-35	55	0.52	0.019	R	P	3	Postcentral

Fig. S6. Significant voxels in the source-sensor tRSA at 100 ms (Permutation test, 10,000 permutations, $p < 0.05$, FDR corrected for all the voxels and time-points). (A) From left to right, axial, sagittal, and coronal slices that go through the POA. Colors indicate correlation coefficients. Only significant voxels are colored. (B) List of significant voxels with MNI coordinates (X, Y and Z), correlation coefficients (rho), *p*-values, hemispheres (HS), lobes, Brodmann areas (BA) and anatomical structures. Anatomical information was obtained using the sLORETA software. Yellow highlights indicate voxels that are in the POA (either piriform cortex, anterior olfactory nucleus or olfactory tubercle). Location of the POA was determined based on a previous study that reported MNI coordinates of the POA (see SI method for details). A, anterior; P, posterior; L, left hemisphere; R, right hemisphere; C, center; F, frontal lobe; Lim, limbic lobe; P, parietal lobe; S, sub-lobar; MedFG, medial frontal gyrus; IFG, inferior frontal gyrus; MidFG, middle frontal gyrus; OrbitalG, orbital gyrus; RectalG, rectal gyrus; SFG, superior frontal gyrus; Subcallosal, subcallosal gyrus; ACC, anterior cingulate cortex; PHG, parahippocampal gyrus; IPL, inferior parietal lobule.



E

source-sensor tRSA

Time (ms)	r	corrected *p*
100-250	0.90	< 0.05
300-350	0.91	< 0.05
400-550	0.96	< 0.05
600-850	0.98	< 0.05
900-1000	0.97	< 0.05
1050-1350	0.97	< 0.05
1450-1500	0.91	< 0.05

pre – int

Time (ms)	r	corrected *p*
500-800	0.78	< 0.05
850-1100	0.80	< 0.05
1150-1500	0.84	< 0.05

unp – int

Time (ms)	r	corrected *p*
300-400	0.87	< 0.05
450-900	0.94	< 0.05
950-1500	0.93	< 0.05

qua – ple – unp – int

Time (ms)	r	corrected *p*
500-600	0.73	< 0.05

Fig. S7. Rank-transformed source-level tRSA. To confirm that a small number of subjects did not bias the RDVs, every RDV was rank-transformed before averaging across subjects for each voxel (see Fig. S4B for the procedure). The remaining processes used for (A) zero-order correlation and (B–D) partial correlation were the same as those used for Fig. 4B and Fig. 4C–E, respectively. (E) Correlation between results of tRSA with and without rank-transformation (i.e., results shown in Fig. S7 and Fig. 4). Pearson's correlation coefficients were obtained using a pair of vectors comprised of Spearman's correlation coefficients for all the 6239 voxels, and *p*-values were Bonfferoni-Holm corrected for the number of time points. Significant strong positive correlations indicated that the spatiotemporal patterns of rank-transformed tRSA were highly similar with those without rank-transformation. Ple, unipolar pleasantness; unp, unipolar unpleasantness; qua, quality; int, intensity; R, right hemisphere; L, left hemisphere.

Table S1. Perceptual rating and trigeminal scores.

	intensity	pleasantness	trigeminal
All	3.3 (0.5)	3.4 (0.7)	57.3 (20.1)
Fru	3.1 (0.6)	2.8 (0.8)	60.4 (18.4)
Cit	2.9 (0.5)	3.5 (0.6)	51.4 (19.0)
Lin	2.9 (0.7)	3.2 (0.5)	51.7 (17.5)
Van	3.2 (0.5)	2.9 (0.7)	59.0 (19.8)
Ace	3.3 (0.8)	2.9 (0.9)	57.3 (19.4)
Hex	3.1 (0.8)	2.6 (0.7)	56.1 (19.5)
Pin	3.4 (0.7)	3.0 (0.8)	62.0 (19.8)
Cyc	4.5 (0.7)	1.4 (0.5)	53.6 (17.0)
4Pe	3.7 (0.8)	2.0 (0.8)	54.1 (17.5)

Inter-subject means and standard deviations (S.D.) of the intensity ratings were obtained using a 5-point scale (1 = odorless; 5 = extremely strong), pleasantness ratings were obtained using a bi-polar 5-point scale (1 = very unpleasant; 5 = very pleasant), and accuracies were obtained using the trigeminal test (% of correct trials). The pleasantness ratings shown here were obtained during the EEG recordings (Fig. 1).

Table S2. Similarity rating scores.

	Ally	Fru	Cit	Lin	Van	Ace	Hex	Pin	Cyc	4Pe
Ally	–	–	–	–	–	–	–	–	–	–
Fru	66.4 (19.4)	–	–	–	–	–	–	–	–	–
Cit	65.3 (13.4)	60.5 (18.6)	–	–	–	–	–	–	–	–
Lin	59.4 (14.7)	54.5 (17.0)	63.5 (19.7)	–	–	–	–	–	–	–
Van	50.5 (21.6)	53.7 (20.2)	53.0 (15.3)	53.7 (20.1)	–	–	–	–	–	–
Ace	58.5 (19.7)	48.3 (16.5)	53.6 (16.1)	47.9 (21.1)	48.8 (16.0)	–	–	–	–	–
Hex	55.3 (18.8)	61.0 (16.4)	52.2 (16.1)	48.8 (16.8)	55.5 (15.8)	47.8 (22.3)	–	–	–	–
Pin	48.9 (21.5)	54.3 (21.5)	56.8 (21.3)	50.9 (18.1)	42.2 (16.9)	54.7 (18.4)	45.8 (15.9)	–	–	–
Cyc	14.2 (13.6)	21.2 (17.8)	18.5 (18.1)	24.3 (18.2)	22.0 (18.4)	24.1 (18.1)	21.5 (18.4)	18.5 (14.7)	–	–
4Pe	42.7 (30.8)	40.6 (22.0)	39.4 (22.5)	37.5 (19.6)	37.3 (19.7)	36.6 (19.7)	38.3 (18.5)	33.1 (17.5)	44.3 (21.7)	–

Inter-subject means and standard deviations (S.D.) of similarity ratings were obtained using a visual analogue scale (0, dissimilar; 100, similar).

Table S3. Correlational matrix for perceptual RDVs.

	intensity	ple	unp	quality	similarity	trigeminal	pleasantness (during EEG)
intensity	–	–	–	–	–	–	–
ple	0.15 p = 0.07	–	–	–	–	–	–
unp	0.29 p < 0.05	0.43 p < 0.05	–	–	–	–	–
quality	0.28 p < 0.05	0.48 p < 0.05	0.64 p < 0.05	–	–	–	–
similarity	0.29 p < 0.05	0.20 p < 0.05	0.48 p < 0.05	0.50 p < 0.05	–	–	–
trigeminal	0.03 p = 2.24	0.04 p = 2.51	0.01 p = 2.61	-0.03 p = 2.30	0.01 p = 3.19	–	–
pleasantness (during EEG)	0.29 p < 0.05	0.45 p < 0.05	0.53 p < 0.05	0.52 p < 0.05	0.38 p < 0.05	0.01 p = 3.19	–

To examine the relationships among perceptual variables, we computed pairwise correlations between perceptual representational dissimilarity vectors (RDVs). For each subject and for each perceptual test, an RDV was constructed by computing the pairwise Euclidean distances of the corresponding perceptual scores for all the possible 45 odor pairs. Then, for each subject for each pair of perceptual RDVs, the Spearman's rank correlation was computed, and group-level statistical significance was examined using the Wilcoxon signed-rank test (one-sided, Bonferroni-Holm corrected). The resulting *p* values and inter-subject mean correlation coefficients are presented. Ple, unipolar pleasantness; unp, unipolar unpleasantness.

Table S4. number of valid trials for each subject and each odors

	Ally	Fru	Cit	Lin	Van	Ace	Hex	Pin	Cyc	4Pe
sub01	30	29	29	30	30	30	30	30	26	28
sub02	27	28	25	28	28	26	27	22	21	25
sub03	30	30	28	26	29	30	29	27	30	29
sub04	30	30	30	30	30	30	30	30	30	30
sub05	24	23	21	22	21	18	26	25	25	23
sub06	27	27	30	29	27	30	28	27	30	29
sub07	21	26	20	20	21	19	19	22	18	23
sub08	30	30	30	30	29	29	30	29	30	30
sub09	30	30	29	30	30	29	29	30	30	30
sub10	28	26	29	23	25	26	22	28	26	26
sub11	28	29	30	27	29	27	28	29	30	28
sub12	30	30	30	30	30	30	30	30	30	29
sub13	30	29	29	30	29	30	30	29	30	30
sub14	30	30	30	30	30	30	29	30	30	30
sub15	29	29	29	29	29	30	24	30	29	27
sub16	30	30	29	30	29	30	30	29	30	29
sub17	29	25	29	27	30	30	23	30	29	30
sub18	30	30	30	30	30	30	29	30	30	29
sub19	30	30	25	29	30	29	30	28	30	30
sub20	30	30	29	30	30	30	29	30	29	30
sub21	30	30	30	30	30	30	29	30	30	30
sub22	30	30	29	29	27	30	30	29	30	30

For decoding analyses, we used a nested cross-validation (CV) procedure, where an inner loop CV responsible for tuning the cost parameter was nested in an outer loop CV, which was responsible for model evaluation. For example, in the case of subject-wise, pairwise decoding for Ally vs Cys for subject 01, the number of trials used for each CV was as follows:

1. In each outer CV, 4 trials of Ally were randomly discarded to match the number of trials with that of Cys.
2. For both odors, 90% of the remaining 26 trials (=23 trials) were used as training data, and the other three trials were used as test data. If the resulting numbers were not integers, the number of trials for the training data was rounded down, and that of the test data was rounded up.
3. The training data were further split into 10 sets for the inner loop, 10-fold CV (23 trials divided by 10), which resulted in 2 or 3 trials for test data, and 21 or 22 training data.

SI references

1. R. C. Oldfield, The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* **9**, 97–113 (1971).
2. A. Dravnieks, *Atlas of Odor Character Profiles* (ASTM International, 1985) (July 24, 2021).
3. G. Kobal, Pain-related electrical potentials of the human nasal mucosa elicited by chemical stimulation. *Pain* **22**, 151–163 (1985).
4. N. Gotow, A. Hoshi, T. Kobayakawa, Expanded olfactometer for measuring reaction time to a target odor during background odor presentation. *Heliyon* **5**, e01254 (2019).
5. E. Becker, *et al.*, Olfactory event-related potentials in psychosis-prone subjects. *Int. J. Psychophysiol.* **15**, 51–58 (1993).
6. T. Hummel, G. Kobal, Differences in human evoked potentials related to olfactory or trigeminal chemosensory activation. *Electroencephalogr. Clin. Neurophysiol. Evoked Potentials* **84**, 84–89 (1992).
7. E. Iannilli, *et al.*, Source localization of event-related brain activity elicited by food and nonfood odors. *Neuroscience* **289**, 99–105 (2015).
8. R. Masago, Y. Shimomura, K. Iwanaga, T. Katsuura, The effects of hedonic properties of odors and attentional modulation on the olfactory event-related potentials. *J. Physiol. Anthropol. Appl. Human Sci.* **20**, 7–13 (2001).
9. G. Kobal, C. Hummel, Cerebral chemosensory evoked potentials elicited by chemical stimulation of the human olfactory and respiratory nasal mucosa. *Electroencephalogr. Clin. Neurophysiol. Evoked Potentials* **71**, 241–250 (1988).
10. T. Tateyama, T. Hummel, S. Roscher, H. Post, G. Kobal, Relation of olfactory event-related potentials to changes in stimulus concentration. *Electroencephalogr. Clin. Neurophysiol. - Evoked Potentials* **108**, 449–455 (1998).
11. G. Kobal, *Elektrophysiologische Untersuchungen des menschlichen Geruchssinns (Thieme Verlag, Stuttgart, 1981).*
12. K. Shibata, T. Watanabe, M. Kawato, Y. Sasaki, Differential Activation Patterns in the Same Brain Region Led to Opposite Emotional States. *PLoS Biol.* **14**, 1–27 (2016).
13. D. F. Salisbury, B. Rutherford, M. E. Shenton, R. W. McCarley, Button-pressing affects P300 amplitude and scalp topography. *Clin. Neurophysiol.* **112**, 1676–1684 (2001).
14. I. Croy, *et al.*, Human olfactory lateralization requires trigeminal activation. *Neuroimage* **98**, 289–295 (2014).
15. T. Hummel, T. Futschik, J. Frasnelli, K. B. Hüttenbrink, Effects of olfactory function, age, and gender on trigeminally mediated sensations: A study based on the lateralization of chemosensory stimuli. *Toxicol. Lett.* **140**–**141**, 273–280 (2003).
16. P. M. Wise, C. J. Wysocki, J. N. Lundström, Stimulus selection for intranasal sensory isolation: Eugenol is an irritant. *Chem. Senses* **37**, 509–514 (2012).
17. T. Hummel, Assessment of intranasal trigeminal function. *Int. J. Psychophysiol.* **36**, 147–155 (2000).
18. G. Kobal, S. Van Toller, T. Hummel, Is there directional smelling? *Experientia* **45**, 130–132 (1989).
19. A. Sorokowska, *et al.*, Odor lateralization and spatial localization: Null effects of blindness. *Attention, Perception, Psychophys.* **81**, 2078–2087 (2019).
20. A. Delorme, S. Makeig, EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).
21. A. Widmann, E. Schröger, B. Maess, Digital filter design for electrophysiological data - a practical approach. *J. Neurosci. Methods* **250**, 34–46 (2015).
22. N. Bigdely-Shamlo, T. Mullen, C. Kothe, K. M. Su, K. A. Robbins, The PREP pipeline: Standardized preprocessing for large-scale EEG analysis. *Front. Neuroinform.* **9**, 1–19 (2015).
23. A. J. Bell, T. J. Sejnowski, An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* **7**, 1129–1159 (1995).
24. I. Winkler, S. Haufe, M. Tangermann, Automatic Classification of Artifactual ICA-Components for Artifact Removal in EEG Signals. *Behav. Brain Funct.* **7**, 30 (2011).

25. M. M. Murray, D. Brunet, C. M. Michel, Topographic ERP analyses: A step-by-step tutorial review. *Brain Topogr.* **20**, 249–264 (2008).
26. B. M. Pause, K. Krauel, Chemosensory event-related potentials (CSERP) as a key to the psychology of odors. *Int. J. Psychophysiol.* **36**, 105–122 (2000).
27. M. Fuchs, J. Kastner, M. Wagner, S. Hawes, J. S. Ebersole, A standardized boundary element method volume conductor model integral equation using analytically integrated elements. *Clin. Neurophysiol.* **113**, 702–712 (2002).
28. J. Mazziotta, *et al.*, A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos. Trans. R. Soc. B Biol. Sci.* **356**, 1293–1322 (2001).
29. R. D. Pascual-Marqui, Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find. Exp. Clin. Pharmacol.* **24 Suppl D**, 5–12 (2002).
30. K. G. Mideksa, *et al.*, Comparison of imaging modalities and source-localization algorithms in locating the induced activity during deep brain stimulation of the STN. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS* **2016-Octob**, 105–108 (2016).
31. S.-H. Jeong, K.-Y. Jung, J.-M. Kim, J. S. Kim, Medial Temporal Activation in Mal de Debarquement Syndrome Revealed by Standardized Low-Resolution Brain Electromagnetic Tomography. *J. Clin. Neurol.* **8**, 238 (2012).
32. G. Petit, *et al.*, Reduced processing of alcohol cues predicts abstinence in recently detoxified alcoholic patients in a three-month follow up period: An ERP study. *Behav. Brain Res.* **282**, 84–94 (2015).
33. G. Zhou, G. Lane, S. L. Cooper, T. Kahnt, C. Zelano, Characterizing functional pathways of the human olfactory system. *Elife* **8**, 1–27 (2019).
34. P. A. Kragel, L. Koban, L. F. Barrett, T. D. Wager, Representation, Pattern Information, and Brain Signatures: From Neurons to Neuroimaging. *Neuron* **99**, 257–273 (2018).
35. Q. Wang, B. Cagna, T. Chaminade, S. Takerkart, Inter-subject pattern analysis: A straightforward and powerful scheme for group-level MVPA. *Neuroimage* **204**, 116205 (2020).
36. N. Kriegeskorte, M. Mur, P. Bandettini, Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, 1–28 (2008).
37. N. Kriegeskorte, R. A. Kievit, Representational geometry: Integrating cognition, computation, and the brain. *Trends Cogn. Sci.* **17**, 401–412 (2013).
38. H. Nili, *et al.*, A Toolbox for Representational Similarity Analysis. *PLoS Comput. Biol.* **10** (2014).
39. T. Okumura, H. Kumazaki, A. K. Singh, K. Touhara, M. Okamoto, Individuals with Autism Spectrum Disorder Show Altered Event-Related Potentials in the Late Stages of Olfactory Processing. *Chem. Senses* **45**, 45–58 (2020).
40. W. J. Evans, G. Kobal, T. S. Lorig, J. D. Prah, Suggestions for collection and reporting of chemosensory (olfactory) event-related potentials. *Chem. Senses* **18**, 751–756 (1993).