Leveraging Multi-Echo EPI to Enhance BOLD Sensitivity in Task-based Olfactory fMRI

Ludwig Sichen Zhao[*1,2], Clara U. Raithel[3,2], M. Dylan Tisdall[4], John A. Detre[2,4], and Jay A. Gottfried[*2,3]

[1]Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, United States
[2]Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States
[3]Department of Psychology, School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, United States
[4]Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

* Corresponding Author: Ludwig Sichen Zhao (ludwigz@seas.upenn.edu), Jay A. Gottfried (jaygottf@pennmedicine.upenn.edu)

Abstract

Functional magnetic resonance imaging (fMRI) using blood-oxygenation-level-dependent (BOLD) contrast relies on gradient echo echo-planar imaging (GE-EPI) to quantify dynamic susceptibility changes associated with the hemodynamic response to neural activity. However, acquiring BOLD fMRI in human olfactory regions is particularly challenging due to their proximity to the sinuses where large susceptibility gradients induce magnetic field distortions. BOLD fMRI of the human olfactory system is further complicated by respiratory artifacts that are highly correlated with event onsets in olfactory tasks.

Multi-echo EPI (ME-EPI) acquires gradient echo data at multiple echo times (TEs) during a single acquisition and can leverage signal evolution over the multiple echo times to enhance BOLD sensitivity and reduce artifactual signal contributions. In the current study, we developed a ME-EPI acquisition protocol for olfactory task-based fMRI and demonstrated significant improvement in BOLD signal sensitivity over conventional single-echo EPI (1E-EPI). The observed improvement arose from both an increase in BOLD signal changes through a T_2^*-weighted echo combination and a reduction in non-BOLD artifacts through the application of the Multi-Echo Independent Components Analysis (ME-ICA) denoising method. This study represents one of the first direct comparisons between 1E-EPI and ME-EPI in high-susceptibility regions and provides compelling evidence in favor of using ME-EPI for future task-based fMRI studies.

Keywords: multi-echo EPI, olfaction, ME-ICA, susceptibility artifacts, task-based fMRI

1 Introduction

Functional magnetic resonance imaging (fMRI) is one of the most common imaging techniques used to non-invasively study both localized neural activation and whole-brain functional connectivity in humans. Blood-oxygenation-level-dependent (BOLD) fMRI measures brain activity by detecting local magnetic susceptibility fluctuations caused by changes in deoxyhemoglobin concentration. Gradient echo echo-planar imaging (GE-EPI) is the dominant BOLD imaging strategy due to its high temporal signal-to-noise ratio and rapid acquisition capabilities (Budde et al., 2014). However, artifacts caused by the local inhomogeneous magnetic field surrounding high-susceptibility regions present a major technical challenge in utilizing GE-EPI to measure BOLD signals. Olfactory regions, including the orbitofrontal cortex (OFC), amygdala, piriform (olfactory) cortex, and entorhinal cortex are particularly affected by susceptibility-induced artifacts and signal drop-out due their location near the ethmoid sinuses (air-filled cavities in the bone surrounding the nose). These artifacts limit the applicability of GE-EPI-based BOLD fMRI in pursuit of understanding the human olfactory system (Catani et al., 2013). Moreover, respiratory artifacts, primarily caused by local field shifts during inhalation and exhalation that are highly correlated with event onsets in tasked-based olfactory-related fMRI, further compromise the sensitivity of olfactory fMRI (Glover & Law, 2001).

Several techniques have been proposed to overcome these challenges. Acquiring scans at a tilted angle allows for partial compensation of local susceptibility gradients (Deichmann et al., 2003; Weiskopf et al., 2006). Alternatively, employing pulse sequences other than GE-EPI, such as T_2-Prepared BOLD fMRI and balanced steady state free precession (bSSFP) (Parrish et al., 2008; Miao et al., 2021).

Recent work on multi-echo EPI (ME-EPI), which acquires multiple gradient echoes at each imaging timepoint, has shown potential to improve BOLD sensitivity and efficiency in resting-state fMRI (Gonzalez-Castillo et al., 2016; Lynch et al., 2020). In conventional single-echo EPI (1E-EPI), a single volume is acquired after each RF excitation at a specific echo time (TE), which determines both image signal intensity as well as BOLD contrast (Donahue et al., 2011; Gati et al., 1997; van der Zwaag et al., 2009). At a

shorter TE, the image is more proton-density weighted, has less BOLD contrast, and less severe signal dropout in regions with high local susceptibility gradients. Conversely, with a longer TE, the image becomes more T_2^*-weighted, has greater BOLD contrast, and more severe signal dropout. The BOLD contrast-to-noise ratio (CNR) is expected to vary with TE, peaking at the T_2^* of each voxel (Donahue et al., 2011; Fera et al., 2004; Gati et al., 1997; Graham et al., 1996; Triantafyllou et al., 2011; van der Zwaag et al., 2009). However, various brain regions exhibit distinct T_2^* values, resulting in varying optimal TEs, making it impossible to optimize TE across the whole brain in a single-echo acquisition.

Unlike 1E-EPI, ME-EPI acquires multiple images across a range of TEs after each RF excitation, resulting in enhanced functional contrast by combining the echo images with different weights (Poser et al., 2006; Posse et al., 1999). The commonly used T_2^*-weighted echo combination method assigns the heaviest weight to the echo that is presumed to have the greatest BOLD CNR, with weighting performed separately for each voxel. By applying a T_2^*-weighted echo combination with ME-EPI, BOLD signal can be maximized at the voxel level while maintaining the temporal resolution needed for fMRI. Moreover, by leveraging the expected exponential signal decay across TEs for true BOLD contrast, noise reduction techniques can be employed to reject artifactual signals and thereby improve the BOLD contrast-to-noise ratio (DuPre et al., 2021; Gonzalez-Castillo et al., 2016; Kundu et al., 2012a, 2013, 2017; Van et al., 2023).

In this study, we implemented and tested a ME-EPI acquisition protocol and analysis pipeline specifically designed for task-based olfactory-related paradigms by incorporating shorter TE values in an effort to improve sensitivity in ventral brain regions with high static susceptibility. Using this approach in comparison to a standard 1E-EPI acquisition, we demonstrate superior sensitivity to olfactory task activation resulting in a significant sample size reduction for detecting group task activation during a representative olfactory three-alternatives forced choice task (3AFC).

2 Method

2.1 Subjects

29 subjects (13 women; aged 22 – 30 years, mean: 26, SD: 3.15) provided informed consent as approved by the University of Pennsylvania Institutional Review Board (#827217). One subject was excluded from the study due to excessive motion (number of excluded volumes, as defined in Statistical Analysis, exceeded 2% of the number of total volumes) during fMRI scanning session, leaving us with a final sample of n = 28. All subjects reported being right-handed nonsmokers with no history of significant neurological/psychiatric disorder or olfactory dysfunction.

2.2 Odor Stimuli and Delivery

Stimuli consisted of two odorants: lemon oil extract (12.3% v/v, Sigma-Aldrich, MO, USA) and benzaldehyde (0.42% v/v, Sigma-Aldrich, MO, USA). Both odorants were diluted in mineral oil (Sigma Aldrich, MO, USA). Mineral oil alone was used as a control stimulus. The odor stimuli were delivered through a custom-built computer-controlled air-diluted 12-channel olfactometer with a flow rate of 4 L/min controlled using the PsychToolbox package (Brainard, 1997) and MATLAB (Version 2020b, The Mathworks Inc., Natick, MA, USA). The medical-grade air (Airgas, Radnor, PA, USA) was first routed through two computer-controlled 6-channel gradient valve manifolds (NResearch Inc., West Caldwell, NJ, USA), regulated by a mass flow controller (MFC) (Alicat Scientific Inc., Tucson, AZ, USA). Subsequently, it passed through the headspace above 2 mL of the mineral oil-diluted odorant before being delivered to the participant through a fluorinated ethylene-propylene (FEP) plastic tube (U.S. Plastic Corp., Lima, OH, USA) placed directly below the nose.

2.3 Behavioral Testing Session

On Day 0, subjects participated in a behavioral testing session during which odor stimuli were presented outside the MRI environment (Figure 1a). Participants were introduced to the odors and asked to label them as well as to rate the perceived odor intensity, via the general Labeled Magnitude Scale (gLMS) (Green et al., 1993, 1996). Valence (likeability) was also assessed via the Labeled Hedonic Scale (LHS) (L. V. Jones et al.,

1955; Lim et al., 2009; Schutz & Cardello, 2001; Stone et al., 2020)], as well as familiarity [Visual Analog Scale (VAS) (Dourado et al., 2021; Flint et al., 2000)]. Ratings were analyzed in MATLAB (Version 2020b, The Mathworks Inc., Natick, MA, USA), using three one-way ANOVAs to evaluate whether there were any differences among stimuli.

2.4 fMRI Scanning Session

2-5 days after the behavioral testing session, the same participants took part in an fMRI scanning session, completing six runs, each of which contained 24 trials. On each trial, participants sniffed one of the three possible odorants (lemon oil extract, benzaldehyde, or mineral oil) and were asked to identify the odor using a three-alternative forced choice task with the labels they provided earlier during the behavioral testing session. Following the 3AFC task, they were also asked to rate the intensity of the odor. Each odor was presented eight times per run in a random order. The experimental paradigm as well as the sequence of events within a given trial is shown in Figure 1.

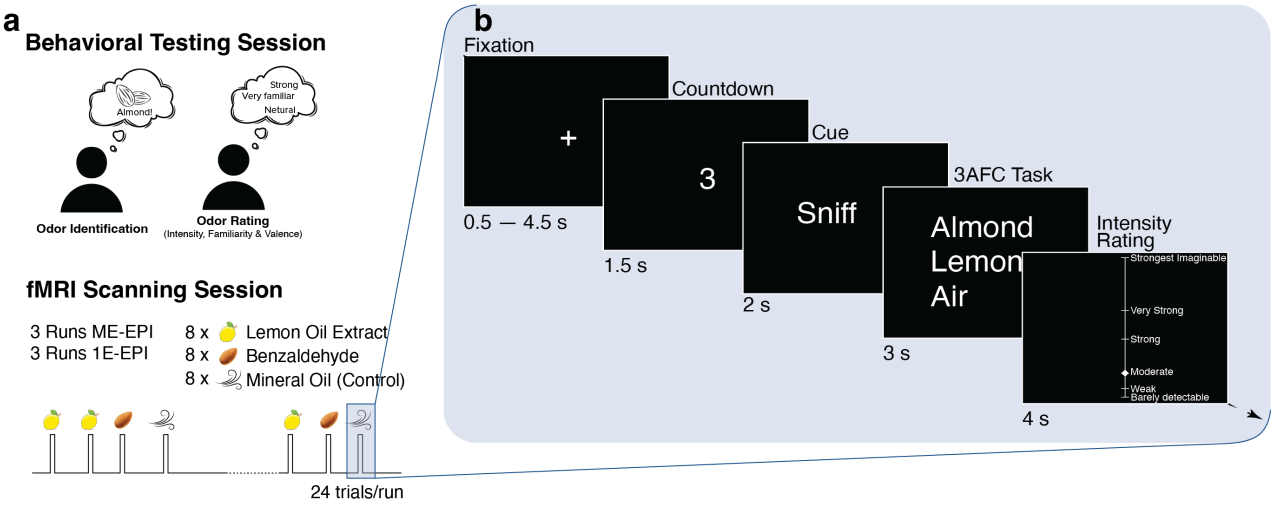


Figure 1. fMRI Experimental Paradigm. a) The experiment was conducted over the course of two *sessions*. On Day 0, the subjects participated in a series of behavioral tests where they were required to identify odors and rate their intensity, valence, and familiarity. Subsequently, an fMRI scanning session was scheduled for the subjects, taking place two to five days after their initial visit. This scanning session consisted of three runs of ME-EPI and three runs of 1E-EPI acquisition. Each run comprised 24 trials, with an equal distribution of three odor stimuli: lemon oil extract, benzaldehyde, and mineral oil. b) Timeline for a single trial. A trial started with a fixation cross presented for 0.5 – 4.5 s, followed by a 1.5 s

countdown. A cue to sniff was then displayed and the odor was delivered for 2 s. The short odor onset and a long intertrial interval (11.5 s on average) were chosen to minimize habituation. Participants were then asked to identify the odor and rate its intensity.

2.5 Structural and Functional MRI Data Acquisition

All magnetic resonance images were acquired using a 3 T scanner (MAGNETOM Prisma, Siemens Healthineers, Germany) using the vendor's 64-channel head and neck coil. T_1-weighted structural images were acquired using ME-MPRAGE (TE = 1.69/3.55/5.41/7.27 ms, TR = 2530 ms, TI = 1100 ms, FA = 8°, 1 mm isotropic).

1E-EPI and ME-EPI fMRI data were acquired using the sequence parameters outlined in Table 1. For both protocols, repetition time (TR), field of view (FOV), number of slices, resolution, flip angle (FA), echo spacing, and multiband factors are identical. To optimize the BOLD signals in OFC regions, the field of view was tilted approximately 25° to the anterior commissure-posterior commissure line (rostral > caudal) (Deichmann et al., 2003). The two sequence types were administered in an interleaved fashion in each participant, with the order being counterbalanced across participants. Of note, for 1E-EPI the echo time was set at 22 ms, which is approximately the same as the second echo time out of five echoes in the ME-EPI sequence (see Table 1). This TE aims to balance signal dropout and BOLD sensitivity for olfactory-related regions and is consistent with other olfactory-related fMRI studies (Bao et al., 2019; de Groot et al., 2021; Raithel et al., 2023; Shanahan et al., 2021; Zhou et al., 2019).

Additionally, a pair of spin-echo EPI (SE-EPI) images with two phase-encode blips are acquired for both fMRI acquisition protocols (1E-EPI and ME-EPI) separately with identical FOV, number of slices, resolution, FA, echo spacing, acceleration and readout bandwidth compared to their respective fMRI acquisition. These pairs of images are then used to estimate the inhomogeneities in the static magnetic field (B_0) during the preprocessing using a similar method described in (Andersson et al., 2003) and implemented in FSL (Smith et al., 2004) as part of fMRIPrep package (Esteban et al., 2019; Gorgolewski et al., 2011).

	1E- EPI	ME-EPI
TR (ms)	2041	2041
TE (ms)	22	10.60/22.92/35.24/ 47.56/59.88
FOV (mm x mm)	216 x 216	216 x 216
Slices	48	48
Resolution (mm x mm x mm)	2.4 x 2.4 x 2.4	2.4 x 2.4 x 2.4
FA (degree)	78	78
Echo Spacing (ms)	0.49	0.49
Bandwidth (Hz/Px)	2526	2646
Acceleration	Multiband 2x	Multiband 2x GRAPPA, 3x in-plane acceleration with 36 reference lines

Table 1. fMRI Sequence Parameters. TR: Repetition time. TE: Echo time. FOV: Field of view. FA: Flip angle. 1E-EPI: single-echo EPI. ME-EPI: multi-echo EPI.

2.6 fMRI Data Preprocessing

fMRI and structural MRI data were preprocessed in each subject's native space using fMRIPrep 21.0.0 based on Nipype 1.6.1 (Esteban et al., 2019; Gorgolewski et al., 2011). All images were then registered to ICBM 152 Nonlinear Asymmetrical template version 2009c (MNI-ICBM2009c) (2.4 mm isotropic resolution, consistent with fMRI acquisitions) before further processing (Fonov et al., 2009, 2011). Spatial smoothing was applied to all data using a Gaussian kernel of 8 mm full width at half maximum (FWHM). A detailed description of these preprocessing steps can be found in Supplementary Materials under Detailed fMRI Data Preprocessing section.

2.7 ME-EPI T_2^* Estimation, Echo Combination, and Multi-Echo Independent Components Analysis (ME-ICA)

Multi-echo EPI produces separate images for each echo time at each timepoint in the functional imaging run. At each timepoint, these images for the multiple echo times were

combined using the T_2^*-weighted combination method (ME-WC) (Posse et al., 1999). For each voxel, T_2^* was estimated by fitting the Eq. (1) using signals from all five echoes,

$$S = S_0 e^{-TE/T_2^*} \tag{1}$$

S_0 is the initial signal magnitude without any T_2^* decay and S is the signal magnitude acquired at $t = TE$.

The images from the five echoes are then combined voxel-wise using a weighted average (ME-WC) and the weights for each echo (W_{TE}) are calculated using Eq. (2) based on estimated T_2^*.

$$W_{TE} = \frac{TE \cdot exp(-TE/T_2^*)}{\sum_{n=1}^{5} TE_n \cdot exp(-TE_n/T_2^*)} \tag{2}$$

Independent components analysis (ICA) is a common strategy to decompose fMRI signals into distinct independent components, allowing the isolation of BOLD-related neuronal components from sources of noise (Beckmann, 2012; Caballero-Gaudes & Reynolds, 2017; Mckeown et al., 1998; McKeown et al., 2003). However, a persistent challenge lies in the manual classification of these independent components, which is labor-intensive and can yield inconsistent results (Kelly et al., 2010). In an attempt to tackle this problem, ME-ICA classifies each component as either a non-BOLD artifact, such as respiratory artifact, or a BOLD-like signal based on their signal characteristics across different echo times (Kundu et al., 2012b, 2013). Here, we specifically employed the TE-dependency analysis pipeline (Tedana 0.0.12) (DuPre et al., 2021) to preprocess multi-echo EPI data before registering them to the MNI template.

In the original TEDANA pipeline, a monoexponential model [Eq. (1)] was fit to the data at each voxel using nonlinear model fitting to estimate T_2^* and S_0 maps, using T_2^*/S_0 estimates from a log-linear fit as initial values. This approach generally performs well across most brain regions. However, in regions with significant signal dropout at longer echo times, particularly due to susceptibility artifacts in areas such as the orbitofrontal cortex and entorhinal cortex, this fitting method tends to overestimate T_2^*. Consequently, we replaced the TEDANA estimate of T2* with our own estimates, generated by fitting the following model voxel-wise:

$$S = \sqrt{(S_0 e^{-TE/T_2^*})^2 + x_0^2} \tag{3}$$

In this model, x_0 represents the thermal noise floor, which does not decay over TE and whose contribution to the overall signal increases in later echoes as the signal diminishes in high-susceptibility regions. This model was fit to the data by minimizing the 2-norm of the weighted residuals [Eq. (4)] via the Trust-Region-Reflective algorithm (Coleman & Li, 1996).

$$r_i = \frac{S_i}{\sum_{j=1}^{n} S_j}\left(\sqrt{(S_0 e^{-TE_i/T_2^*})^2 + x_0^2} - S_i\right) \tag{4}$$

The weighted residuals are calculated as the difference between the estimated signals $\widehat{S_i}$ and the observed signals S_i, weighted by the observed signals S_i across all n = 1000 observations (200 volumes/run, 5 echoes/volume).

2.8 Estimating Motion for Different Odor Stimuli

The framewise displacement (*FD*), calculated during fMRI preprocessing using the method described in (Power et al., 2012), and the onset of odor stimulus for each condition were extracted for each run. The coefficient of determination (R^2) was then estimated by calculating the Pearson correlation coefficient between the *FD* and the odor onset for each stimulus condition convolved with canonical hemodynamic response function (HRF) for each subject (Pearson, 1895; Siegel, 1956). To compare whether there is a group-level difference in movement parameters for different odor stimuli, two-sided paired *t*-tests were performed on the R^2, across all three stimulus

229 conditions ([Lemon vs. Control], [Benzaldehyde vs. Control], and [Lemon vs.
230 Benzaldehyde]) for both 1E-EPI and ME-EPI acquisitions.

231 *2.9 Physiological Data Analysis*

232 The respiratory trace was recorded with a respiratory transducer using a pressure
233 sensor (BIOPAC System, CA, USA) affixed around the subject's chest, and then
234 preprocessed before being added as a nuisance regressor: the trace was smoothed
235 using a moving average with a 250 ms window, high-pass filtered at 0.05 Hz with an
236 infinite impulse response (IIR) filter, detrended, and standardized such that the mean of
237 the signals is 0 and the standard deviation is 1 (Liu et al., 2024; Raithel et al., 2023).
238 The preprocessed trace was also downsampled to 1/2.4 Hz to match the *TR* of the fMRI
239 acquisitions and included as a nuisance regressor. Two additional nuisance regressors
240 were introduced: the trial-by-trial sniff volume was defined as the amplitude difference
241 between the peak and the trough within the cued sniff, and the sniff duration was
242 defined as the time difference between the peak and the trough.

243 *2.10 Statistical Analysis*

244 All fMRI data were analyzed in SPM12 using a general linear model (GLM) to estimate
245 the main effects of the odor stimulation. Each experimental condition (control, lemon oil,
246 and benzaldehyde) was modeled using an independent regressor for each run. This
247 results in a 2 (multi-echo and single-echo) by 3 (control, lemon oil, and benzaldehyde)
248 factorial design. The model included the following nuisance regressors: trial-by-trial sniff
249 volume and sniff duration, both convolved with the hemodynamic response function
250 (HRF), as well as the down-sampled breathing trace and 24 movement parameters (six
251 movement parameters derived from spatial realignment, their squares, derivatives, and
252 squares of derivatives). Additional nuisance regressors were introduced to exclude
253 volumes with excessive motion, defined as those with at least one of six movement
254 parameters derived from spatial realignment deviating by more than 6 standard
255 deviations from the mean. In all GLMs, data were high-pass filtered at 1/128 Hz; the
256 temporal autocorrelation was estimated and removed as an autoregressive AR(1)
257 process globally (Friston et al., 2002).

Contrasts of interest were defined as [Odor > Control], [Lemon > Control] and [Benzaldehyde > Control]. Whole brain analysis was performed to validate that odor-evoked activations were present in the data acquired through both protocols (1E-EPI and ME-EPI). Small volume corrections (SVC) and ROI analysis were separately performed on OFC, amygdala, piriform cortex and entorhinal cortex, as suggested by previous studies highlighting the involvement of these regions in human olfactory processing (Anderson et al., 2003). Amygdala and piriform cortex were defined anatomically using a human brain atlas (Mai et al., 2016). Delineation of OFC and entorhinal cortex were based on the MNI-ICBM2009c atlas (Manera et al., 2020).

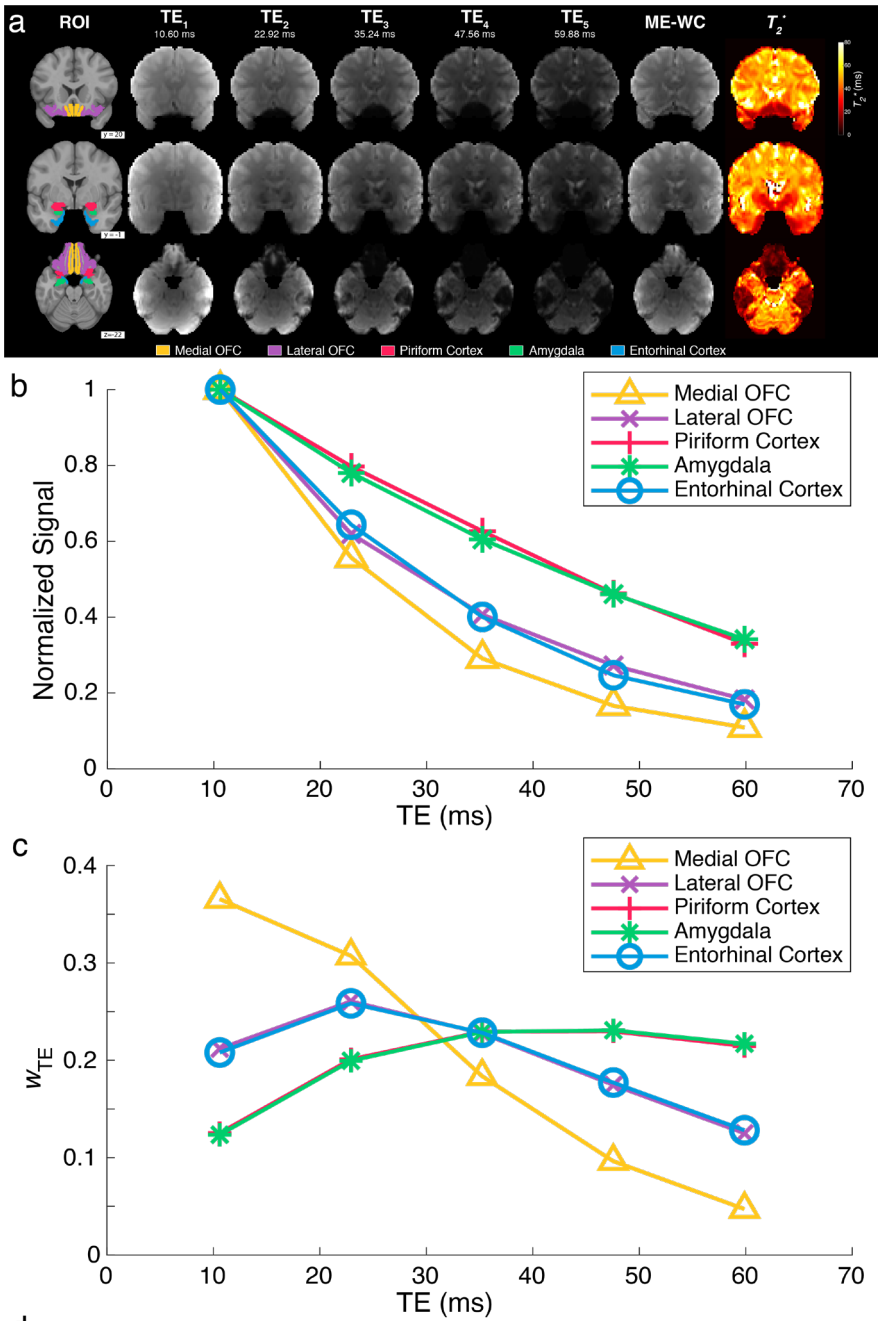
2.11 Post-hoc Power Analysis

We used our ROI analysis results to estimate the sample sizes required to detect significant olfactory-related group fMRI activation in subsequent studies. A priori power analyses were performed with $\alpha = 0.05$ and $\beta = 0.80$ using G*Power (Faul et al., 2009) for both ME-EPI and 1E-EPI data in the same ROIs described in the previous section.

3 Results

3.1 Echo-combined Images Reduce Signal Dropout

We first examined the estimated T_2^* values for each olfactory-related region. As expected, signal dropouts are reduced in shorter echo times (Figure 2a, TE_1 and TE_2), while T_2^*-weighted contrast is greater in later echo times (Figure 2a, TE_3 – TE_5). In particular, the medial and lateral OFC along with the entorhinal cortex, exhibited notably smaller T_2^* values due to susceptibility artifacts (Figure 2a & 2b), and in turn the echo-combined image more heavily weights the shorter echoes (Figure 2c). Using these weights, the combined image (Figure 2a, ME-WC) effectively mitigated signal dropout while retaining notable T_2^*-weighted contrast.



d

	Medial OFC	Lateral OFC	Piriform Cortex	Amygdala	Entorhinal Cortex
T_2^* (ms)	13.04 ± 0.15	21.81 ± 0.19	41.29 ± 1.11	42.18 ± 1.03	22.24 ± 0.30

Figure 2. Multi-Echo fMRI Reduces Signal Dropout and Improves BOLD Contrast through T_2^*-Weighted Combination in a Single Subject. a) Two representative coronal slices containing the regions of interest are shown with a variety of contrasts: ROIs overlayed on the T_1-weighted image (ROI); estimated T_2^* maps; all five individual echo images (TE_1 – TE_5) from the ME-EPI acquisition; and the T_2^*-weighted combined images (ME-WC). b) Plot of signal vs echo time for each anatomical ROI within a single

subject. Within each ROI, mean signals for each individual echo were normalized by dividing by the mean signal from the first echo. c) Within-ROI mean of the voxel-wise weights used to compute the T_2^*-weighted combined images for the same subject. d) Estimated mean T_2^* ± 95% confidence interval for each ROI.

3.2 BOLD Activation in Diverse Olfactory Brain Regions During Odor Stimulation

We next investigated the regions that were significantly activated when the odors were delivered under ME-EPI acquisition with the ME-ICA denoising method [Odor > Control] (Figure 3a) (DuPre et al., 2021). Consistent with previous literature, we observed significant activation in the piriform cortex, amygdala, entorhinal cortex, insula, and OFC bilaterally ($p < 0.001$, uncorrected) (Anderson et al., 2003; Dikecligil & Gottfried, 2024; Gottfried et al., 2003; Howard & Gottfried, 2014; Patin & Pause, 2015; Plailly et al., 2008). We next examined the BOLD effects of individual odor stimulus, namely [Benzaldehyde > Control] and [Lemon > Control]. Despite statistical tests indicating no significant differences in perceived intensity ($p = 0.38$), valence ($p = 0.45$), and familiarity ($p = 0.11$) between benzaldehyde and lemon oil and no significant differences in motion among three experimental conditions ([Lemon vs. Control] $p = 0.60$, [Benzaldehyde vs. Control] $p = 0.33$, and [Lemon vs. Benzaldehyde] $p = 0.63$), our analysis revealed that the main effect [Odor > Control] was driven mainly by the contrast [Lemon > Control]. In comparison, the contrast [Benzaldehyde > Control] produced significant levels of activation only in the bilateral insular region (Figure 3a).

The difference in BOLD response to these stimuli was not anticipated though not wholly unprecedented, as previous research has demonstrated that different odors can elicit distinct responses within olfactory brain regions even in the absence of differences in low-level stimulus features (Fournel et al., 2016; Howard & Gottfried, 2014; Savic, 2002; Savic et al., 2002). In the single-echo fMRI, a similar trend is observed, with more active regions being detected with lemon oil ($p < 0.05$, uncorrected, Figure S2). To our knowledge, there is no literature in humans that directly elicits piriform cortex BOLD activation with a similar sample size for benzaldehyde. However, rodent studies have reported that benzaldehyde demonstrates BOLD activation in the amygdala, as well as the piriform and entorhinal cortices (Kulkarni et al., 2012). Accordingly, to compare and observe the differences between acquisition methods within olfactory regions, we will

focus on the activation under lemon oil condition [Lemon > Control] for our further analyses.

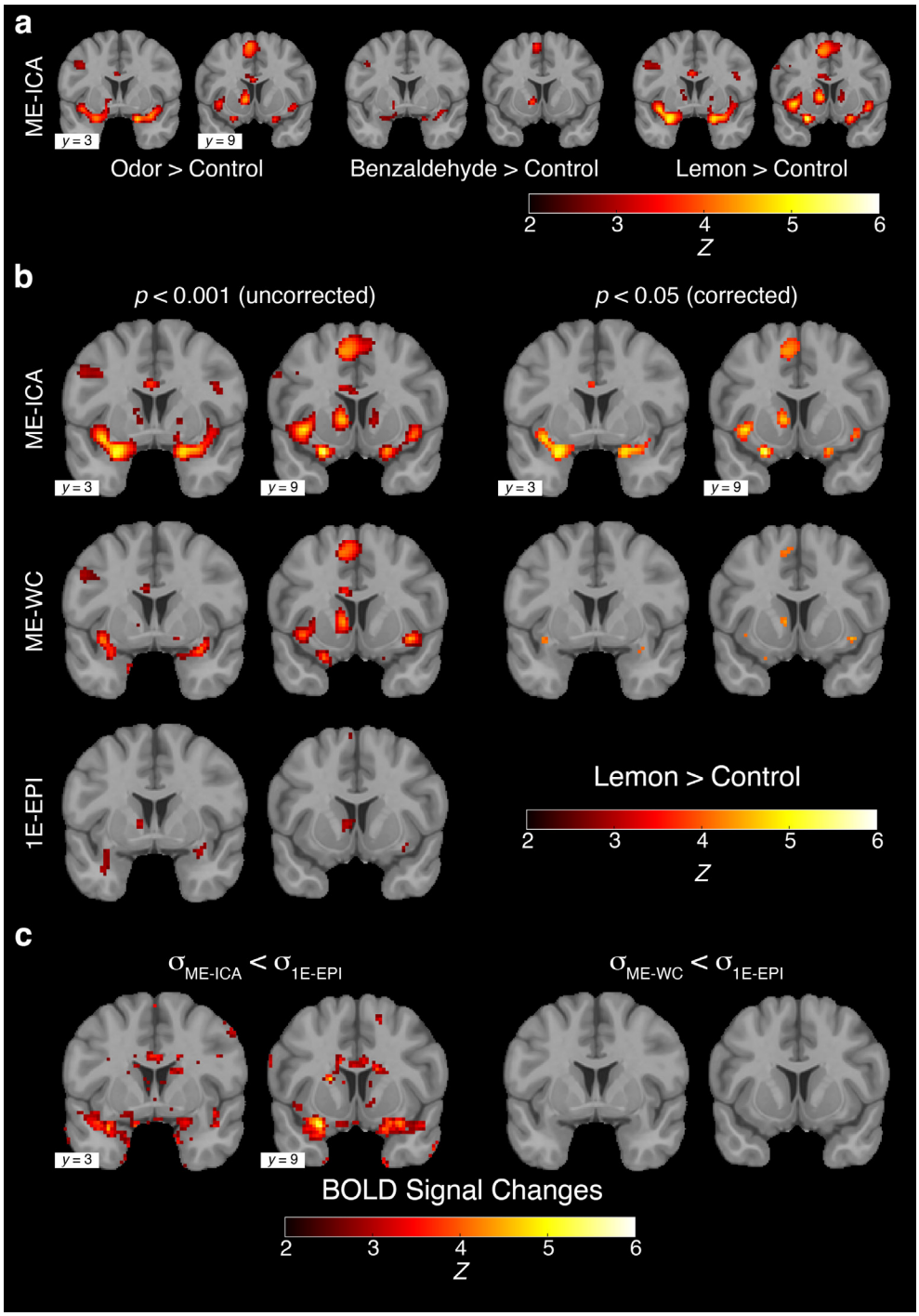


Figure 3. Activation Maps from Olfactory-Related Task for Both 1E-EPI and ME-EPI. a) Two representative coronal slices of activation maps for all three contrasts ([Odor > Control], [Benzaldehyde > Control], [Lemon > Control]) using ME-EPI acquisition with ME-ICA denoising methods ($p < 0.05$, FDR

corrected). b) Activation maps for [Lemon > Control] contrast for ME-EPI acquisition with ME-ICA denoising methods (ME-ICA), and without ME-ICA denoising methods (ME-WC) as well as single-echo fMRI acquisition (1E-EPI). The left two columns show the maps without multiple comparison correction ($p < 0.001$) while the right two columns show the maps corrected for multiple comparison ($p < 0.05$, FDR corrected). No activation survived correction for multiple comparisons for 1E-EPI (not shown). c) *Z*-score statistical maps yielded from Levene's test to examine whether BOLD signal variance was lower in ME-WC [$\sigma_{ME\text{-}WC} < \sigma_{1E\text{-}EPI}$] and ME-EPI [$\sigma_{ME\text{-}ICA} < \sigma_{1E\text{-}EPI}$] compared to 1E-EPI ($p < 0.05$, FDR corrected).

3.3 ME-EPI Improves fMRI Signal Detection by Both Increasing BOLD Signal Sensitivity and Specificity

Having validated that our ME-EPI acquisition protocol is suitable to capture odor-evoked BOLD responses, we next compared BOLD activation among different acquisition methods (1E-EPI vs. ME-EPI) and analysis pipelines (ME-WC vs. ME-ICA). A whole-brain analysis using either whole-brain multiple comparisons correction with false discovery rate (FDR) (Benjamini & Hochberg, 1995; Benjamini & Yekutieli, 2001; Yekutieli & Benjamini, 1999) or small volume correction (SVC) did not yield any supra-threshold voxels for 1E-EPI. However, the same whole-brain analysis using the T_2^*-weighted combined ME-EPI (ME-WC) highlighted activation in all ROIs (piriform cortex, amygdala, OFC, and entorhinal cortex) (Figure 3b). Moreover, by using ME-ICA denoising methods (DuPre et al., 2021), we observed a further increase in both the number of activation voxels and peak *Z* values in all ROIs (Table 2).

	Piriform Cortex		**Amygdala**		**Entorhinal Cortex**		**OFC**	
	# Voxel	**Peak *Z***	**# Voxel**	**Peak *Z***	**# Voxel**	**Peak *Z***	**# Voxel**	**Peak *Z***
ME-ICA	158	5.43	6	4.21	11	4.78	27	5.43
ME-WC	4	4.32	2	3.63	1	3.82	20	4.62
1E-EPI	0	N/A	0	N/A	0	N/A	0	N/A

Table 2. Number of Voxels and Peak Z-Score Activated by Lemon Oil for each ROI ($p < 0.05$, SVC).

We next examined potential reasons for the observed differences. Specifically, we hypothesized the improvement of the BOLD sensitivity in the case of the ME-EPI data was due to larger BOLD signal changes (i.e., greater sensitivity), as we performed a locally weighted combination of all echoes for each individual voxel instead of using one single TE across the brain. BOLD signal changes could be calculated by contrast

351 estimates from the GLM. Therefore, we performed a paired *t*-test on contrast estimates
352 [Lemon > Control] between 1E- and ME-EPI with ME-ICA [ME-ICA > 1E-EPI] among all
353 ROIs. Only piriform cortex (p = 0.023) and lateral OFC (p = 0.006) showed greater
354 estimates for the contrast [Lemon > Control] for ME-ICA compared to 1E-EPI.

355 An alternative explanation for the observed increase in BOLD detection for ME-EPI
356 versus 1E-EPI, may lie in the reduction of the background noise due to non-BOLD
357 signals (i.e., greater specificity). To test the hypothesis that ME-EPI signals have lower
358 variance, we performed a Levene's test (Levene, 1960). This test is commonly
359 employed to assess whether two or more groups have equal variances in a dataset
360 (Chiarello et al., 2009; Farinha et al., 2022; Im et al., 2010; T. B. Jones et al., 2010;
361 Swarup et al., 2013). When using a whole-brain analysis to compare the signal variance
362 between ME-ICA and 1E-EPI, we observed suprathreshold clusters (p < 0.05, FDR
363 corrected) overlapping with the voxels activated for the [Lemon > Control] contrast (p <
364 0.05, FDR corrected) from ME-ICA (Figure 3c). We then generated a mask using those
365 activated voxels and found that 18.89% of the masked voxels reached the significance
366 threshold under Levene's test. This suggests that ME-ICA decreases the signal
367 variance, thereby reducing non-BOLD noise and increasing BOLD specificity. We did
368 not observe any suprathreshold clusters when comparing ME-WC and 1E-EPI.

369 Taken together, these tests suggest that our observed increase in BOLD detection is
370 not only due to increased sensitivity BOLD signals, but also the reduction of non-BOLD
371 artifacts enhancing BOLD specificity.

372 *3.4 ME-EPI Improves Statistical Power and Reduces Sample Sizes*

373 So far, we have established that ME-EPI improved BOLD signal sensitivity by both
374 enhancing BOLD signal and reducing non-BOLD artifacts. An outstanding question is to
375 what degree BOLD sensitivity is improved when using our suggested method in the
376 context of task-based olfactory fMRI. To quantify the change in BOLD sensitivity, we
377 examined the contrast [Lemon > Control] in the brain regions central to olfaction,
378 namely the piriform cortex, amygdala, lateral and medial OFC, and entorhinal cortex for
379 both acquisition methods separately and then compared the statistical outcomes.
380 Whereas activation in medial OFC did not reach statistical significance using data

resulting from either acquisition method, all remaining regions show much lower *p*-values for ME-ICA compared to 1E-EPI (Table 3). Furthermore, to determine the minimum sample size required to observe significance in each region, post-hoc power analyses were carried out (α = 0.05 and β = 0.80) to estimate the effect of each acquisition method on the sample sizes needed to detect effects in similar olfactory task fMRI studies. ME-ICA led to a substantial reduction in the sample size, ranging from by half to one-fourth, compared to 1E-EPI, depending on the specific regions (Figure 4).

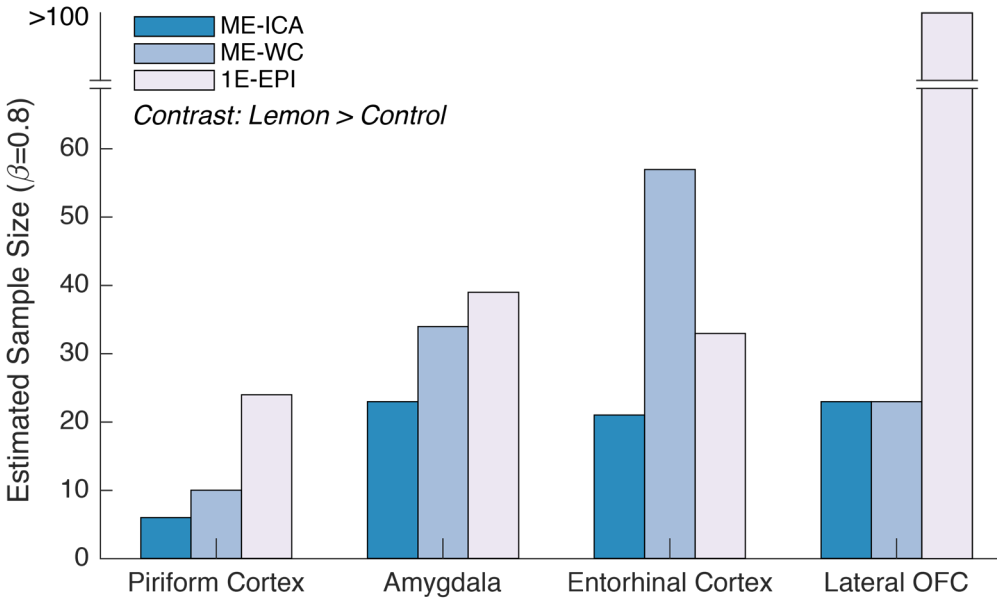


Figure 4. Estimated Sample Size for ROI Analyses Based on Post-hoc Power Analyses (α = 0.05 and β = 0.80) in all Olfactory-Related ROIs under Contrast [Lemon > Control] for Single-echo fMRI (1E-EPI), ME-EPI BOLD Acquisition with ME-ICA Denoising Methods (ME-ICA), and without ME-ICA Denoising Methods (ME-WC). The estimated sample sizes for medial OFC for 1E-EPI, ME-ICA and ME-WC were not shown since they were all above 100.

	Piriform Cortex	Amygdala	Entorhinal Cortex	Lateral OFC	Medial OFC
ME-ICA	< 0.0001***	0.0052**	0.0039**	0.0058**	0.7663
ME-WC	0.0001***	0.0173*	0.0505	0.0054**	0.5607
1E-EPI	0.0066**	0.0243*	0.0157*	0.5429	0.6687

Table 3. Statistical Significance of ROI Analyses for Each Olfactory-related ROI under the Contrast [Lemon > Control] for 1E-EPI, ME-WC and ME-ICA.

4 Discussion

Susceptibility artifacts pose a significant challenge when using GE-EPI for fMRI studies. The primary olfactory cortex not only suffers from some of the most severe static susceptibility artifacts within the human brain but is also affected by respiration-related artifacts. Here, we have developed and validated a ME-EPI protocol and analysis pipeline specifically tailored for olfactory task-based fMRI. In the context of a basic olfactory 3AFC task, we found that ME-EPI, alongside ME-ICA denoising, systemically improves BOLD sensitivity and reduces the required sample sizes needed to infer a significant effect at a group level compared to conventional 1E-EPI. This is achieved by both increasing BOLD signal sensitivity and reducing the contributions of non-BOLD signal changes using the additional denoising features ME-EPI provided.

The echo times used for ME-BOLD provided a range of contrasts across different TEs for each ROI of interest (Figure 2a). When evaluating the T_2^*-derived weights calculated for the echo-combined image, we confirmed that high-susceptibility regions with short T_2^* values, namely, lateral and medial OFC, and entorhinal cortex, had higher weights assigned to the first two echoes. In contrast, the remaining ROIs had higher weights assigned to the remaining (and later) three echoes. As a result, the T_2^*-weighted combined (ME-WC) images were able to produce more uniform T_2^*-weighted contrast across the whole brain (Figure 2a, ME-WC).

When investigating the region activated by odor onset using ME-ICA, we detected activation in all major primary olfactory cortex regions (piriform cortex, amygdala, entorhinal cortex, and OFC bilaterally) in addition to the insular region (Figure 3a). These findings are consistent with previous literature (Anderson et al., 2003; Dikecligil & Gottfried, 2024; Gottfried et al., 2003; Howard & Gottfried, 2014; Patin & Pause, 2015; Plailly et al., 2008). Further analysis revealed that the activation was primarily driven by the lemon extract stimulus in comparison to benzaldehyde. Previous studies have demonstrated that different odor stimuli can activate distinct primary olfactory regions to different degrees (Fournel et al., 2016; Howard & Gottfried, 2014; Savic, 2002; Savic et al., 2002). In this study we found that that benzaldehyde activates the olfactory system

425 to a lesser extent than lemon oil, such that the observed activation did not reach
426 significance after correcting for multiple comparisons in the whole brain analysis.

427 When comparing the results of 1E-EPI and ME-EPI, we observed significantly increased
428 sensitivity in both whole brain analysis and ROI analysis, which are two commonly used
429 univariate fMRI analysis methods. In the case of whole brain analysis (Figure 3b), we
430 found that ME-EPI exhibited higher peak *Z* values and a greater number of activated
431 voxels compared to 1E-EPI, as shown in Table 2. It is worth noting that even without
432 using the ME-ICA denoising method, ME-EPI still outperformed 1E-EPI by performing a
433 T_2^*-weighted combination of data from all echoes. This improvement was expected, as
434 the weighting of data from each echo time was designed to enhance the sensitivity to
435 BOLD contrast (Posse et al., 1999).

436 For ROI analysis and sequential power analysis, which were used to estimate minimal
437 sample sizes, ME-EPI with the ICA denoising method demonstrated superiority over 1E-
438 EPI, as shown in Table 3 and Figure 4. We attribute these improvements to two factors:
439 the increase of BOLD signal intensity through T_2^*-weighted echo combination and the
440 reduction of noise by leveraging signal evaluation across the echo times. To support
441 these hypotheses, we conducted a series of statistical tests (Figure 3c) to show the
442 mean signals are significantly higher and the variance is significantly lower in ME-ICA
443 acquisition compared to 1E-EPI.

444 While both ME-EPI and ME-WC overwhelmingly increased the statistical power for ROI
445 analysis, we observed an unexpected decrease in statistical power for ME-WC in the
446 entorhinal cortex. Previous studies have demonstrated that local B_0 field shifts affect T_2^*
447 values near the ethmoid sinuses during respiration, resulting in fluctuations of T_2^* values
448 across respiratory cycles (Barry & Menon, 2005; Raj et al., 2000, 2001; Van de
449 Moortele et al., 2002). The T_2^*-weighted combination method used in our analyses
450 adopted a fixed weighting approach throughout a single scan, potentially failing to
451 account for these dynamic T_2^* fluctuations. However, our additional analyses showed
452 that, although statistically significant in most ROIs, T_2^* values exhibit only subtle
453 differences between inhalation and exhalation phases (Table S1). This, in return,

demonstrates the robustness of the echo combination method against respiratory phase variations.

We also evaluated whether nonexponential signal decay contributed to suboptimal echo combination weightings, leading to the unexpected result. As shown in Figure S2, signals from most regions fit well with exponential decay, but regions with high susceptibility, such as the entorhinal cortex, medial, and lateral OFC, exhibited more pronounced nonexponential decay. Yet, this observation does not fully explain the unexpected decreases observed exclusively in the entorhinal cortex.

Furthermore, we examined the effects of ME-WC and ME-ICA on the distribution of statistical significance within each ROI (Figure S1). We have shown that ME-WC increases sensitivity to both positively correlated BOLD activation and anti-correlated signals. Moreover, ME-WC and ME-ICA both increases the peak *Z*-scores and the number of statistically significant voxels, consistent with voxel-wise whole-brain analysis (Table 2 and Figure 3b).

Additionally, we investigated the effect of the number of echoes on the main effect, considering whether later echoes contribute significantly. Using the same analysis pipeline, we compared whole-brain analysis results from data including the first 3, 4, and all 5 echoes (Figure S4). As the number of echoes increased, the size of the clusters also increased.

The current analysis pipeline utilized a specific method of combining multiple echo image series into a single image series: T_2^*-weighted combination, also known as "optimal combination", is used in most ME-EPI studies as it is designed to maximize BOLD-contrast sensitivity (Baracchini et al., 2021; Gonzalez-Castillo et al., 2016; Kundu et al., 2017; Lynch et al., 2020; Roth et al., 2020). However, an interesting finding emerged when we processed the ME-EPI data separately and treated each individual echo series as a 1E-EPI dataset. Surprisingly, no supra-threshold activation was detected at the first echo ($p < 0.05$, uncorrected), while the majority of activations were observed at the later echoes ($p < 0.05$, uncorrected) (Figure S5). When examining the *Z* values for the ROI analyses across all 5 individual echoes, we also observe a similar trend (Table S2). However, regions with short T_2^* (such as the lateral and medial OFC

484 and entorhinal cortex in this study) predominantly assigned higher weights to the first
485 two echoes.

486 Interestingly, in the process of understanding the T_2^*-weighted combination method
487 implemented in the TEDANA package, we observed that directly estimating T_2^* by fitting
488 the Eq. (1) tends to overestimate T_2^* in regions with high susceptibility (Table S3). As
489 described in the Method, we proposed using Eq. (3) as the objective function to
490 estimate T_2^* for each voxel. While this approach improved some estimations, work is
491 needed to improve the accuracy of these estimations in the future.

492 There have also been recent advances in denoising methods that leverage phase
493 information, such as Noise Reduction with Distribution Corrected (NORDIC) PCA
494 (Dowdle et al., 2023; Moeller et al., 2021; Vizioli et al., 2021). These methods have
495 demonstrated the ability to further enhance BOLD sensitivity in resting-state fMRI.
496 However, our study did not acquire phase data, which prevented us from testing these
497 advanced denoising methods in our data.

498 In this study, we utilized a limited set of two odor stimuli, along with a control stimulus.
499 Although this approach was adequate to compare the ME-EPI and 1E-EPI acquisition
500 methods, future studies will aim to incorporate a wider range of odor stimuli. By
501 including multiple odor stimuli, we can expand our analysis to encompass multivariate
502 analysis, such as Multivariate Pattern Analysis (MVPA), which has been commonly
503 utilized in other olfaction task-based fMRI studies (Donoshita et al., 2021; Howard et al.,
504 2009; Howard & Gottfried, 2014; Perszyk et al., 2023), though we expect also to see an
505 improvement in multivariate analysis using ME-ICA.

506 In conclusion, our study demonstrated the utility of a ME-EPI protocol combined with a
507 pre-configured analysis pipeline for fMRI in olfactory cortex. This approach enhanced
508 sensitivity to task-dependent BOLD contrast changes in olfactory regions, and post-hoc
509 power analyses indicate that group effects can be detected with a reduced sample size
510 compared to standard 1E-EPI. Although our protocol and pipeline were specifically
511 designed for task-based fMRI in the presence of large static susceptibility, a
512 longstanding challenge in the field of fMRI, they showcase the superior denoising

capabilities of ME-EPI for task-based fMRI more generally. These findings strongly support the adoption of ME-EPI over 1E-EPI for future task-based fMRI studies.

Data and Code Availability

The unthresholded *Z*-maps showing the whole-brain results for univariate analyses were made available on Neurovault (https://neurovault.org/collections/KGZMOMWR/). The modified TEDANA codes with updated T2* estimation methods described in Section 2.7 are available on GitHub (https://github.com/luxwig/tedana/tree/altT2s).

Author Contributions

Ludwig Sichen Zhao: Conceptualization, Methodology, Software, Formal Analysis, Investigation, Data Curation, Visualization, Writing – Original Draft, Project administration. **Clara U. Raithel**: Conceptualization, Methodology, Investigation, Writing – Review & Editing. **M. Dylan Tisdall**: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Resources. **John A. Detre**: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Resources. **Jay A. Gottfried**: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Resources, Funding acquisition

Funding

This work was supported by the National Institutes of Health award R01DC019405 [awarded to JAG].

Declaration of Competing Interests

The authors have declared that no competing interests exist.

References

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., Gabrieli, J. D. E., & Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, *6* (2), Article 2. https://doi.org/10.1038/nn1001

Andersson, J. L. R., Skare, S., & Ashburner, J. (2003). How to correct susceptibility distortions in spin-echo echo-planar images: Application to diffusion tensor imaging. *NeuroImage*, *20* (2), 870–888. https://doi.org/10.1016/S1053-8119(03)00336-7

Bao, X., Gjorgieva, E., Shanahan, L. K., Howard, J. D., Kahnt, T., & Gottfried, J. A. (2019). Grid-like Neural Representations Support Olfactory Navigation of a Two-Dimensional Odor Space. *Neuron*, *102* (5), 1066-1075.e5. https://doi.org/10.1016/j.neuron.2019.03.034

Baracchini, G., Mišić, B., Setton, R., Mwilambwe-Tshilobo, L., Girn, M., Nomi, J. S., Uddin, L. Q., Turner, G. R., & Spreng, R. N. (2021). Inter-regional BOLD signal variability is an organizational feature of functional brain networks. *NeuroImage*, *237*, 118149. https://doi.org/10.1016/j.neuroimage.2021.118149

Barry, R. L., & Menon, R. S. (2005). Modeling and suppression of respiration-related physiological noise in echo-planar functional magnetic resonance imaging using global and one-dimensional navigator echo correction. *Magnetic Resonance in Medicine*, *54* (2), 411–418. https://doi.org/10.1002/mrm.20591

Beckmann, C. F. (2012). Modelling with independent components. *NeuroImage*, *62* (2), 891–901. https://doi.org/10.1016/j.neuroimage.2012.02.020

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57* (1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, *29* (4), 1165–1188. https://doi.org/10.1214/aos/1013699998

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10* (4), 433–436. https://doi.org/10.1163/156856897X00357

Budde, J., Shajan, G., Zaitsev, M., Scheffler, K., & Pohmann, R. (2014). Functional MRI in human subjects with gradient-echo and spin-echo EPI at 9.4 T. *Magnetic Resonance in Medicine*, *71* (1), 209–218. https://doi.org/10.1002/mrm.24656

Caballero-Gaudes, C., & Reynolds, R. C. (2017). Methods for cleaning the BOLD fMRI signal. *NeuroImage*, *154*, 128–149. https://doi.org/10.1016/j.neuroimage.2016.12.018

Catani, M., Dell'Acqua, F., & Thiebaut de Schotten, M. (2013). A revised limbic system model for memory, emotion and behaviour. *Neuroscience & Biobehavioral Reviews*, *37* (8), 1724–1737. https://doi.org/10.1016/j.neubiorev.2013.07.001

Chiarello, C., Welcome, S. E., Halderman, L. K., Towler, S., Julagay, J., Otto, R., & Leonard, C. M. (2009). A large-scale investigation of lateralization in cortical anatomy and word reading: Are there sex differences? *Neuropsychology*, *23* (2), 210–222. https://doi.org/10.1037/a0014265

Coleman, T. F., & Li, Y. (1996). An Interior Trust Region Approach for Nonlinear Minimization Subject to Bounds. *SIAM Journal on Optimization*, *6* (2), 418–445. https://doi.org/10.1137/0806023

de Groot, J. H. B., Kirk, P. A., & Gottfried, J. A. (2021). Titrating the Smell of Fear: Initial Evidence for Dose-Invariant Behavioral, Physiological, and Neural Responses. *Psychological Science*, *32* (4), 558–572. https://doi.org/10.1177/0956797620970548

Deichmann, R., Gottfried, J. A., Hutton, C., & Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage*, *19* (2), Article 2. https://doi.org/10.1016/S1053-8119(03)00073-9

Dikecligil, G. N., & Gottfried, J. A. (2024). What Does the Human Olfactory System Do, and How Does It Do It? *Annual Review of Psychology*, *75* (1), null. https://doi.org/10.1146/annurev-psych-042023-101155

Donahue, M. J., Hoogduin, H., van Zijl, P. C. M., Jezzard, P., Luijten, P. R., & Hendrikse, J. (2011). Blood oxygenation level-dependent (BOLD) total and extravascular signal changes and ΔR2* in human visual cortex at 1.5, 3.0 and 7.0 T. *NMR in Biomedicine*, *24* (1), 25–34. https://doi.org/10.1002/nbm.1552

Donoshita, Y., Choi, U.-S., Ban, H., & Kida, I. (2021). Assessment of olfactory information in the human brain using 7-Tesla functional magnetic resonance imaging. *NeuroImage*, *236*, 118212. https://doi.org/10.1016/j.neuroimage.2021.118212

Dourado, G. B., Volpato, G. H., de Almeida-Pedrin, R. R., Pedron Oltramari, P. V., Freire Fernandes, T. M., & de Castro Ferreira Conti, A. C. (2021). Likert scale vs

visual analog scale for assessing facial pleasantness. *American Journal of Orthodontics and Dentofacial Orthopedics*, *160* (6), 844–852. https://doi.org/10.1016/j.ajodo.2020.05.024

Dowdle, L. T., Vizioli, L., Moeller, S., Akçakaya, M., Olman, C., Ghose, G., Yacoub, E., & Uğurbil, K. (2023). Evaluating increases in sensitivity from NORDIC for diverse fMRI acquisition strategies. *NeuroImage*, *270*, 119949. https://doi.org/10.1016/j.neuroimage.2023.119949

DuPre, E., Salo, T., Ahmed, Z., Bandettini, P. A., Bottenhorn, K. L., Caballero-Gaudes, C., Dowdle, L. T., Gonzalez-Castillo, J., Heunis, S., Kundu, P., Laird, A. R., Markello, R., Markiewicz, C. J., Moia, S., Staden, I., Teves, J. B., Uruñuela, E., Vaziri-Pashkam, M., Whitaker, K., & Handwerker, D. A. (2021). TE-dependent analysis of multi-echo fMRI with *tedana*. *Journal of Open Source Software*, *6* (66), 3669. https://doi.org/10.21105/joss.03669

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nature Methods*, *16* (1), Article 1. https://doi.org/10.1038/s41592-018-0235-4

Farinha, M., Amado, C., Morgado, P., & Cabral, J. (2022). Increased Excursions to Functional Networks in Schizophrenia in the Absence of Task. *Frontiers in Neuroscience*, *16*. https://www.frontiersin.org/articles/10.3389/fnins.2022.821179

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41* (4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fera, F., Yongbi, M. N., van Gelderen, P., Frank, J. A., Mattay, V. S., & Duyn, J. H. (2004). EPI-BOLD fMRI of human motor cortex at 1.5 T and 3.0 T: Sensitivity dependence on echo time and acquisition bandwidth. *Journal of Magnetic Resonance Imaging*, *19* (1), 19–26. https://doi.org/10.1002/jmri.10440

Flint, A., Raben, A., Blundell, J. E., & Astrup, A. (2000). Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies. *International Journal of Obesity*, *24* (1), Article 1. https://doi.org/10.1038/sj.ijo.0801083

Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., & Collins, D. L. (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, *54* (1), 313–327. https://doi.org/10.1016/j.neuroimage.2010.07.033

Fonov, V., Evans, A., McKinstry, R., Almli, C., & Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage*, *47*, S102. https://doi.org/10.1016/S1053-8119(09)70884-5

Fournel, A., Ferdenzi, C., Sezille, C., Rouby, C., & Bensafi, M. (2016). Multidimensional representation of odors in the human olfactory cortex. *Human Brain Mapping*, *37* (6), 2161–2172. https://doi.org/10.1002/hbm.23164

Friston, K. J., Glaser, D. E., Henson, R. N. A., Kiebel, S., Phillips, C., & Ashburner, J. (2002). Classical and Bayesian Inference in Neuroimaging: Applications. *NeuroImage*, *16* (2), 484–512. https://doi.org/10.1006/nimg.2002.1091

Gati, J. S., Menon, R. S., Uğurbil, K., & Rutt, B. K. (1997). Experimental determination of the BOLD field strength dependence in vessels and tissue. *Magnetic Resonance in Medicine*, *38* (2), 296–302. https://doi.org/10.1002/mrm.1910380220

Glover, G. H., & Law, C. S. (2001). Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. *Magnetic Resonance in Medicine*, *46* (3), Article 3. https://doi.org/10.1002/mrm.1222

Gonzalez-Castillo, J., Panwar, P., Buchanan, L. C., Caballero-Gaudes, C., Handwerker, D. A., Jangraw, D. C., Zachariou, V., Inati, S., Roopchansingh, V., Derbyshire, J. A., & Bandettini, P. A. (2016). Evaluation of multi-echo ICA denoising for task based fMRI studies: Block designs, rapid event-related designs, and cardiac-gated fMRI. *NeuroImage*, *141*, 452–468. https://doi.org/10.1016/j.neuroimage.2016.07.049

Gorgolewski, K., Burns, C., Madison, C., Clark, D., Halchenko, Y., Waskom, M., & Ghosh, S. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Frontiers in Neuroinformatics*, *5*. https://www.frontiersin.org/articles/10.3389/fninf.2011.00013

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

Graham, S. J., Stanchev, P. L., & Bronskill, M. J. (1996). Criteria for analysis of multicomponent tissue T2 relaxation data. *Magnetic Resonance in Medicine*, *35* (3), 370–378. https://doi.org/10.1002/mrm.1910350315

Green, B. G., Dalton, P., Cowart, B., Shaffer, G., Rankin, K., & Higgins, J. (1996). Evaluating the 'Labeled Magnitude Scale' for Measuring Sensations of Taste and Smell. *Chemical Senses*, *21* (3), 323–334. https://doi.org/10.1093/chemse/21.3.323

Green, B. G., Shaffer, G. S., & Gilmore, M. M. (1993). Derivation and evaluation of a semantic scale of oral sensation magnitude with apparent ratio properties. *Chemical Senses*, *18* (6), 683–702. https://doi.org/10.1093/chemse/18.6.683

Howard, J. D., & Gottfried, J. A. (2014). Configural and Elemental Coding of Natural Odor Mixture Components in the Human Brain. *Neuron*, *84* (4), Article 4. https://doi.org/10.1016/j.neuron.2014.10.012

Howard, J. D., Plailly, J., Grueschow, M., Haynes, J.-D., & Gottfried, J. A. (2009). Odor quality coding and categorization in human posterior piriform cortex. *Nature Neuroscience*, *12* (7), Article 7. https://doi.org/10.1038/nn.2324

Im, K., Jo, H. J., Mangin, J.-F., Evans, A. C., Kim, S. I., & Lee, J.-M. (2010). Spatial Distribution of Deep Sulcal Landmarks and Hemispherical Asymmetry on the Cortical Surface. *Cerebral Cortex*, *20* (3), 602–611. https://doi.org/10.1093/cercor/bhp127

Jones, L. V., Peryam, D. R., & Thurstone, L. L. (1955). Development of a scale for measuring soldiers' food preferences. *Journal of Food Science*, *20* (5), 512–520. https://doi.org/10.1111/j.1365-2621.1955.tb16862.x

Jones, T. B., Bandettini, P. A., Kenworthy, L., Case, L. K., Milleville, S. C., Martin, A., & Birn, R. M. (2010). Sources of group differences in functional connectivity: An

investigation applied to autism spectrum disorder. *NeuroImage*, *49* (1), 401–414. https://doi.org/10.1016/j.neuroimage.2009.07.051

Kelly, R. E., Alexopoulos, G. S., Wang, Z., Gunning, F. M., Murphy, C. F., Morimoto, S. S., Kanellopoulos, D., Jia, Z., Lim, K. O., & Hoptman, M. J. (2010). Visual inspection of independent components: Defining a procedure for artifact removal from fMRI data. *Journal of Neuroscience Methods*, *189* (2), 233–245. https://doi.org/10.1016/j.jneumeth.2010.03.028

Kulkarni, P., Stolberg, T., SullivanJr., J. M., & Ferris, C. F. (2012). Imaging evolutionarily conserved neural networks: Preferential activation of the olfactory system by food-related odor. *Behavioural Brain Research*, *230* (1), 201–207. https://doi.org/10.1016/j.bbr.2012.02.002

Kundu, P., Brenowitz, N. D., Voon, V., Worbe, Y., Vértes, P. E., Inati, S. J., Saad, Z. S., Bandettini, P. A., & Bullmore, E. T. (2013). Integrated strategy for improving functional connectivity mapping using multiecho fMRI. *Proceedings of the National Academy of Sciences*, *110* (40), 16187–16192. https://doi.org/10.1073/pnas.1301725110

Kundu, P., Inati, S. J., Evans, J. W., Luh, W.-M., & Bandettini, P. A. (2012a). Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *NeuroImage*, *60* (3), 1759–1770. https://doi.org/10.1016/j.neuroimage.2011.12.028

Kundu, P., Inati, S. J., Evans, J. W., Luh, W.-M., & Bandettini, P. A. (2012b). Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo

EPI. *NeuroImage*, *60* (3), 1759–1770. https://doi.org/10.1016/j.neuroimage.2011.12.028

Kundu, P., Voon, V., Balchandani, P., Lombardo, M. V., Poser, B. A., & Bandettini, P. A. (2017). Multi-echo fMRI: A review of applications in fMRI denoising and analysis of BOLD signals. *NeuroImage*, *154*, 59–80. https://doi.org/10.1016/j.neuroimage.2017.03.033

Levene, H. (1960). Robust tests for equality of variances. *Contributions to Probability and Statistics*, 278–292.

Lim, J., Wood, A., & Green, B. G. (2009). Derivation and Evaluation of a Labeled Hedonic Scale. *Chemical Senses*, *34* (9), 739–751. https://doi.org/10.1093/chemse/bjp054

Liu, Q., Zhao, Y., Attanti, S., Voss, J. L., Schoenbaum, G., & Kahnt, T. (2024). Midbrain signaling of identity prediction errors depends on orbitofrontal cortex networks. *Nature Communications*, *15* (1), 1704. https://doi.org/10.1038/s41467-024-45880-1

Lynch, C. J., Power, J. D., Scult, M. A., Dubin, M., Gunning, F. M., & Liston, C. (2020). Rapid Precision Functional Mapping of Individuals Using Multi-Echo fMRI. *Cell Reports*, *33* (12), 108540. https://doi.org/10.1016/j.celrep.2020.108540

Mai, J. K., Majtanik, M., & Paxinos, G. (2016). *Atlas of the human brain* (4th edition). Academic Press.

Manera, A. L., Dadar, M., Fonov, V., & Collins, D. L. (2020). CerebrA, registration and manual label correction of Mindboggle-101 atlas for MNI-ICBM152 template. *Scientific Data*, *7* (1), Article 1. https://doi.org/10.1038/s41597-020-0557-9

McKeown, M. J., Hansen, L. K., & Sejnowsk, T. J. (2003). Independent component analysis of functional MRI: What is signal and what is noise? *Current Opinion in Neurobiology*, *13* (5), 620–629. https://doi.org/10.1016/j.conb.2003.09.012

Mckeown, M. J., Makeig, S., Brown, G. G., Jung, T.-P., Kindermann, S. S., Bell, A. J., & Sejnowski, T. J. (1998). Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, *6* (3), 160–188. https://doi.org/10.1002/(SICI)1097-0193(1998)6:3<160::AID-HBM5>3.0.CO;2-1

Miao, X., Paez, A. G., Rajan, S., Cao, D., Liu, D., Pantelyat, A. Y., Rosenthal, L. I., van Zijl, P. C. M., Bassett, S. S., Yousem, D. M., Kamath, V., & Hua, J. (2021). Functional Activities Detected in the Olfactory Bulb and Associated Olfactory Regions in the Human Brain Using T2-Prepared BOLD Functional MRI at 7T. *Frontiers in Neuroscience*, *15*. https://www.frontiersin.org/articles/10.3389/fnins.2021.723441

Moeller, S., Pisharady, P. K., Ramanna, S., Lenglet, C., Wu, X., Dowdle, L., Yacoub, E., Uğurbil, K., & Akçakaya, M. (2021). NOise reduction with DIstribution Corrected (NORDIC) PCA in dMRI with complex-valued parameter-free locally low-rank processing. *NeuroImage*, *226*, 117539. https://doi.org/10.1016/j.neuroimage.2020.117539

Parrish, T., Chen, Y., Li, W., Howard, J., & Gottfried, J. (2008). High resolution 3D functional images of the human olfactory bulb using passband SSFP at 3T. *Proceedings of International Society of Magnetic Resonance in Medicine*, 3567.

Patin, A., & Pause, B. M. (2015). Human amygdala activations during nasal chemoreception. *Neuropsychologia*, *78*, 171–194. https://doi.org/10.1016/j.neuropsychologia.2015.10.009

Pearson, K. (1895). Note on Regression and Inheritance in the Case of Two Parents. *Proceedings of the Royal Society of London*, *58*, 240–242.

Perszyk, E. E., Davis, X. S., & Small, D. M. (2023). Olfactory decoding is positively associated with ad libitum food intake in sated humans. *Appetite*, *180*, 106351. https://doi.org/10.1016/j.appet.2022.106351

Plailly, J., Howard, J. D., Gitelman, D. R., & Gottfried, J. A. (2008). Attention to Odor Modulates Thalamocortical Connectivity in the Human Brain. *Journal of Neuroscience*, *28* (20), 5257–5267. https://doi.org/10.1523/JNEUROSCI.5607-07.2008

Poser, B. A., Versluis, M. J., Hoogduin, J. M., & Norris, D. G. (2006). BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity-desensitized fMRI. *Magnetic Resonance in Medicine*, *55* (6), 1227–1235. https://doi.org/10.1002/mrm.20900

Posse, S., Wiese, S., Gembris, D., Mathiak, K., Kessler, C., Grosse-Ruyken, M.-L., Elghahwagi, B., Richards, T., Dager, S. R., & Kiselev, V. G. (1999). Enhancement of BOLD-contrast sensitivity by single-shot multi-echo functional MR imaging. *Magnetic Resonance in Medicine*, *42* (1), 87–97. https://doi.org/10.1002/(SICI)1522-2594(199907)42:1<87::AID-MRM13>3.0.CO;2-O

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, *59* (3), 2142–2154. https://doi.org/10.1016/j.neuroimage.2011.10.018

Raithel, C. U., Miller, A. J., Epstein, R. A., Kahnt, T., & Gottfried, J. A. (2023). Recruitment of grid-like responses in human entorhinal and piriform cortices by odor landmark-based navigation. *Current Biology*, *33* (17), 3561-3570.e4. https://doi.org/10.1016/j.cub.2023.06.087

Raj, D., Anderson, A. W., & Gore, J. C. (2001). Respiratory effects in human functional magnetic resonance imaging due to bulk susceptibility changes. *Physics in Medicine & Biology*, *46* (12), 3331. https://doi.org/10.1088/0031-9155/46/12/318

Raj, D., Paley, D. P., Anderson, A. W., Kennan, R. P., & Gore, J. C. (2000). A model for susceptibility artefacts from respiration in functional echo-planar magnetic resonance imaging. *Physics in Medicine & Biology*, *45* (12), 3809. https://doi.org/10.1088/0031-9155/45/12/321

Roth, Z. N., Ryoo, M., & Merriam, E. P. (2020). Task-related activity in human visual cortex. *PLOS Biology*, *18* (11), e3000921. https://doi.org/10.1371/journal.pbio.3000921

Savic, I. (2002). Imaging of brain activation by odorants in humans. *Current Opinion in Neurobiology*, *12* (4), 455–461. https://doi.org/10.1016/S0959-4388(02)00346-X

Savic, I., Gulyás, B., & Berglund, H. (2002). Odorant differentiated pattern of cerebral activation: Comparison of acetone and vanillin. *Human Brain Mapping*, *17* (1), 17–27. https://doi.org/10.1002/hbm.10045

Schutz, H. G., & Cardello, A. V. (2001). A Labeled Affective Magnitude (lam) Scale for Assessing Food Liking/Disliking1. *Journal of Sensory Studies*, *16* (2), 117–159. https://doi.org/10.1111/j.1745-459X.2001.tb00293.x

Shanahan, L. K., Bhutani, S., & Kahnt, T. (2021). Olfactory perceptual decision-making is biased by motivational state. *PLOS Biology*, *19* (8), e3001374. https://doi.org/10.1371/journal.pbio.3001374

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences.* (pp. xvii, 312). McGraw-Hill.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23*, S208–S219. https://doi.org/10.1016/j.neuroimage.2004.07.051

Stone, H., Bleibaum, R. N., & Thomas, H. A. (2020). *Sensory evaluation practices*. Academic press.

Swarup, S., Huang, W., Mackay, T. F. C., & Anholt, R. R. H. (2013). Analysis of natural variation reveals neurogenetic networks for Drosophila olfactory behavior. *Proceedings of the National Academy of Sciences*, *110* (3), 1017–1022. https://doi.org/10.1073/pnas.1220168110

Triantafyllou, C., Wald, L. L., & Hoge, R. D. (2011). Echo-Time and Field Strength Dependence of BOLD Reactivity in Veins and Parenchyma Using Flow-

Normalized Hypercapnic Manipulation. *PLOS ONE*, *6* (9), e24519. https://doi.org/10.1371/journal.pone.0024519

Van, A. N., Tisdall, M. D., Laumann, T. O., Montez, D. F., Fair, D. A., & Dosenbach, N. U. F. (2023). Framewise distortion correction in Multi-echo Echo Planar Imaging. *Proceedings of International Society of Magnetic Resonance in Medicine*.

Van de Moortele, P.-F., Pfeuffer, J., Glover, G. H., Ugurbil, K., & Hu, X. (2002). Respiration-induced B0 fluctuations and their spatial distribution in the human brain at 7 Tesla. *Magnetic Resonance in Medicine*, *47* (5), 888–895. https://doi.org/10.1002/mrm.10145

van der Zwaag, W., Francis, S., Head, K., Peters, A., Gowland, P., Morris, P., & Bowtell, R. (2009). fMRI at 1.5, 3 and 7 T: Characterising BOLD signal changes. *NeuroImage*, *47* (4), 1425–1434. https://doi.org/10.1016/j.neuroimage.2009.05.015

Vizioli, L., Moeller, S., Dowdle, L., Akçakaya, M., De Martino, F., Yacoub, E., & Uğurbil, K. (2021). Lowering the thermal noise barrier in functional brain mapping with magnetic resonance imaging. *Nature Communications*, *12* (1), Article 1. https://doi.org/10.1038/s41467-021-25431-8

Weiskopf, N., Hutton, C., Josephs, O., & Deichmann, R. (2006). Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: A whole-brain analysis at 3 T and 1.5 T. *NeuroImage*, *33* (2), 493–504. https://doi.org/10.1016/j.neuroimage.2006.07.029

Yekutieli, D., & Benjamini, Y. (1999). Resampling-based false discovery rate controlling multiple test procedures for correlated test statistics. *Journal of Statistical*

Planning and Inference, *82* (1), 171–196. https://doi.org/10.1016/S0378-3758(99)00041-5

Zhou, G., Lane, G., Cooper, S. L., Kahnt, T., & Zelano, C. (2019). Characterizing functional pathways of the human olfactory system. *eLife*, *8*, e47177. https://doi.org/10.7554/eLife.47177

Supplementary Materials

Leveraging Multi-Echo EPI to Enhance BOLD Sensitivity in Task-based Olfactory fMRI

Ludwig Sichen Zhao[*1,2], Clara U. Raithel[3,2], M. Dylan Tisdall[4], John A. Detre[2,4], and Jay A. Gottfried[*2,3]

[1] *Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, United States*
[2] *Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States*
[3] *Department of Psychology, School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, United States*
[4] *Department of Radiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States*

* Corresponding Author: Ludwig Sichen Zhao (ludwigz@seas.upenn.edu), Jay A. Gottfried (jaygottf@pennmedicine.upenn.edu)

Table of Contents

1 Detailed fMRI Data Preprocessing 1

1.1 Preprocessing of B_0 Inhomogeneity Mappings 1

1.2 Anatomical Data Preprocessing 1

1.3 Functional Data Preprocessing 2

2 Impacts of ME-WC and ME-ICA on the Distribution of Statistical Significance 4

3 T_2^* Changes Due to Respiration 6

4 Nonexponential Decay of Multi-Echo GE-EPI Signal 7

5 Supplementary Figures and Tables 9

6 Supplementary References 13

1 Detailed fMRI Data Preprocessing

Results included in this manuscript come from preprocessing performed using *fMRIPrep* 21.0.0 (Esteban, Markiewicz, et al. (2018); Esteban, Blair, et al. (2018); RRID:SCR_016216), which is based on *Nipype* 1.6.1 (K. Gorgolewski et al. (2011); K. J. Gorgolewski et al. (2018); RRID:SCR_002502).

1.1 Preprocessing of B_0 Inhomogeneity Mappings

A total of 2 fieldmaps were found available within the input BIDS structure for this particular subject. A B_0-nonuniformity map (or *fieldmap*) was estimated based on two (or more) echo-planar imaging (EPI) references with topup (Andersson, Skare, and Ashburner (2003); FSL 6.0.5.1:57b01774).

1.2 Anatomical Data Preprocessing

A total of 1 T1-weighted (T1w) images were found within the input BIDS dataset. The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al. 2010), distributed with ANTs 2.3.3 (Avants et al. 2008, RRID:SCR_004757), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a *Nipype* implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 6.0.5.1:57b01774, RRID:SCR_002823, Zhang, Brady, and Smith 2001). Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, RRID:SCR_001847, Dale, Fischl, and Sereno 1999), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (RRID:SCR_002438, Klein et al. 2017). Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial

normalization: *ICBM 152 Nonlinear Asymmetrical template version 2009c* [Fonov et al. (2009), RRID:SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym].

1.3 Functional Data Preprocessing

For each of the 6 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated by aligning and averaging 5 single-band references (SBRefs). For single-echo dataset, the reference volume is the single-band references (SBRefs). Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 6.0.5.1:57b01774, Jenkinson et al. 2002). The estimated *fieldmap* was then aligned with rigid-registration to the target EPI (echo-planar imaging) reference run. The field coefficients were mapped on to the reference EPI using the transform. BOLD runs were slice-time corrected to 0.969s (0.5 of slice acquisition range 0s-1.94s) using 3dTshift from AFNI (Cox and Hyde 1997, RRID:SCR_005927). The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration (Greve and Fischl 2009). Co-registration was configured with six degrees of freedom. First, a reference volume and its skull-stripped version were generated using a custom methodology of *fMRIPrep*. Several confounding time-series were calculated based on the *preprocessed BOLD*: framewise displacement (FD), DVARS and three region-wise global signals. FD was computed using two formulations following Power (absolute sum of relative motions, Power et al. (2014)) and Jenkinson (relative root mean square displacement between affines, Jenkinson et al. (2002)). FD and DVARS are calculated for each functional run, both using their implementations in *Nipype* (following the definitions by Power et al. 2014). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Additionally, a set of physiological regressors were extracted to allow for component-based noise correction (*CompCor*, Behzadi et al. 2007). Principal components are estimated after high-pass filtering the *preprocessed BOLD* time-series (using a discrete cosine filter with 128s cut-off) for the two *CompCor* variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor

components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM and combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by 2 pixels on BOLD space, the aCompCor masks are subtracted a mask of pixels that likely contain a volume fraction of GM. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's *aseg* segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the *k* components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50 percent of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration. The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each (Satterthwaite et al. 2013). Frames that exceeded a threshold of 0.5 mm FD or 1.5 standardised DVARS were annotated as motion outliers. All resamplings can be performed with *a single interpolation step* by composing all the pertinent transformations (i.e. head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos 1964). Non-gridded (surface) resamplings were performed using mri_vol2surf (FreeSurfer).

Many internal operations of *fMRIPrep* use *Nilearn* 0.8.1 (Abraham et al. 2014, RRID:SCR_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in *fMRIPrep*'s documentation.

2 Impacts of ME-WC and ME-ICA on the Distribution of Statistical Significance

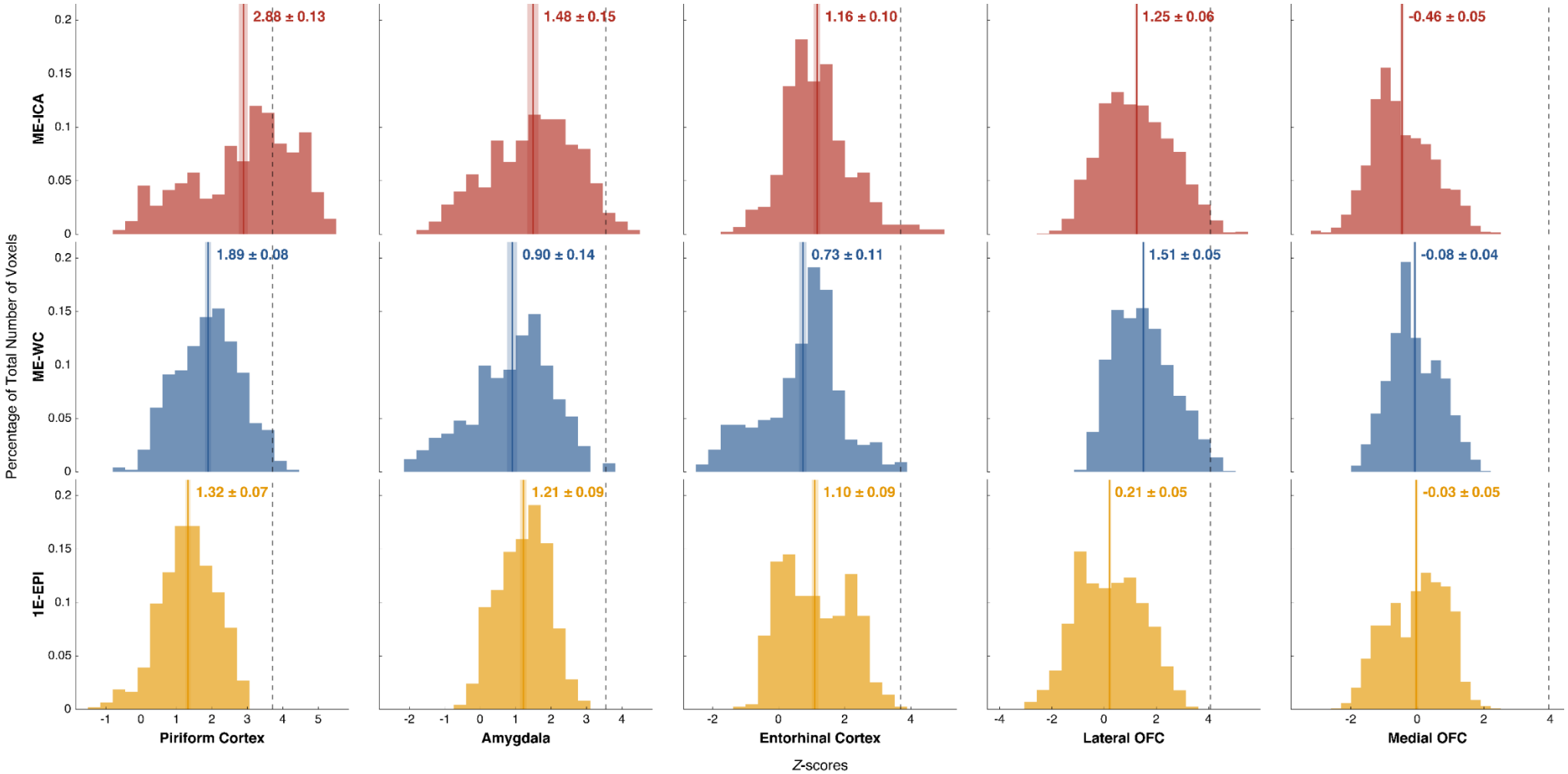


Figure S1. Histograms of Group Level Z-Scores for All ROIs for [Lemon > Control] Contrast. The mean ± 95% confidence interval of the *Z*-scores within each ROI for each acquisition is annotated. The mean is also marked by the solid line surrounded by a shaded area indicating the 95% confidence interval. The dashed line indicates the threshold corresponding to statistical significance ($p < 0.05$, SVC).

In Figure S1, the number of voxels (cluster size) exceeding the statistical significance thresholds (indicated by the dashed lines) and the peak *Z*-scores are higher for ME-WC compared to 1E-EPI. ME-ICA further increases the number of voxels above these thresholds, which is consistent with the results from voxel-wise whole brain analysis (Table 2 and Figure 3b). When examining the mean values and the 95%

confidence intervals (CI) of the distribution, we observed that ME-WC has more voxels shifting left past the thresholds, increasing the suprathreshold cluster size as well as peak *Z*-scores and thus widening the distribution and increasing variance. In other words, ME-WC tends to have a wider CI, indicating larger variance and a broader spread of the distribution.

The significance of the ROI analysis depends on both the mean and the variance of the distribution. In some regions, such as the entorhinal cortex, ME-WC increases the within-ROI variance without increasing the mean, and thus the statistical significance decreases compared to 1E-EPI. We interpret this as meaning that while ME-WC has increased sensitivity to positively correlated BOLD activation, it has also increased the sensitivity to anti-correlated signals.

3 T_2^* Changes Due to Respiration

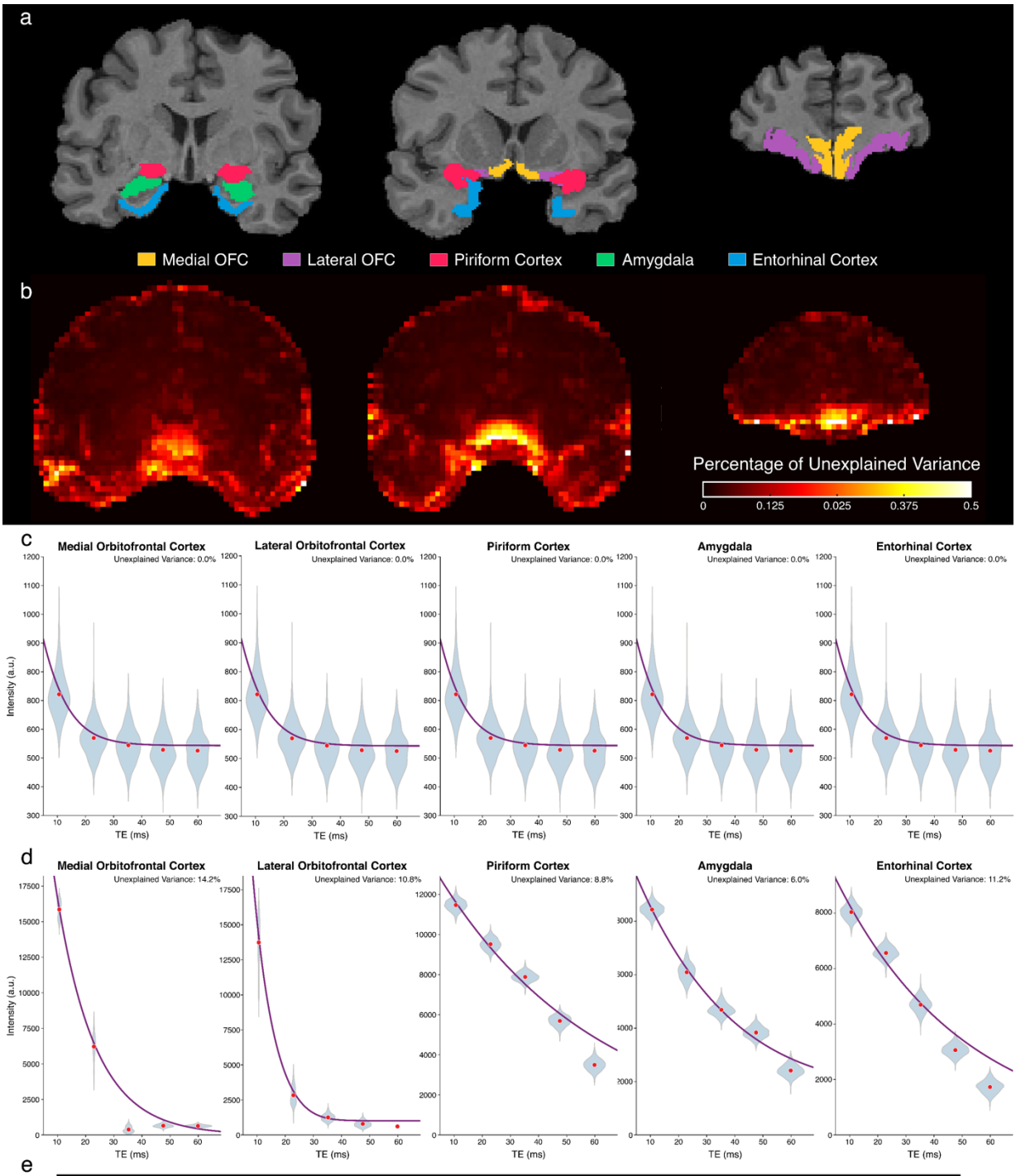
	p-value	T_2^* [Inhalation] (ms)	T_2^* [Exhalation] (ms)	ΔT_2^* [Inhalation – Exhalation] (ms)
Piriform Cortex	0.0109*	54.50 ± 1.48	54.13 ± 1.45	0.36 ± 0.26
Amygdala	0.0011**	52.32 ± 1.76	52.12 ± 1.75	0.20 ± 0.11
Entorhinal Cortex	0.0004***	33.25 ± 2.48	33.12 ± 2.45	0.13 ± 0.06
Lateral OFC	0.0521	33.45 ± 1.09	33.38 ± 1.11	0.06 ± 0.06
Medial OFC	0.0334*	24.92 ± 1.08	24.87 ± 1.08	0.05 ± 0.04

Table S1. T_2^* Changes during Inhalation and Exhalation Phase in All Five ROIs. The *p*-values indicate the statistical significance of the two-sample, two-sided *t*-tests comparing T_2^* values estimated from volumes during the inhalation phase with those from the exhalation phase. T_2^* values and their differences (ΔT_2^*) are reported as mean values within the each ROI across all subjects ± 95% confidence intervals.

To investigate T_2^* changes during respiration, we estimated T_2^* maps for each subject at both inhalation and exhalation phases. As described in the Methods section, the respiratory trace was downsampled to match the *TR* of the fMRI acquisition. Inhalation phase volumes are defined as those 0.8 standard deviations above the mean of the downsampled respiratory trace, while exhalation phase volumes are defined as those 0.8 standard deviations below the mean.

We extracted the mean T_2^* values for each subject at both inhalation and exhalation phases and performed two-sample, two-sided *t*-tests to determine if there were statistically significant differences between the two phases. All ROIs, except the lateral orbitofrontal cortices (OFC), showed significant, though subtle, differences, with increased T_2^* values during the inhalation phase, as shown in Table S1. As a result, the weights in the T_2^*-weighted combination method (ME-WC) remain relatively stable across different respiratory phases.

4 Nonexponential Decay of Multi-Echo GE-EPI Signal



e

	Medial OFC	Lateral OFC	Piriform Cortex	Amygdala	Entorhinal Cortex
Unexplained Variance (%)	14.15 ± 0.82	10.84 ± 0.36	8.79 ± 0.60	5.98 ± 0.59	11.11 ± 0.54

Figure S2. Percentage of Unexplained Variance in the Exponential Decay Signal across Different Echo Times for a Single Subject. a) Three representative coronal slices of a single subject with ROIs overlaid on T_1-weighted anatomical images. b) The percentage of unexplained variance map for the representative coronal slices. c) Signal across different echo times for voxels with the minimum unexplained variance in each ROI. The red dots represent the mean signal intensity over 600 measurements (200 volumes per run, 3 runs in total), and the blue violin plots show the distribution of signal intensity. The purple lines indicate the fitted model using the method described in Section 2.7. d) Signal across different echo times for voxels with the unexplained variance closest to the mean unexplained variance within each ROI. The red dots, blue violin plots, and purple lines represent the mean signal intensity, distributions, and fitted model, respectively. e) The mean ± 95% confidence interval of unexplained variance for each ROI.

Theoretically, the BOLD signal under GE-EPI should exhibit an exponential decay across echo times. To examine deviations from this expected exponential decay, we calculated the voxel-wise percentage of unexplained variance. This calculation was performed by dividing the root mean square (RMS) of the residuals by the RMS of the observed signal, as shown in Figure S3. Regions with high susceptibility, such as the medial and lateral orbitofrontal cortices (OFC) and the entorhinal cortex, demonstrated a more pronounced nonexponential decay.

5 Supplementary Figures and Tables

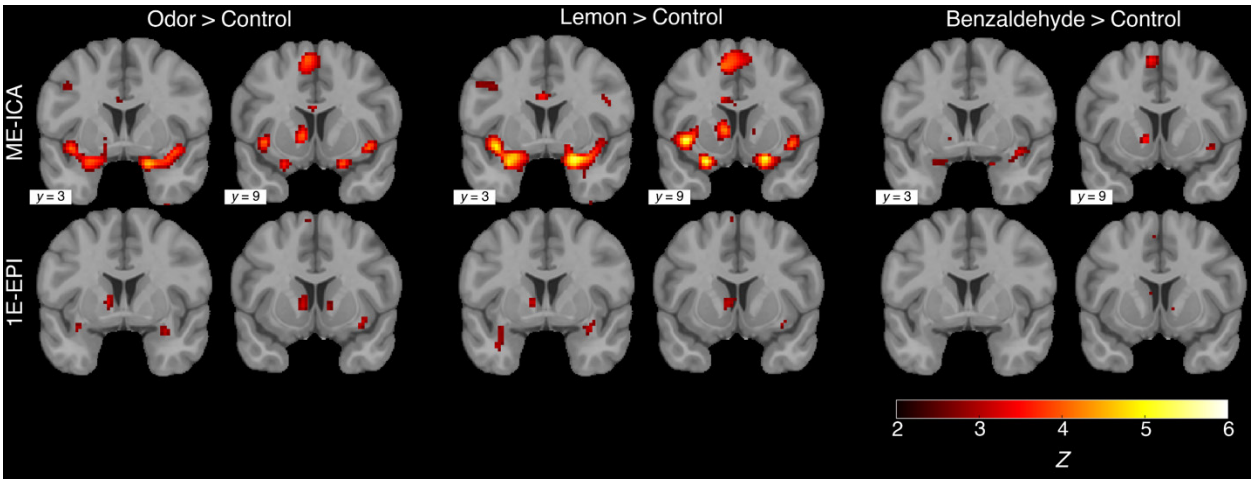


Figure S3. Comparison of Activation Maps among Different Contrasts for Olfactory-Related Tasks in Both 1E-EPI and ME-EPI. Two representative coronal slices of activation maps for all three contrasts ([Odor > Control], [Benzaldehyde > Control], [Lemon > Control]) using ME-EPI acquisition with ME-ICA denoising methods (ME-ICA) and using single-echo fMRI (1E-EPI) acquisition ($p < 0.001$, uncorrected).

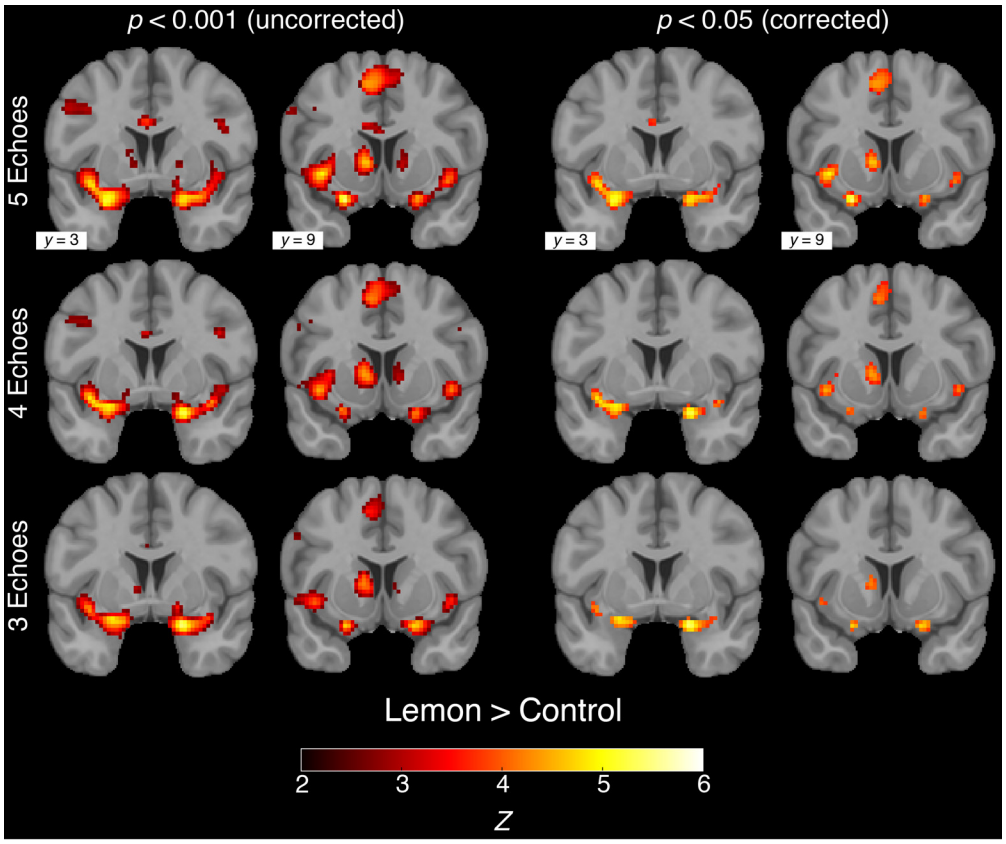


Figure S4. Comparison of Activation Maps for ME-EPI Acquisition with ME-ICA Denoising among Different Number of Echoes for Olfactory-Related Tasks. Two representative coronal slices of activation maps for [Lemon > Control] using ME-EPI acquisition with ME-ICA denoising methods (ME-ICA) with first 3, 4 and all 5 echoes. The left two columns show the maps without multiple comparison correction ($p < 0.001$) while the right two columns show the maps corrected for multiple comparison ($p < 0.05$, FDR corrected).

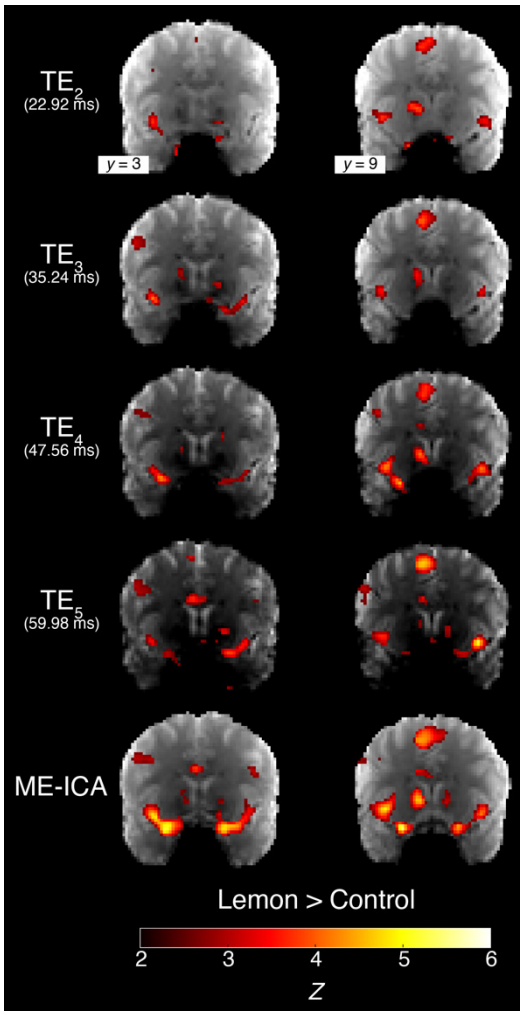


Figure S5. Comparison of Activation Maps When Analyzing Individual Echoes from ME-ICA Separately for Olfactory-Related Tasks. Two representative coronal slices of activation maps for [Lemon > Control] when analyzing individual echoes from ME-EPI acquisition with ME-ICA denoising methods (ME-ICA) separately on the group level ($p < 0.05$, uncorrected). The activation maps are overlaid on the mean image of a single subject across one run, registered to MNI template space, to illustrate signal dropout across different echoes.

	Piriform Cortex	Amygdala	Entorhinal Cortex	Orbitofrontal Cortex
TE_1 (10.60 ms)	1.945	1.010	0.756	1.205
TE_2 (22.92 ms)	2.383	1.957	2.183	1.605
TE_3 (35.24 ms)	3.457	1.583	1.789	1.724
TE_4 (47.56 ms)	3.819	1.355	0.973	1.835
TE_5 (59.98 ms)	3.693	2.604	1.265	1.705
1E-EPI (22.00 ms)	2.479	1.972	2.152	-0.279

Table S2. Z-Scores of ROI Analyses for Each Olfactory-related ROI under the Contrast [Lemon > Control] When Analyzing 5 Individual Echoes from ME-ICA Separately.

	Medial OFC	Lateral OFC	Piriform Cortex	Amygdala	Entorhinal Cortex
Unmodified T_2^* Estimation (ms)	21.9 ± 0.6	30.2 ± 0.6	45.8 ± 1.1	45.6 ± 1.1	29.0 ± 1.0
Modified T_2^* Estimation (ms)	13.0 ± 0.2	21.8 ± 0.2	41.3 ± 1.1	42.2 ±1.0	22.2 ± 0.3

Table S3. Comparison of Estimated T_2^* Values between Unmodified and Modified Fitting Methods for All ROIs within a Single Subject. Unmodified T_2^* estimation method fits Eq. (1) for each voxel. Modified T_2^* estimation method fits Eq. (3) instead.

6 Supplementary References

Abraham, Alexandre, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Mueller, Jean Kossaifi, Alexandre Gramfort, Bertrand Thirion, and Gael Varoquaux. 2014. "Machine Learning for Neuroimaging with Scikit-Learn." *Frontiers in Neuroinformatics* 8. https://doi.org/10.3389/fninf.2014.00014.

Andersson, Jesper L. R., Stefan Skare, and John Ashburner. 2003. "How to Correct Susceptibility Distortions in Spin-Echo Echo-Planar Images: Application to Diffusion Tensor Imaging." *NeuroImage* 20 (2): 870–88. https://doi.org/10.1016/S1053-8119(03)00336-7.

Avants, B. B., C. L. Epstein, M. Grossman, and J. C. Gee. 2008. "Symmetric Diffeomorphic Image Registration with Cross-Correlation: Evaluating Automated Labeling of Elderly and Neurodegenerative Brain." *Medical Image Analysis* 12 (1): 26–41. https://doi.org/10.1016/j.media.2007.06.004.

Behzadi, Yashar, Khaled Restom, Joy Liau, and Thomas T. Liu. 2007. "A Component Based Noise Correction Method (CompCor) for BOLD and Perfusion Based fMRI." *NeuroImage* 37 (1): 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042.

Cox, Robert W., and James S. Hyde. 1997. "Software Tools for Analysis and Visualization of fMRI Data." *NMR in Biomedicine* 10 (4-5): 171–78. https://doi.org/10.1002/(SICI)1099-1492(199706/08)10:4/5<171::AID-NBM453>3.0.CO;2-L.

Dale, Anders M., Bruce Fischl, and Martin I. Sereno. 1999. "Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction." *NeuroImage* 9 (2): 179–94. https://doi.org/10.1006/nimg.1998.0395.

Esteban, Oscar, Ross Blair, Christopher J. Markiewicz, Shoshana L. Berleant, Craig Moodie, Feilong Ma, Ayse Ilkay Isik, et al. 2018. "fMRIPrep." *Software*. https://doi.org/10.5281/zenodo.852659.

Esteban, Oscar, Christopher Markiewicz, Ross W Blair, Craig Moodie, Ayse Ilkay Isik, Asier Erramuzpe Aliaga, James Kent, et al. 2018. “fMRIPrep: A Robust Preprocessing Pipeline for Functional MRI.” *Nature Methods*. https://doi.org/10.1038/s41592-018-0235-4.

Fonov, VS, AC Evans, RC McKinstry, CR Almli, and DL Collins. 2009. “Unbiased Nonlinear Average Age-Appropriate Brain Templates from Birth to Adulthood.” *NeuroImage* 47, Supplement 1: S102. https://doi.org/10.1016/S1053-8119(09)70884-5.

Gorgolewski, K., C. D. Burns, C. Madison, D. Clark, Y. O. Halchenko, M. L. Waskom, and S. Ghosh. 2011. “Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python.” *Frontiers in Neuroinformatics* 5: 13. https://doi.org/10.3389/fninf.2011.00013.

Gorgolewski, Krzysztof J., Oscar Esteban, Christopher J. Markiewicz, Erik Ziegler, David Gage Ellis, Michael Philipp Notter, Dorota Jarecka, et al. 2018. “Nipype.” *Software*. https://doi.org/10.5281/zenodo.596855.

Greve, Douglas N, and Bruce Fischl. 2009. “Accurate and Robust Brain Image Alignment Using Boundary-Based Registration.” *NeuroImage* 48 (1): 63–72. https://doi.org/10.1016/j.neuroimage.2009.06.060.

Jenkinson, Mark, Peter Bannister, Michael Brady, and Stephen Smith. 2002. “Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images.” *NeuroImage* 17 (2): 825–41. https://doi.org/10.1006/nimg.2002.1132.

Klein, Arno, Satrajit S. Ghosh, Forrest S. Bao, Joachim Giard, Yrjö Häme, Eliezer Stavsky, Noah Lee, et al. 2017. “Mindboggling Morphometry of Human Brains.” *PLOS Computational Biology* 13 (2): e1005350. https://doi.org/10.1371/journal.pcbi.1005350.

Lanczos, C. 1964. “Evaluation of Noisy Data.” *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis* 1 (1): 76–85. https://doi.org/10.1137/0701007.

Posse, Stefan, Stefan Wiese, Daniel Gembris, Klaus Mathiak, Christoph Kessler, Maria-Liisa Grosse-Ruyken, Barbara Elghahwagi, Todd Richards, Stephen R. Dager, and Valerij G. Kiselev. 1999. “Enhancement of BOLD-Contrast Sensitivity by Single-Shot Multi-Echo Functional MR Imaging.” *Magnetic Resonance in Medicine* 42 (1): 87–97. https://doi.org/10.1002/(SICI)1522-2594(199907)42:1<87::AID-MRM13>3.0.CO;2-O.

Power, Jonathan D., Anish Mitra, Timothy O. Laumann, Abraham Z. Snyder, Bradley L. Schlaggar, and Steven E. Petersen. 2014. “Methods to Detect, Characterize, and Remove Motion Artifact in Resting State fMRI.” *NeuroImage* 84 (Supplement C): 320–41. https://doi.org/10.1016/j.neuroimage.2013.08.048.

Satterthwaite, Theodore D., Mark A. Elliott, Raphael T. Gerraty, Kosha Ruparel, James Loughead, Monica E. Calkins, Simon B. Eickhoff, et al. 2013. “An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data.” *NeuroImage* 64 (1): 240–56. https://doi.org/10.1016/j.neuroimage.2012.08.052.

Tustison, N. J., B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee. 2010. “N4itk: Improved N3 Bias Correction.” *IEEE Transactions on Medical Imaging* 29 (6): 1310–20. https://doi.org/10.1109/TMI.2010.2046908.

Zhang, Y., M. Brady, and S. Smith. 2001. “Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm.” *IEEE Transactions on Medical Imaging* 20 (1): 45–57. https://doi.org/10.1109/42.906424