An Efficient Multifunction fMRI Localizer for High-Level Visual, Auditory, and Cognitive Regions in Humans

Ammar I. Marvi† Sam Hutchinson† Evelina Fedorenko Rebecca R. Saxe
Frederik S. Kamps Tamar I. Regev Emily M. Chen
Nancy G. Kanwisher

MIT

*Correspondence: ngk@mit.edu

†These authors contributed equally

August 28, 2025

Keywords: human brain, functional regions of interest and functional localization

Abstract: One of the most robust findings in human cognitive neuroscience is the discovery that many regions of the cortex are engaged in distinctive, often very specific, functions. Although these regions are found in approximately the same location in almost all typical participants, their exact location varies from one individual to the next. Thus the first step in studying these regions is to identify them in each participant individually. Standard functional localizers have been devised to accomplish this goal, but most localizers identify only a few regions. Many important questions in modern neuroscience can only be answered by measuring the responses of multiple cortical regions at the same time. Here we introduce a new Efficient Multifunction fMRI localizer (EMFL) in which visual and auditory conditions are presented simultaneously, enabling the identification, in just 14 minutes of fMRI scan time, of fourteen of the most widely-studied cortical regions: those selectively engaged in perceiving faces, places, bodies, words, objects, and speech sounds; understanding language and other people's thoughts; and engaging broadly in demanding cognitive tasks (the "multiple demand" system). We validate the EMFL by showing that it identifies the major functional regions of interest as well as the standard localizers do, in a quarter of the scan

time. The stimuli and presentation code for this new localizer are publicly available, enabling future studies to efficiently identify the major functional regions of interest with the same procedure across labs.

1 Introduction

The last few decades of research in human cognitive neuroscience have revealed the functional organization of the human brain in considerable detail. Prominent in this organization is a set of cortical regions with functionally distinctive and often highly specific response profiles that are present in approximately the same location in virtually every neurotypical participant. The most robust and widely replicated of these regions respond selectively to visual images of faces, places, bodies, or text, or to auditory clips containing speech sounds, or during high-level cognitive operations like understanding language or thinking about other people's thoughts. Other regions of the parietal and frontal lobes are distinguished by their lack of functional specificity, i.e. by their broad engagement across multiple cognitively demanding tasks. The existence of this highly systematic functional organization of the cortex invites a deeper set of questions into the representations, computations, connectivity, cytoarchitecture, development, and evolution of these regions, as well as the interactions among them and their possible alterations in clinical disorders. To approach these questions, the first step is to find these regions in new participants. There is just one problem.

Although each of these functionally distinctive cortical regions is found in approximately the same location across individuals, their exact location varies considerably from brain to brain. As a result, purely anatomical landmarks are insufficient to precisely identify the location of these regions in new participants. The only way to identify the exact location and extent of any of these regions in a particular individual is to scan them with fMRI on a "localizer" scan that includes the functional contrast by which that region or set of regions is defined (e.g., a higher response to images of faces than objects in the lateral fusiform gyrus identifies the fusiform face area). For this purpose, standard localizer contrasts and paradigms have been developed to identify these regions, enabling a cumulative research program in which studies across participants and labs can refer to the same region because it is identified by the same functional contrast. However, many scientific questions can only be answered by measuring responses of many of these regions at once. Running the standard localizer scans to identify a wide selection of the most-commonly studied regions can take upwards of an hour, placing a significant burden on both the participant's patience and on the PI's budget. Here we present a new Efficient Multifunction Localizer (EMFL) that enables robust localization of all of the regions with the functional selectivities listed above in each individual participant

in as little as 14 minutes of scan time.

To accomplish this, we implement a blocked design (Figure 1) in which participants watch sets of short video clips from five different visual categories (faces, bodies, scenes, objects, and words), while simultaneously listening to and performing unrelated tasks on five different kinds of audio stimuli (false belief stories, false photo stories, arithmetic problems, non-word strings, and 'quilted' or scrambled, speech). We counterbalance the pairing of each visual-auditory combination across five runs to unconfound the responses to stimuli in one modality from responses to the other. We then assess the efficacy of this new localizer by testing whether it reliably identifies voxels whose selectivity cross-validates with held-out data from the EMFL and with the standard localizers for these regions. We find that the EMFL identifies similar fROIs with similar response profiles to those observed in standard localizers, thereby enabling the main functional regions to be identified in a unified and short localizer scan.

Condition	**Task**	**First statement**	**Second statement**
False belief stories	Is the second statement consistent with the first?	Laura didn't have time to braid her horse's mane before going to camp. While she was at camp, William braided the horse's mane for her.	Laura returns assuming that her horse's mane is braided.
False photo stories	Is the second statement consistent with the first?	The traffic camera snapped an image of the black car as it sped through the stoplight. Soon after, the car was painted red.	According to the traffic camera, the car is black.
Non-word strings	Is the gender of the second speaker the same as the first speaker?	(in female voice) Alf nimes beas brole inca coaker nour tunt ang drare sha doga. bave heers titer, sha bagners wole uld ake weps rho freir coaker fom breekmost.	(in female voice) Sha bawners ank surskills tu fand sha somo sussing fror sha coaker.
Audio quilts	Is the gender of the second speaker the same as the first speaker?	(in female voice) *unintelligible audio*	(in male voice) *unintelligible audio*
Arithmetic	Is the given solution to the equation correct?	Nine. times three. plus three. minus six. divided by eight. times two	Equals six.

Table 1: **Example stimuli of the EMFL auditory task.** Auditory stimuli occurred in five Conditions. Participants performed variants of a *'Does this come next?'* Task, comprised of a First statement followed by a Second statement. Correct responses (top to bottom): no, yes, yes, no, yes.

2 Methods

This study was approved by and in full compliance with the WCG IRB Connexus board. The methods and analyses of this study were pre-registered on OSF (materials available here: https://doi.org/10.17605/OSF.IO/GJSDB) prior to data collection and analysis. Data from an initial pilot phase (15 subjects, 9 females) were collected and analyzed to refine the experimental design prior to pre-registration. Scripts for the EMFL and relevant analyses are available at these GitHub repositories: https://github.mit.

edu/kanlab/efficient_localizer and https://github.com/aimarvi/emfl_analysis. fMRI data are available on OpenNeuro: https://openneuro.org/datasets/ds006179/versions/1.0.1.

2.1 EMFL design

The EMFL uses a blocked design in which five visual conditions are crossed with five unrelated but simultaneously-presented auditory conditions (see Figure 1). Tasks are performed only on the auditory stimuli, on the hypothesis (validated here) that visual category-selective regions show selective activation for their preferred stimuli even when participants are conducting a demanding orthogonal task on auditory stimuli. Visual conditions consist of dynamic videos of 1) faces, 2) scenes, 3) objects, 4) body parts (excluding faces), and 5) five-letter consonant strings superimposed on a scrambled-object background. Auditory conditions (Table 1) consist of 1) stories describing false beliefs, 2) stories describing a false physical representation ("false photo" stories), 3) spoken non-word strings, 4) synthesized audio quilts (Overath et al., 2015) of the non-word strings, and 5) arithmetic problems.

Each of the 25 possible combinations of a visual and an auditory condition occurs twice across five runs, with each run lasting 274 seconds (ten 22-second stimulus blocks and three 18-second fixation blocks). Each condition occurs twice per run in palindromic (for visual conditions) or semi-palindromic (for auditory conditions) order. A single block consists of a visual component and an auditory component presented simultaneously. The visual component consists of seven three-second clips of a single visual condition (e.g., videos of human faces). Face, body, and object videos were taken from a widely used dynamic localizer (Pitcher et al., 2011); scene videos were egocentric navigation videos selected as those producing the strongest response in the three scene-selective regions in a prior study (Chen et al., 2024); words were superimposed on grid-scrambled versions of the object videos. At the same time, a single exemplar from an auditory condition is played for 20.5 seconds, consisting of a long main clip and a second, shorter clip, followed by a silent response period of 1.5 seconds.

Participants are instructed to fixate on the visual stimuli while performing a *'Does this come next?'* task on the auditory stimuli. After listening to the main clip of audio, subjects hear a short continuation of the clip and are asked to decide whether the second clip is consistent with the first, responding *yes/no* via a button box. The exact meaning of *'Does this come next?'* varies depending on the content of the auditory stimulus. Example stimuli are shown in Table 1. A yellow outline appears around the edges of the screen indicating the time in which participants can submit their response to the *'Does this come next?'* task and remains present for the duration of the response period.

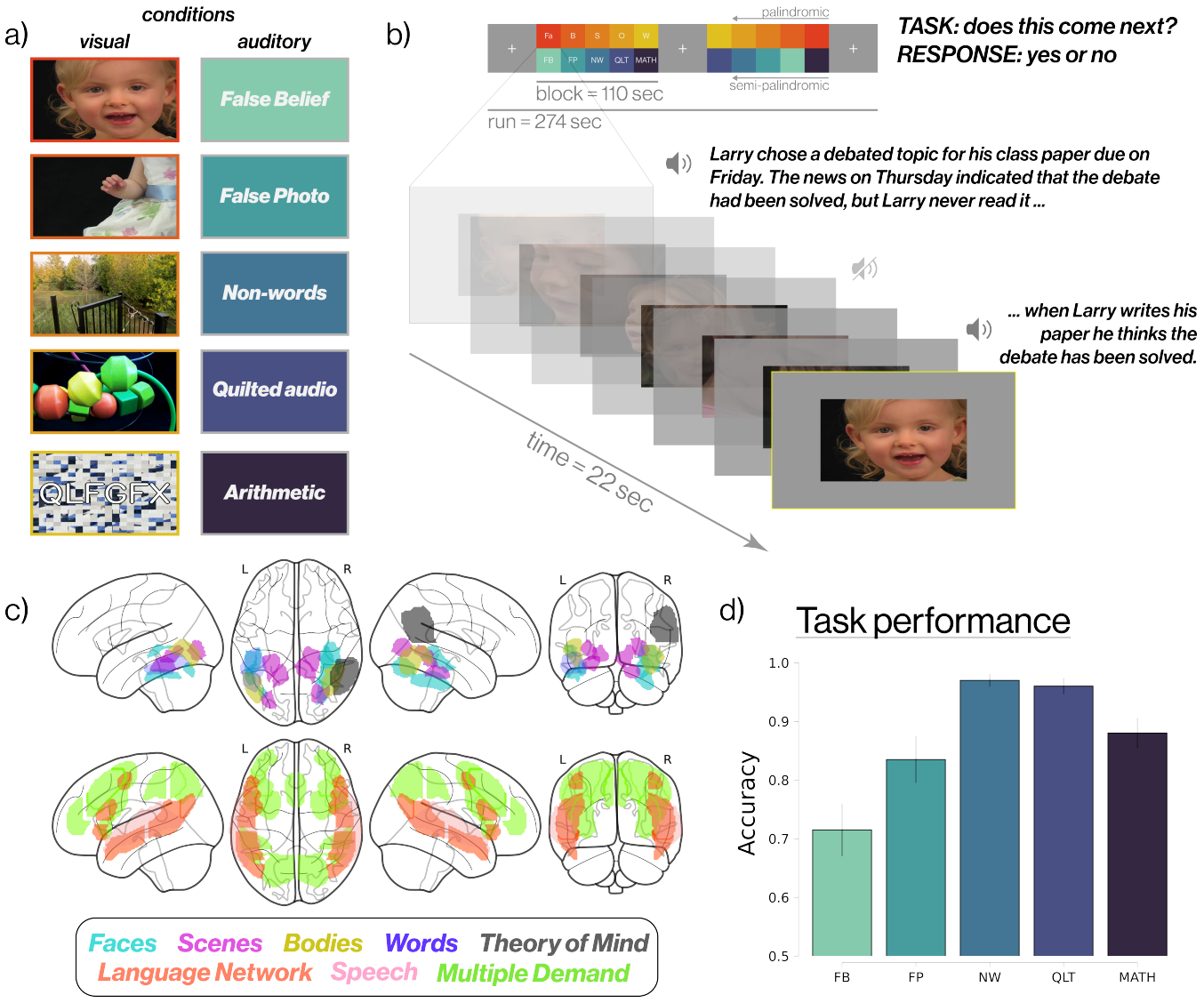


Figure 1: **Design of the Efficient Multifunction Localizer (EMFL)** a) the five visual (left) and five auditory (right) conditions used in EMFL. b) blocked structure of an example run (top) and presentation of stimuli within a block (bottom). Each block implements one visual and one auditory condition; Participants perform a *'Does this come next?'* task only on the auditory stimuli. Across five runs of the experiment each combination of visual and auditory conditions occurs twice, unconfounding the visual and auditory conditions. c) brain regions targeted by the EMFL; "glass brains" show the anatomical constraint parcels used for each contrast. d) accuracy on the 'Does this come next?' task, by auditory condition. *Condition abbreviations: Faces (Fa), Scenes (S), Bodies (B), Objects (O), Words/scrambled (W), False belief (FB), False photo (FP), Non-words (NW), Quilted audio (QLT), Arithmetic (MATH).*

2.2 Experimental data collection

We scanned 20 participants, each over two imaging sessions, on the EMFL as well as a battery of standard functional localizers and anatomical scans. The details of the standard localizers—including targeted fROIs, task, and stimuli—are described in Table 2. The standard visual category localizer we used contains static images of faces, objects, scenes, and scrambled objects, aka "FOSS" (Epstein & Kanwisher, 1998), but not bodies or words. We thus created our own additional localizer for the EBA and VWFA that contained blocks of words, line drawings of bodies and line drawings of objects. Participants were familiarized with the stimuli and task in our EMFL localizer by performing a short practice run outside the scanner before their scanning session.

Participants were recruited from the Massachusetts Institute of Technology and the Greater Boston area (n=20, 10 females. 18-50 years old). All participants provided written informed consent of the protocol approved by the Allen Institute for Brain Science WCB IRB Connexus and were subsequently compensated $75 for each two-hour scan session. Data from five additional participants were excluded from analyses because they did not return for a second scanning session.

Standard localizer	Targeted fROI(s)	Stimuli	Task	Duration
Epstein and Kanwisher (1998)	FFA, OFA, fSTS, PPA, OPA, RSC, LOC	Static images of faces, scenes, objects, and scrambled objects	Button press after one-back presentation	16 min
Fedorenko et al. (2010)	High-level language processing regions	Written English sentences and non-word sequences	Button press after on-screen indication	10
Jacoby et al. (2016)	rTPJ	Written stories of characters experiencing physical or emotional pain	Button press (1–4) rating main character's pain or suffering	10
Fedorenko et al. (2013)	Multiple demand network (frontal and parietal MD)	Sequences of visual spatial working memory grids	Button press choosing correct grid pattern	14
Regev et al., in prep.	Speech processing regions	Spoken non-words and scrambled audio	Button press after on-screen indication	4
Newly developed for this study	EBA, VWFA	Static black & white line drawings of bodies, objects, and words	Button press after one-back presentation	18

Table 2: **Overview of standard functional localizers used here.** Battery of standard functional localizers that have been used across numerous prior studies, as well as two newer localizers used here to validate the EMFL. *fROI abbreviations: fusiform face area (FFA), occipital face area (OFA), face-selective regions of the superior temporal sulcus (fSTS), parahippocampal place area (PPA), occipital place area (OPA), retrosplenial cortex (RSC), lateral occipital complex (LOC), right (rTPJ) and left (ltPJ) temporal parietal junction, right STS (rSTS), medial prefrontal cortex (mPFC), extrastriate body area (EBA), visual word form area (VWFA).*

2.3 fMRI data acquisition and preprocessing

All imaging was performed on a Siemens 3T MAGNETOM Prisma scanner with a 32-channel head coil at the Athinoula A. Martinos Imaging Center at MIT. For each subject, a high-resolution T1-weighted anatomical image (MPRAGE: TR = 2.53 s; TE = 3.57 ms; $\alpha = 9°$; FOV = 256 mm; matrix = 256 $\times$ 256; slice thickness = 1 mm; 176 slices; acceleration factor = 3; 24 reference lines) was collected in addition to whole-brain functional data using a T2-weighted echo planar imaging pulse sequence (TR = 2 s; TE = 30 ms; $\alpha = 90°$; FOV = 208 mm; matrix = 104 $\times$ 104; slice thickness = 2 mm; voxel size = 2 $\times$ 2 mm in-plane; slice gap = 0 mm; 52 slices).

fMRI data were analyzed using the FreeSurfer Functional Analysis Stream (FS-FAST) software (version 6.0.0) and run via custom MATLAB (version 2018b) scripts. The native anatomical space was first reconstructed for each participant using the standard FS-FAST pipeline. Pre-processing of functional data included motion correction, smoothing with a 3 mm FWHM Gaussian kernel, binary brain mask creation, intensity normalization, and registration to the native anatomical space with 6 degrees of freedom.

First-level analysis of functional data was performed in the native space of each participant for each experiment. Conditions were included as covariates of interest along with six nuisance regressors for motion correction in a general linear model (GLM), modeled as a boxcar function convolved with a canonical hemodynamic response function (HRF). Functional contrasts—including t- and F- statistics, significance values, and effect sizes—were then estimated using the fitted GLM in the FS-FAST first-level analysis pipeline, which constructs design and contrast matrices, concatenates all functional runs, fits regression coefficients in the model to the voxel-wise time course, and computes significance values (Fischl (2012); https://surfer.nmr.mgh.harvard.edu/fswiki/FsFast). The following contrasts were computed for the visual stimuli: Faces > Objects, Scenes > Objects, Bodies > Objects, Words/scrambled > Objects, and Objects > Words/scrambled (n.b. words were superimposed on a background of scrambled object videos to maximize stimulus utility). The following contrasts were computed for the auditory stimuli: False belief > False photo, English > Non-words, Non-words > Quilted audio, and Arithmetic > English. The 'English' condition was created by averaging over responses to both the 'False belief' and 'False photo' stimuli but not the 'Arithmetic' stimuli. See Figures 2, 3 for whole-brain activation maps in example participants.

2.4 Definition of functional regions of interest (fROIs)

Functional regions of interest (fROIs) were constrained to lie within previously-published anatomical parcels (see Figure 1C for illustration and Table 3 for details). We first transformed the parcels into each participant's native space, and then selected the top 10% of voxels with strongest significance for the relevant contrast within each parcel. First, we defined these fROIs using all five runs of the EMFL for maximum statistical power. Next, we used the three odd-numbered runs to test whether three runs were sufficient to localize the fROIs and to use held-out data to test the selectivity of the responses.

fROI	Hemi.	Source of parcel	EMFL contrast	Standard contrast	Processing domain
FFA, OFA, fSTS	Right	Julian et al. (2012)	Faces > Objects	Faces > Objects	Faces
PPA, OPA, RSC	Right	Julian et al. (2012)	Scenes > Objects	Scenes > Objects	Scenes
LOC	Right	Julian et al. (2012)	Objects > Words/scrambled	Objects > Scrambled	Non-specific objects
EBA	Right	Julian et al. (2012)	Bodies > Objects	Bodies > Objects	Bodies
VWFA	Left	Saygin et al. (2016)	Words/scrambled > Objects	Words > Objects	Visual word forms
rTPJ	Right	Dufour et al. (2013)	False belief > False photo	Emotional pain > Physical pain	Social cognition (Theory of Mind)
Language network (IFGorb, IFG, MFG, AntTemp, PostTemp, AG)	Left	Fedorenko et al. (2010)	English (false belief + false photo) > Non-words	Sentences > non-words	Language (word retrieval, syntax)
STG	Bilateral	Regev et al., in prep.	Non-words > Quilted audio	Non-words > Degraded speech	Speech (phonemes)
Frontal MD (IFGop, SFG, MFG1, MFG2, MFGorb)	Bilateral	Fedorenko et al. (2013)	Arithmetic > English (false belief + false photo)	Hard > Easy	Domain-general executive demands
Parietal MD (APG, PPG, MPG)	Bilateral	Fedorenko et al. (2013)	Arithmetic > English (false belief + false photo)	Hard > Easy	Domain-general executive demands

Table 3: **fROIs targeted in this study.** Functional regions of interest (fROIs) targeted by the EMFL (first column), source of anatomical parcels used to constrain fROI bounds (second column), functional contrast of stimulus conditions used in the EMFL to define an fROI (third column), and the comparable contrast used in standard localizers to define the same fROI (fourth). Processing domain in which the fROI is involved (fifth). *fROI abbreviations: inferior and middle frontal gyrus (IFG and MFG), orbital region of IFG (IFGorb), anterior and posterior temporal areas (AntTemp and PostTemp), angular gyrus (AG), superior temporal gyrus (STG), opercular region of IFG (IFGop), superior and middle frontal gyrus (SFG, MFG1 and MFG2), orbital region of MFG (MFGorb), anterior, middle, and posterior parietal cortex (APG, MPG, and PPG). See Tables 2, 3 legend for additional abbreviations.*

As noted in our pre-registration, it was not clear in advance whether the Theory of Mind contrast (False belief > False photo) would be stronger when analyzing the full block, or only the second half of the block (which contained more mental state content). We therefore tried both and found that using the full block worked marginally better, and used that for all analyses reported here.

2.5 Estimating fROI selectivity

To test whether the EMFL successfully identified fROIs with the expected selectivities, we first defined each fROI as described above using just the three odd-numbered runs. We then measured the response of this fROI to each of the ten conditions (five visual and five auditory) using the held-out data from the even-numbered runs. This procedure was repeated using the even-numbered runs to select voxels and the odd runs to measure responses in these voxels. The two estimates were then averaged to derive a single value for the response in each subject for each fROI for each condition (see Figure 4).

2.6 Does the EMFL identify fROIs and produce selective responses in these fROIs as effectively as Standard Localizers?

To compare the EMFL with standard localizers—both in their ability to identify selective voxels, and in the selectivity of the response in those voxels—we measured the selectivity of each fROI as a function of (i) whether it was identified with the EMFL or a standard localizer, and (ii) whether its response was measured in the EMFL or a standard localizer. For this analysis, we defined fROIs using a subset of runs from either the EMFL or a standard localizer in individual subjects. Next, we measured the fROI's response in the remaining held-out runs of both localizers to its preferred and non-preferred conditions (e.g. faces and objects for the FFA). Finally, we tested the significance of contrast ('preferred' vs 'non-preferred'), localizer used to define an fROI (EMFL vs. standard), and localizer used to measure its response (EMFL vs. standard) separately for each fROI via a 2x2x2 ANOVA with these three factors. We also analyzed the entire dataset at once by adding a factor for fROI, resulting in an omnibus 2x2x2x14 ANOVA. These analyses enabled us to test whether the strength of the measured contrast in each fROI depended on the localizer used to identify voxels or the localizer used to measure responses or both, and whether the answer to this question varied across fROIs.

2.7 Selectivity of fROIs as a function of fROI Volume

The selectivity of an fROI will naturally depend on the volume of the region selected: a hotspot at the center of each individual's activation is likely to be very selective, but as the volume included in the fROI increases, more voxels with lower selectivity will be included. Thus, an anatomical parcel large enough to accommodate the variable position of fROIs across participants will necessarily include many voxels with weak or no selectivity. (That is why anatomical parcels must be intersected with individual activation maps to adequately define each fROI.) Our next analysis quantifies the dependence of fROI selectivity

on fROI volume using a "growing window" method as described in previous work (Kosakowski et al., 2022).

First, we rank-ordered voxels by their significance value for a given contrast in a subset of functional runs. We then computed the average response magnitude to the localizer conditions in held-out runs in a growing selection of voxels, starting from the most selective voxel and gradually increasing the selection size by including less selective voxels, in steps of one voxel at a time. We performed this procedure on individual participants using data from the EMFL (Figure 6), but also using data across different localizers with the same or similar contrasts. For example, we rank-ordered voxels by the significance of their Faces > Objects contrast during even runs of the EMFL and measured the responses to Faces and Objects conditions, in those same voxels, during odd runs of the FOSS localizer (see Supplemental Figure 1). This alternative analysis also tests whether our EMFL cross-validates to standard localizers by measuring the response magnitude of conditions in the standard localizer in fROIs defined (as described above) by the EMFL.

2.8 Voxel overlap within fROIs measured via Dice coefficient

To determine whether EMFL and standard localizers identify similar voxels as belonging to each fROI, we calculated the Dice coefficients between the EMFL and the standard localizers to quantify the degree to which they spatially overlap. First we defined each fROI separately for each individual, specifically the top 10% of voxels within the relevant parcel based on the relevant functional contrast, separately for the EMFL and standard localizers. We then calculated the Dice coefficient of the overlap both within and between localizers. Dice coefficients range between 0 (meaning no overlap) and 1 (perfect overlap). Dice coefficients within a localizer necessarily required splitting the data into even and odd runs, so for a comparison with equivalent power, we calculated between-localizer Dice coefficients using the same even-odd splits of the data. In addition, to quantify the between-localizer Dice coefficients with maximum power, we also calculated overlap using all runs of the data for each localizer.

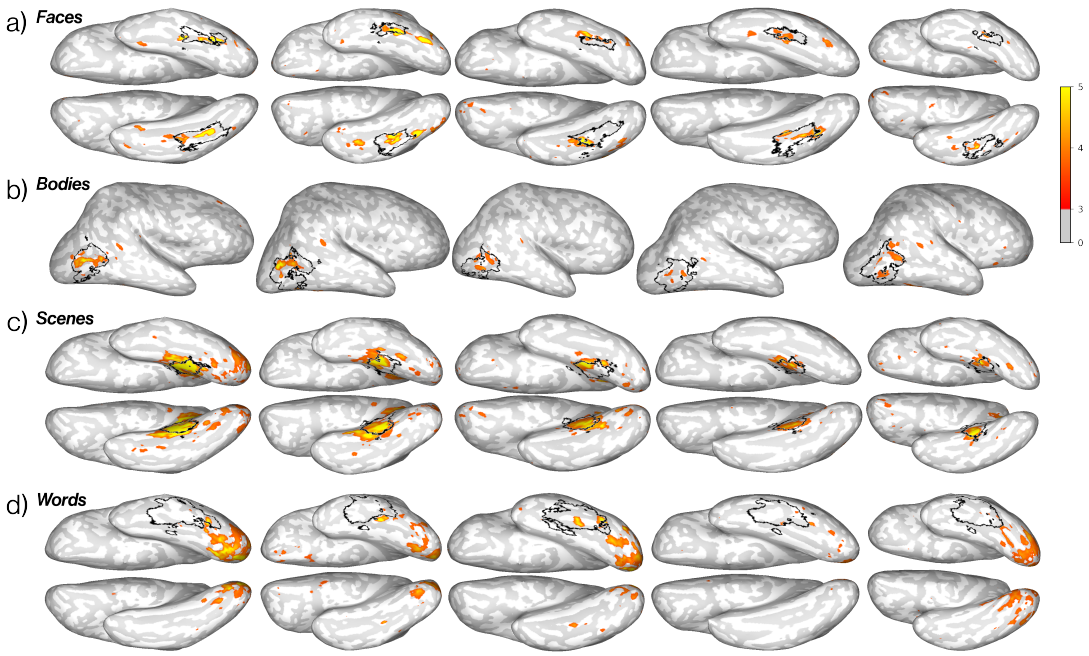


Figure 2: **EMFL captures expected functional topography in visually selective regions.** Brain activation maps of four visual contrasts in the same five participants using data from all five runs of the EMFL. Each brain shows the functional contrast for a given condition as $-log(p\text{-}value) * sgn(t\text{-}test)$, thresholded at $+3$. Surface projections of relevant anatomical constraint parcels are outlined in black. Contrasts: (a) Faces $>$ Objects, (b) Scenes $>$ Objects, (c) Bodies $>$ Objects, and (d) Words $>$ Objects. Although fROI-based analyses of these regions were conducted in only one hemisphere per region (see Table 3), ventral surface views here show activations in both hemispheres. Analyses were conducted in the volume and are here projected to the surface for visualization only.

3 Results

3.1 EMFL data reveal the expected activations for both visual and auditory contrasts in most individual participants

First, as an initial reality check on the ability of the EMFL to produce the expected activations, we analyzed the fMRI BOLD responses from each participant with each of the fROI-defining contrasts (see Table 3) using the full set of five runs. Whole-brain activation maps for category-selective contrasts for faces, places, bodies, and words (all compared to objects) are shown for five randomly-selected participants, along with the relevant anatomical constraint parcels, in Figure 2. Visual inspection reveals activations that qualitatively resemble those in the extensive literature on these regions. This finding shows that the simultaneous performance of a relatively demanding task on auditory stimuli does not noticeably compromise the expected visual category-selective activations. Further, for our various contrasts of auditory conditions, whole-brain activation maps for the same five participants reveal the expected activations in most participants as shown in Figure 3. Comparisons of activations for EMFL and standard localizers within the same participants are shown for Theory of Mind and Speech contrasts in Supplemental Figure 2 and for Language in Supplemental Figure 3.

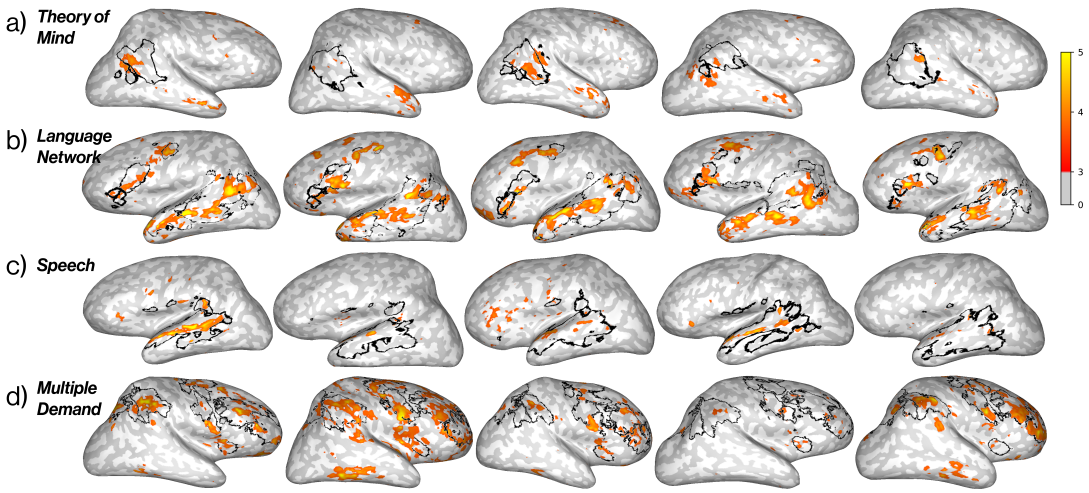


Figure 3: **EMFL captures expected functional topography in speech and higher-level processing regions.** Brain activation maps of five auditory contrasts in five representative subjects after five runs of the EMFL, plotted in the individual's native space. Each brain shows the functional contrast for a given condition as $-log(p\text{-}value) * sgn(t\text{-}test)$, thresholded above +3. Surface projections of relevant anatomical parcels are outlined in black. Contrasts: (a) False belief > False photo, (b) English > Non-words, (c) Non-words > Quilted audio, and (d) Arithmetic > English. Analyses were conducted in the volume and are here projected to the surface for visualization only.

3.2 How many participants show each fROI individually?

We next asked how effective EMFL is by asking in how many participants it identifies each fROI using the relatively stringent significance threshold of $p < 0.001$ uncorrected and a requirement of at least 10 voxels in the relevant anatomical parcel for each fROI. By this measure, the EMFL identifies the main fROIs in at least 18/20 participants when all five runs are used (see Table 4); exceptions are the rOFA, rSTS, and VWFA, with fROIs successfully defined by this measure in only 12-14 out of 20 participants (fewer subjects than in the standard localizers for these regions). However, many fMRI studies now define fROIs not using a hard statistical threshold, but instead by choosing the top 10% of most-significant voxels within an anatomical constraint parcel, whether or not these voxels reach a fixed significance level. This method has the advantage of including all participants in an analysis. Indeed, as shown in Supplementary Figure 4, even in participants who did not have 10 voxels teaching the $p < 0.001$ significance level for the weakest fROIs (STS, OFA, and VWFA), most of these participants nonetheless showed the expected selective response profile in held-out data when their fROIs were defined as the top 10% of voxels within the parcel.

Can the main fROIs be defined via the hard statistical threshold with even less scan time? Indeed, as shown in Table 4, just three runs of the EMFL (14 minutes of data collection) was still sufficient to identify most fROIs in most participants. Taken together, we conclude that three runs (14 minutes) are sufficient to define most regions in almost all subjects using the hard significance criterion, but five runs are recommended if localization of VWFA, rOFA, or rSTS are important for the study.

3.3 The EMFL Effectively Identifies Voxels Selectively Responsive to the Predicted Condition(s)

Our primary question is whether the EMFL successfully identifies fROIs that reveal the expected functional selectivities in held-out data. To answer this question we identified each fROI as the top 10% of voxels within the relevant anatomical parcel in a subset of EMFL runs, and quantified responses using the held-out data from the other EMFL runs, separately for each participant. Figure 4 shows the response to each of the ten conditions in EMFL in each of fourteen well-established fROIs, showing the expected selective responses in each fROI. Results for left hemisphere fROIs are shown in Supplemental Figure 7 and for individual language and MD regions in Supplemental Figure 6) These findings are supported by statistical tests in Supplementary Table 1, which show (with very few exceptions) that the response to the preferred condition was significantly greater than to each of the non-preferred conditions in each fROI.

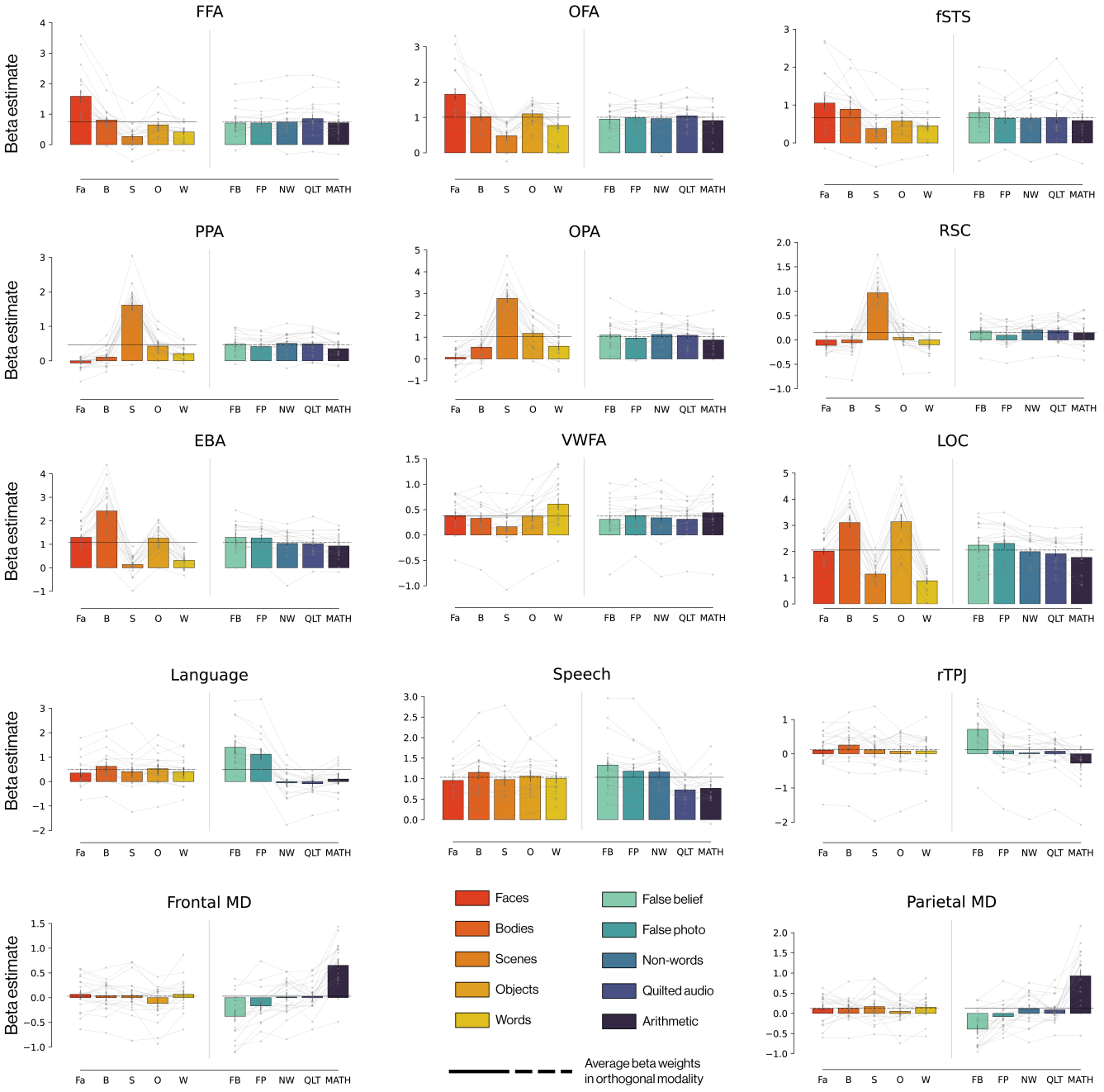


Figure 4: **EMFL-defined fROIs are functionally specific for their preferred condition.** Response of each functional region of interest defined by EMFL to each of the ten conditions in EMFL. We selected voxels by their functional selectivity during even numbered runs (run 2, 4) and measured their responses in three held-out runs (run 1, 3, 5) and vice versa; bar charts here show the average response across the two even-odd splits of the data, averaged across the 20 participants. Horizontal lines indicate the average response across the five conditions in the modality used to define the region (solid black portion), used as the baseline (dotted black portion) against which to consider responses in the modality not used to define the region. *See Figure 1 and Tables 2, 3 legend for condition and ROI abbreviations.*

	EMFL				**Standard localizers**			
Runs	odd		all		odd		all	
Hemisphere	left	right	left	right	left	right	left	right
FFA	17	20	18	20	17	18	19	19
OFA	4	8	5	12	9	16	11	20
STS	4	12	2	14	1	16	2	17
PPA	18	18	19	18	16	16	17	17
OPA	17	18	19	20	15	17	19	19
RSC	16	18	18	18	11	15	16	17
EBA	17	18	17	18	18	19	18	20
VWFA	11		14		17		18	
LOC	18	20	20	20	20	19	20	20
RTPJ		8		17		11		17
Speech	20	20	20	20	19	19	20	20
Language network (1)	17		19		8		12	
Language network (2)	19		19		14		18	
Language network (3)	19		20		14		16	
Language network (4)	20		20		17		18	
Language network (5)	20		20		18		19	
Language network (6)	17		18		4		8	
Frontal MD	19	20	20	20	18	16	19	19
Parietal MD	20	19	20	20	16	18	20	20

Table 4: **Number of participants (out of 20) who show at least ten voxels that are significant for the relevant functional contrast ($p < 0.001$) within the relevant anatomical constraint parcel.** Contrast significance was computed using either all five runs, or just the odd-numbered runs of either the EMFL or the relevant standard localizer. Empty cells indicate fROIs that are not expected to be robust in the hemisphere indicated (e.g. the VWFA in the right hemisphere).

One exception was the contrast of faces versus bodies, which did not reach significance in the fSTS ($p = 0.07$). This result was expected due to the implied presence of social interaction in videos of bodies, and the role of fSTS in social information processing. Notably, the contrast of visual faces over auditory false belief and false photo stories reached only modest significance in the fSTS ($p = 0.04$ and $p = 0.005$, respectively), perhaps reflecting amodal processing and/or interference in the STS produced by the required mentalization during these stories.

The other exception was the speech region, which did not respond more to speech than to the visual body condition. This result can be explained by the fact that most of the auditory conditions contained speech sounds, and the average response of all of these conditions served as the baseline for the visual responses

(because the simultaneous presentation of visual and auditory conditions resulted in the response of each condition in one modality including the average of all the conditions in the opposite modality). For the same reason we did not expect the response measured for the irrelevant modality (e.g., the response to auditory stimuli for visual category-selective selective regions) to be zero: each auditory condition occurs simultaneously with visual conditions, and vice versa. The average response across all preferred-modality conditions for each fROI (solid black line in Figure 4) thus serves as the relevant baseline for the non-preferred modality conditions (dashed black line in Figure 4) in that fROI.

In sum, as shown in Figure 4, three runs of the EMFL localizer is sufficient to identify each of the standard fROIs, which show the expected selective response profile in the two held-out runs. Supplemental Figure 7 shows that the same is true when the first three runs as used to identify the fROI, and the last two to quantify its response, showing that just three runs suffices to identify fROIs.

3.4 Quantifying the Similarity of fROIs obtained from EMFL and Standard Localizers

Are the same voxels picked out by the EMFL and standard localizers? A qualitative visualization of the activation overlap between EMFL and standard localizers for face, place, and language-selective actions is shown in Figure 5A. To quantify this overlap, we calculated the Dice coefficients between the top 10% most selective voxels for each region identified by the EMFL versus the standard localizer, using all data in each case (Figure 5B). All Dice coefficients were significantly greater than chance (all p-values < 0.001). We then asked whether Dice coefficients were higher within a localizer than between localizers, using equivalent splits of the data within each localizer (see Methods), and found significantly higher overlap between runs of the same localizer compared to the overlap of runs between different localizers.

We also measured the similarity of the relevant stimulus contrast across voxels within the relevant parcel for each fROI via Pearson's correlation coefficient (Supplementary Figure 8). The pattern of response across voxels for the localizer contrast within the relevant parcel was significantly correlated between EMFL and standard localizers for all fROIs (all p-values < 0.001). However, the correlation between the pattern of response across voxels was significantly higher between splits of the data within a localizer than between localizers for all fROIs, with an exception for the fSTS.

Thus, both Dice and correlation measures show that fROIs identified with EMFL and standard localizers are similar, showing generalization across the many differences between EMFL and the standard localizers (static images versus movies, presence of stimuli in other modality, etc).

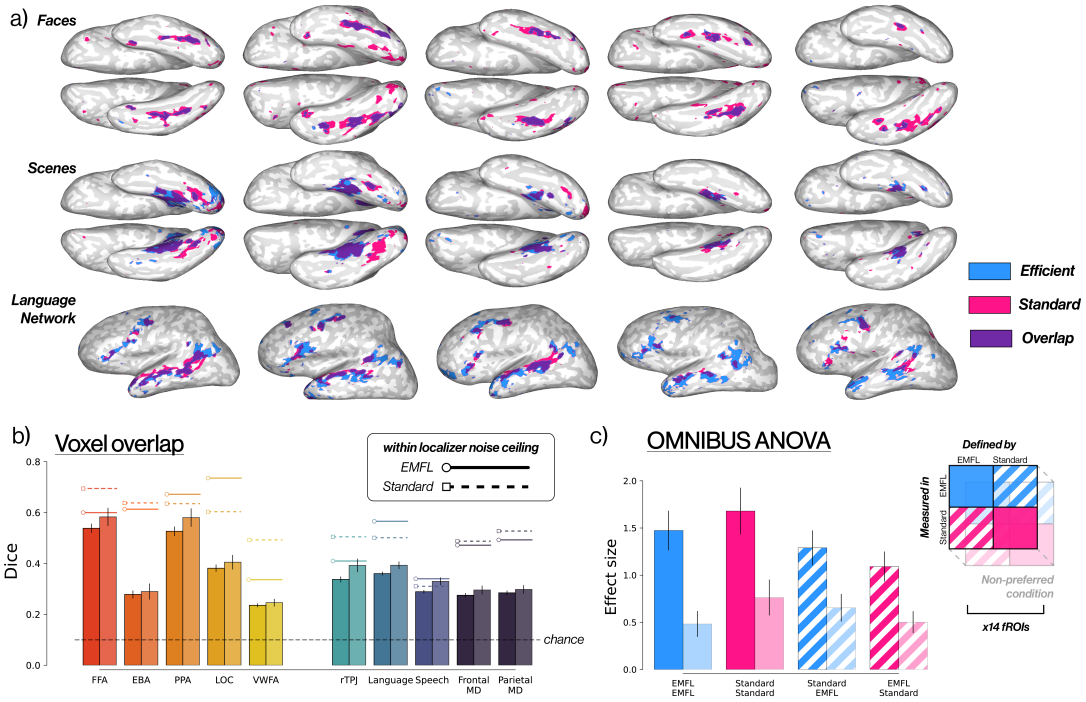


Figure 5: **EMFL and standard localizers produce similar brain activity.** (a) brain activation maps of three different conditions during EMFL and a standard localizer, plotted in the individual's native space (same five subjects as Figure 2). Voxels are colored if they reached a significance threshold of $p < 0.001$ in EMFL (cyan), a standard localizer (magenta), or both (violet) for a given functional contrast. (top) Faces > objects by EMFL and Epstein and Kanwisher (1998), (middle) scenes > objects by EMFL and Epstein and Kanwisher (1998), (bottom) English > non-words in EMFL and sentences > non-words in Fedorenko et al. (2010) (b) Dice coefficient of the top 10% most-selective voxels identified by the EMFL and standard localizers for each region, calculated within an anatomical parcel. Dark bars average across four even-odd splits, light bars use all runs available. Floating bars represent the Dice coefficient within even and odd runs of the same localizer. Dashed line indicates overlap by chance. (c) results of the 2x2x2x14 omnibus ANOVA. Response magnitudes are averaged across fROIs for preferred (bright) versus non-preferred (pale) conditions, measured in the EMFL (blue) or standard (pink) localizer and defined by either the same localizer (solid) or the other localizer (striped). See Tables 2, 3 legend for ROI abbreviations.

However, unsurprisingly, the various pairs of localizers do not produce exactly the same pattern of response for each contrast. These differences between localizers remind us that the response of voxels in these regions is not solely determined by the presence versus absence of the preferred stimulus or task.

The lowest similarities between EMFL and standard localizers were found for word and body selective regions. We speculate that this dissimilarity resulted from the dissimilar depiction of word and body stimuli in the EMFL (colored movies of bodies and words on dynamic scrambled background) and the standard localizer (static images of line drawings and text on a white background). Consistent with this interpretation, pilot data from 5 participants who were brought back and scanned on a more standard localizer for body-selective regions, using photographs instead of line drawings, showed a higher Dice coefficient with EMFL than did our previous less-standard localizer based on line drawings.

Similarly, the relatively low correlation between the EMFL Theory of Mind contrast and the "standard" localizer presumably reflects the fact that the standard localizer in this case was a contrast of reading stories about emotional pain versus physical pain, whereas the EMFL localizer uses the classic contrast of false belief $>$ false photo, which isolates similar but not identical mental processes (Jacoby et al., 2016). This nonidentical localizer was used because the standard false belief localizer is less effective when repeated within the same participant.

3.5 The Selectivity of Response Does not Depend on the Localizer used to Define a fROI, but does depend (for some fROIs) on which Localizer is used to Measure Responses

To test whether the selectivity of response of each fROI depends on either (i) the localizer used to identify the fROI or (ii) the localizer used to measure responses in that fROI, we performed an omnibus 2x2x2x14 ANOVA across all participants, with factors for contrast (preferred versus non-preferred condition), localizer used to define the fROI (EMFL versus standard), localizer used to measure responses (EMFL versus standard), and parcel (the 14 main functional regions of interest). The results of this ANOVA are shown in Table 5 (see Figure 5c for summary). The key finding is that the selectivity measured ("contrast") does not depend on whether the region is defined with EMFL or standard localizers ("definer"), i.e. neither the interaction of contrast x definer nor contrast x definer x fROI reached significance. However, for some fROIs, the strength of the contrast measured does depend on whether responses are measured using the EMFL or standard localizers, as reflected in a significant interaction of contrast x measurer x parcel ($p = 2 * 10^{-6}$). We therefore performed 2x2x2 ANOVAs separately within each of the 14 fROIs, which are shown in Supplemental Figure 9. These analyses show that although the EMFL and standard

localizers are similarly effective for defining all fROIs (the contrast x definer interaction did not reach significance for any of the fROIs), the EMFL produces more selective responses for some fROIs including the PPA and RSC (supported by significant interactions of contrast by measurer, $p < 0.0001$ for each) - likely due to the optimization procedure used to select the particular scene stimuli used in this study (Chen et al., 2024). In only one fROI, the OFA, did the EMFL produce a less selective response than did the standard localizer (interaction of contrast x measurer $p = 0.02$). In sum, these analyses show that the EMFL identifies fROIs as effectively as the standard localizers do, and (with the one exception of the OFA) produces responses in those fROIs that are as selective or more selective than the standard localizers.

	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1	370.0	370.0	990.0	1.5e-178
C(definer)	1	28.0	28.0	76.0	5.4e-18
C(measurer)	1	0.44	0.44	1.2	0.27
C(parcel)	13	580.0	45.0	120.0	2.5e-244
C(contrast):C(definer)	1	0.012	0.012	0.032	0.86
C(contrast):C(measurer)	1	0.33	0.33	0.89	0.35
C(definer):C(measurer)	1	26.0	26.0	69.0	1.4e-16
C(contrast):C(parcel)	13	45.0	3.4	9.3	5.1e-19
C(definer):C(parcel)	13	27.0	2.1	5.6	3.7e-10
C(measurer):C(parcel)	13	170.0	13.0	35.0	2.5e-80
C(contrast):C(definer):C(measurer)	1	18.0	18.0	48.0	6.9e-12
C(contrast):C(definer):C(parcel)	13	1.8	0.14	0.37	0.98
C(contrast):C(measurer):C(parcel)	13	19.0	1.5	4.0	2e-06
C(definer):C(measurer):C(parcel)	13	22.0	1.7	4.6	6e-08
C(contrast):C(definer):C(measurer):C(parcel)	13	8.3	0.64	1.7	0.05
Residual	2100	780.0	0.37		

Table 5: **Summary of 2x2x2x14 omnibus ANOVA.** With main effects for contrast (preferred vs non-preferred), localizer used to define ROIs (definer, EMFL vs standard), localizer used to measure responses (measurer, EMFL vs standard), and fROI (parcel), as well as their interactions.

3.6 Characterizing the Selectivity of each fROI as a Function of fROI Volume

The ultimate test of a functional localizer is its ability to identify voxels that produce the intended selectivity in held-out data. However, the exact number of voxels which reach a fixed significance threshold for a given contrast can vary considerably across participants. As noted above, a standard method for addressing this problem is to define the fROIs as the top 10% most-significant voxels within a contrast-relevant constraint parcel, regardless of whether those voxels meet a fixed significance threshold. Here, we provide support for that approach by showing how the selectivity in held-out data decreases

as less-significant voxels are included in the fROI. We quantify in Figure 6 this drop-off in selectivity of each fROI as the volume of that fROI increases. As expected, fROIs show their strongest selectivity in the top 10 percent of voxels, and most fROIs still show a twofold response to their preferred condition compared to any non-preferred condition, though a few show more modest patterns of selectivity. For all fROIs, if 100% of the parcel is included then selectivity is weak (which is why functional localizers are needed!). In a similar analysis, we further cross-validate the selectivity of each fROI by using the EMFL to define voxels and then measure their responses in a standard localizer (Supplemental Figure 1).

4 Discussion

The discovery that the human brain contains a set of cortical regions with distinctive functional response profiles, each present in virtually every neurotypical individual, invites research into the developmental and evolutionary origins, as well as the representations, computations and behavioral relevance of these regions. Yet the individual variability in the precise anatomical location of these regions means that they can only be studied precisely if they are first functionally defined in each participant individually. Here, we present and evaluate an efficient functional localizer that reliably identifies, in just 14 minutes of fMRI scan time, brain regions in each individual that are selectively engaged in the visual processing of faces, scenes, bodies, objects, and words, as well as regions engaged in speech perception, language understanding, theory of mind, and domain-general cognitive demand. We further cross-validate this localizer with the corresponding established localizers by showing substantial topographic overlap and similar expected response selectivity in held-out data. We find that the selectivity of these regions in held-out data is as strong when they are defined using our new efficient multifunction localizer (EMFL) as they are when defined by standard established localizers. The EMFL will be made available on GitHub (see Methods for the repository link) and should be widely applicable in human cognitive neuroscience research, enabling multiple regions to be studied in each experiment, and making possible a cumulative research program in which studies across participants and labs can refer to the same regions.

Our results further reinforce the functional selectivity of each region, not just with respect to the standard conditions it is often contrasted with (e.g., faces versus objects for the FFA), but with respect to a broader set of conditions spanning vision, hearing, social cognition, and domain-general demand. Indeed, almost all of our fROIs showed a significantly higher response to their preferred condition than to each of the other non-preferred conditions. (One exception was the fSTS, for which the higher response to visually-presented faces over bodies did not reach significance, and speech regions, which did not respond more to speech than some of the visual conditions presumably because speech sounds are present in almost

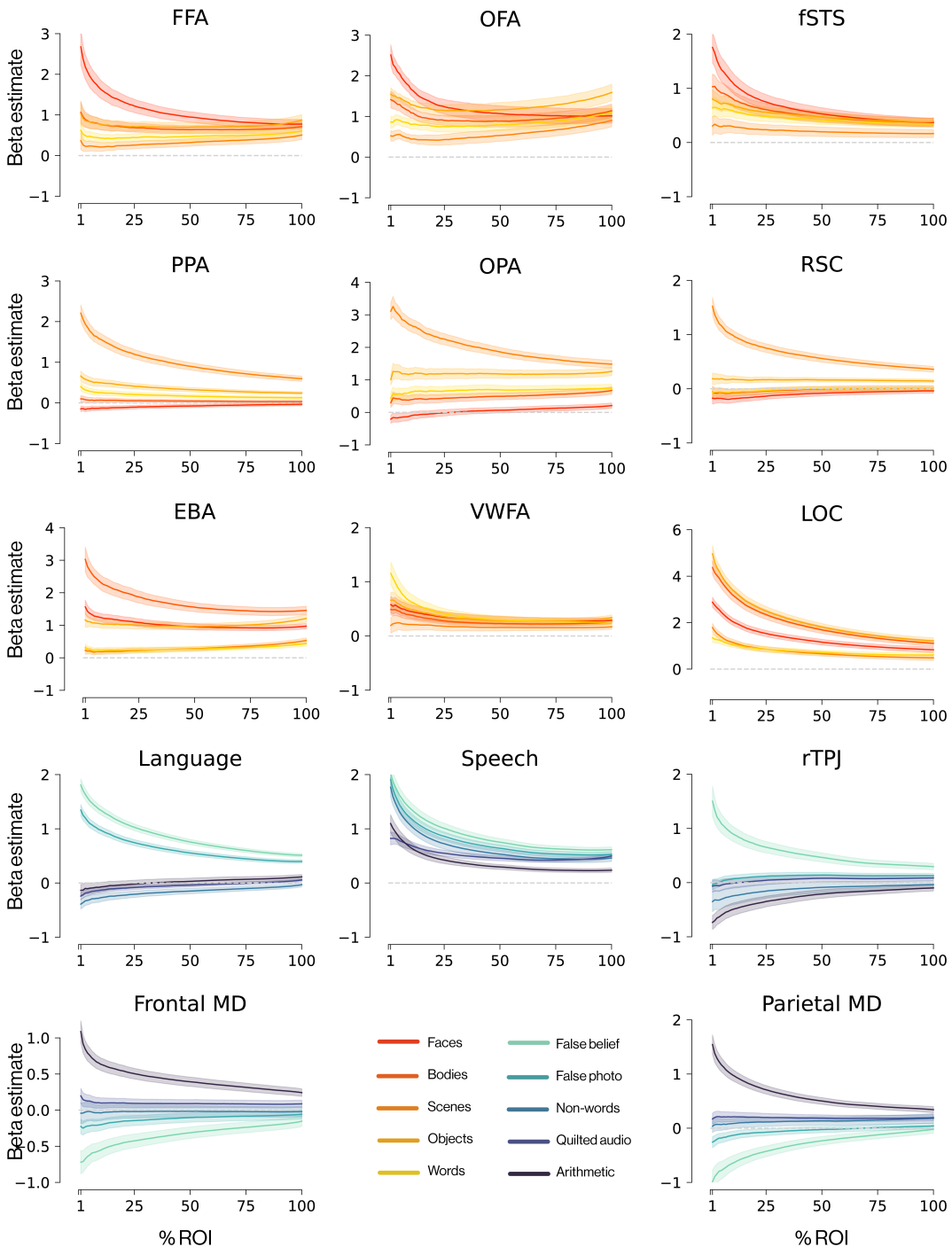


Figure 6: **Selectivity of EMFL-defined fROIs decreases as a function of fROI size.** Response of each functional region of interest defined by and measured during EMFL. We selected voxels by their functional selectivity during the first three runs and measured their beta estimates during the final two held-out runs. Response magnitudes (y-axis) are plotted as a function of fROI size (x-axis, % of voxels in the parcel that are included in the fROI)

all visual blocks). While the signature response profiles of these regions have long been widely reported (and in some cases, extended to animal physiology work; e.g. Tsao et al., 2006), recent discussion oddly seems to question the existence or significance of functional specificity in the brain (e.g. Pessoa, 2023; Westlin et al., 2023). Our data, paradigm and analysis tools are publicly available, enabling anyone to easily replicate these findings for themselves.

Beyond demonstrating the efficacy of an efficient localizer and showing the functional selectivity of the regions it identifies across a broad swath of stimuli and tasks, several findings from our study are notable. First, we found strong selectivity for visual conditions even when visual stimuli were task-irrelevant and participants were engaged in an unrelated and demanding auditory task. Although attentional modulation of category-selective regions is well established (e.g. O'Craven et al., 1999; Wojciulik et al., 1998), few if any prior studies have shown as we do here that strongly selective responses nonetheless remain for unattended visual stimuli. Additional pilot data (shown in Supplemental Figure 10) from five participants who were brought back to perform the EMFL experiment without the concurrent auditory stimuli or tasks showed no clear increase in selectivity of category-selective regions (see also Bugatus et al., 2017; Harel et al., 2014; Mur et al., 2025). Second, the face-selective region in the STS did not respond significantly to speech sounds, in contrast to prior reports (Deen et al., 2015). However, task-relevant mentalization produced by auditory false belief and false photo stories modulated their contrast significance against visually-presented faces. Conversely speech-selective regions did not respond more to faces than other stimuli, suggesting a segregation of these two responses in the STS. Third, the multiple demand regions responded strongly to auditory arithmetic but barely above a fixation baseline for the other auditory contrasts including speech perception, language understanding, and theory of mind, even though these tasks were at least as difficult as the auditory arithmetic task (Figure 1D). This finding shows that although multiple demand regions respond across a wide range of cognitive tasks, this does not include speech perception, language understanding, or theory of mind (Fedorenko & Shain, 2021; Saxe et al., 2006).

Of course functionally distinct cortical regions do not act alone, but receive input from, send outputs to, and presumably interact online with numerous other areas. Indeed, analyses of the correlations of the fMRI signal at rest (or during task performance) have identified "functional networks" spanning many regions. Although these fMRI correlations between regions do not necessarily reflect structural connectivity, they do reliably identify groups of regions that apparently work together (Yeo et al., 2011). This macroscopic level of analysis of brain-wide networks is complementary to and synergistic with more fine-grained analyses of individual cortical regions; both levels of analysis will be important for understanding brain function. Further, resting functional correlation analyses have also been shown to allow efficient functional localization of some functionally distinct brain regions, as demonstrated by

cross-validation with standard task-based functional localizers (Braga et al., 2020; Du et al., 2024). Language regions can even be accurately identified from functional correlation analyses of single runs of resting state data (Du et al., 2025; Shain & Fedorenko, 2025).

Other studies have proposed efficient methods for localizing functional regions using task and movie data. Lee et al. (2024) demonstrated the ability to localize language regions in just 3.5 minutes of scan time while maintaining the fidelity of its spatial layout and functional response profile. Tuckute et al. (2024) found similar results with their 3.5 minute localizer based on speeded reading, which additionally showed a functional dissociation of the language network from the topographically-nearby multiple demand network. The hyperalignment method introduced by Haxby et al. (2011) can also be used to identify functionally specific regions from fMRI data while participants watch movies (Jiahui et al., 2020). Finally, one recent study found that many high-level functionally specific visual areas could be accurately identified from a ten-minute resting functional scan (Molloy et al., 2024). While all of these methods are useful, our functional localizer is the only one with broad coverage, identifying the mostly widely studied perceptual and cognitive regions, in as little as 14 minutes of scan time. As such, we hope to encourage researchers to include this localizer in a wide range of fMRI studies, enabling them to complement other planned analyses with a quantification of their effects within the main established fROIs.

Future research should compare the efficacy of EMFL with other methods such as individual-subject rs-fMRI methods (Braga et al., 2020; Du et al., 2025; Shain & Fedorenko, 2025), connectivity fingerprints (Molloy et al., 2024), group-based fROIs (see Figure 2D, also Saxe et al., 2006), multimodal anatomical parcels (Glasser et al., 2016), and movie-based hyperalignment (Jiahui et al., 2020).

Despite its advantages, our localizer has some limitations. First, the *'Does this come next?'* task feels at first counterintuitive and demanding, as participants must perform an auditory task while salient and distracting but irrelevant visual movies are presented. However, most participants quickly realized that they can in fact successfully selectively attend to the auditory stimuli, performing above chance in every condition. While very young, very old, or cognitively impaired individuals may not be able to perform our task adequately, our pilot data from three high-functioning elderly participants indicate that they had no trouble with the paradigm, and strong activations were obtained. A second limitation is that our localizer is less effective for identifying some fROIs like the VWFA and OFA, so studies focusing on these regions may wish to use localizers optimized for identifying these regions. Nonetheless, these regions—when defined by the EMFL—show similar profiles of selectivity as those found in recent studies (Li et al., 2024). Finally, EMFL does not localize all cortical regions implicated in distinct mental functions. Functional contrasts not targeted in this localizer include visual perception of color (Lafer-Sousa et al.,

2016), motion (Tootell et al., 1995), hands (Bracci et al., 2010), tools (Bracci et al., 2012), and third-party social interactions (Isik et al., 2017); auditory perception of music (Norman-Haignere et al., 2015) or pitch (Norman-Haignere et al., 2013); and motor functions like reaching, grasping, and eye movements (Gallivan & Culham, 2015), episodic projection (DiNicola et al., 2020), and intuitive physical reasoning (Fischer et al., 2016). Note however that because we provide the code and stimuli for the experiment, it would be straightforward to sub in alternate stimulus conditions or tasks to localize other functional regions of interest.

In sum, we show here that fourteen of the most widely studied function regions of interest can be functionally localized in each participant individually in as little as 14 minutes of scan time. Our method cross-validates with standard localizers for individual functional contrasts, and shows selectivity in held-out data as strong as found for standard localizers that take four times as long to run. Our tasks, stimuli, and analysis methods are available online, so anyone can easily add this short localizer to an existing scanning protocol, enabling them to test how these established fROIs respond in their paradigm. In this fashion we hope to help build a cumulative research program across participants, methods, and labs.

Author contributions

AIM: conceptualization, methodology, software, writing, visualization, validation. SH: conceptualization, methodology, software, writing, visualization, validation. EF: conceptualization, editing. RRS: conceptualization, editing. FSK: methodology. TIR: methodology. EMC: methodology. NGK: conceptualization, methodology, software, writing, visualization, validation, supervision.

Funding information

This research was funded by NIH Project Number: UM1 MH130981-01, "Functionally guided adult whole brain cell atlas in human and NHP", PI: Ed Lein.

Declaration of competing interests

The authors declare no competing interests.

Data & code availability

Scripts for the EMFL and relevant analyses are available at these GitHub repositories: https://github.mit.edu/kanlab/efficient (localizer) and https://github.com/aimarvi/emfl (analysis). fMRI data are available on OpenNeuro: https://openneuro.org/datasets/ds006179/versions/1.0.1.

References

Bracci, S., Cavina-Pratesi, C., Ietswaart, M., Caramazza, A., & Peelen, M. V. (2012). Closely overlapping responses to tools and hands in left lateral occipitotemporal cortex. *J Neurophysiol*, *107*, 1443–1456. https://doi.org/10.1152/jn.00619.2011.-The

Bracci, S., Ietswaart, M., Peelen, M. V., & Cavina-Pratesi, C. (2010). Dissociable neural responses to hands and non-hand body parts in human left extrastriate visual cortex. *Journal of Neurophysiology*, *103*, 3389–3397. https://doi.org/10.1152/jn.00215.2010

Braga, R. M., DiNicola, L. M., Becker, H. C., & Buckner, R. L. (2020). Situating the left-lateralized language network in the broader organization of multiple specialized large-scale distributed networks. *Journal of Neurophysiology*, *124*, 1415–1448. https://doi.org/10.1152/jn.00753.2019

Bugatus, L., Weiner, K. S., & Grill-Spector, K. (2017). Task alters category representations in prefrontal but not high-level visual cortex. *NeuroImage*, *155*, 437–449. https://doi.org/10.1016/j.neuroimage.2017.03.062

Chen, E. M., Kamps, F. S., & Saxe, R. R. (2024). An open dataset of functional mri responses to egocentric navigation through natural scenes. *Cognitive Computational Neuroscience*.

Deen, B., Koldewyn, K., Kanwisher, N., & Saxe, R. (2015). Functional organization of social perception and cognition in the superior temporal sulcus. *Cerebral Cortex*, *25*, 4596–4609. https://doi.org/10.1093/cercor/bhv111

DiNicola, L. M., Braga, R. M., & Buckner, R. L. (2020). Parallel distributed networks dissociate episodic and social functions within the individual. *Journal of Neurophysiology*, *123*, 1144–1179. https://doi.org/10.1152/jn.00529.2019

Du, J., DiNicola, L. M., Angeli, P. A., Saadon-Grosman, N., Sun, W., Kaiser, S., Ladopoulou, J., Xue, A., Yeo, B. T. T., Eldaief, M. C., & Buckner, R. L. (2024). Organization of the human cerebral cortex estimated within individuals: Networks, global topography, and function. *Journal of Neurophysiology*, *131*, 1014–1082. https://doi.org/10.1152/jn.00308.2023

Du, J., DiNicola, L. M., Angeli, P. A., Saadon-Grosman, N., Sun, W., Kaiser, S., Ladopoulou, J., Xue, A., Yeo, B. T., Eldaief, M. C., & Buckner, R. L. (2025). Within-individual organization of the

493 human cerebral cortex: Networks, global topography, and function. *bioRxiv*. https://doi.org/10.
494 1101/2023.08.08.552437
495 Dufour, N., Redcay, E., Young, L., Mavros, P. L., Moran, J. M., Triantafyllou, C., Gabrieli, J. D. E., &
496 Saxe, R. (2013). Similar brain activation during false belief tasks in a large sample of adults with
497 and without autism. *PLoS ONE*, *8*, e75468. https://doi.org/10.1371/journal.pone.0075468
498 Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*,
499 *392*, 598–601. https://doi.org/10.1038/33402
500 Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal
501 and parietal cortex. *Proceedings of the National Academy of Sciences of the United States of*
502 *America*, *110*, 16616–16621. https://doi.org/10.1073/pnas.1315235110
503 Fedorenko, E., Hsieh, P.-J., Nieto-Castañón, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New
504 method for fmri investigations of language: Defining rois functionally in individual subjects. *Journal*
505 *of Neurophysiology*, *104*, 1177–1194. https://doi.org/10.1152/jn.00032.2010
506 Fedorenko, E., & Shain, C. (2021). Similarity of computations across domains does not imply shared im-
507 plementation: The case of language comprehension. *Current Directions in Psychological Science*,
508 *30*, 526–534. https://doi.org/10.1177/09637214211046955
509 Fischer, J., Mikhael, J. G., Tenenbaum, J. B., & Kanwisher, N. (2016). Functional neuroanatomy of
510 intuitive physical inference. *Proceedings of the National Academy of Sciences*, *113*. https://doi.
511 org/10.1073/pnas.1610344113
512 Fischl, B. (2012). Freesurfer. *NeuroImage*, *62*, 774–781. https://doi.org/10.1016/j.neuroimage.2012.
513 01.021
514 Gallivan, J. P., & Culham, J. C. (2015). Neural coding within human brain areas involved in actions.
515 *Current Opinion in Neurobiology*, *33*, 141–149. https://doi.org/10.1016/j.conb.2015.03.012
516 Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., Ugurbil, K.,
517 Andersson, J., Beckmann, C. F., Jenkinson, M., Smith, S. M., & Essen, D. C. V. (2016). A
518 multi-modal parcellation of human cerebral cortex. *Nature*, *536*, 171–178. https://doi.org/10.
519 1038/nature18933
520 Harel, A., Kravitz, D. J., & Baker, C. I. (2014). Task context impacts visual object processing differentially
521 across the cortex. *Proceedings of the National Academy of Sciences*, *111*. https://doi.org/10.
522 1073/pnas.1312567111
523 Haxby, J. V., Guntupalli, J. S., Connolly, A. C., Halchenko, Y. O., Conroy, B. R., Gobbini, M. I., Hanke,
524 M., & Ramadge, P. J. (2011). A common, high-dimensional model of the representational space
525 in human ventral temporal cortex. *Neuron*, *72*, 404–416. https://doi.org/10.1016/j.neuron.2011.
526 08.026

Isik, L., Koldewyn, K., Beeler, D., & Kanwisher, N. (2017). Perceiving social interactions in the posterior superior temporal sulcus. *Proceedings of the National Academy of Sciences*, *114*. https://doi.org/10.1073/pnas.1714471114

Jacoby, N., Bruneau, E., Koster-Hale, J., & Saxe, R. (2016). Localizing pain matrix and theory of mind networks with both verbal and non-verbal stimuli. *NeuroImage*, *126*, 39–48. https://doi.org/10.1016/j.neuroimage.2015.11.025

Jiahui, G., Feilong, M., di Oleggio Castello, M. V., Guntupalli, J. S., Chauhan, V., Haxby, J. V., & Gobbini, M. I. (2020). Predicting individual face-selective topography using naturalistic stimuli. *NeuroImage*, *216*. https://doi.org/10.1016/j.neuroimage.2019.116458

Julian, J., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *NeuroImage*, *60*, 2357–2364. https://doi.org/10.1016/j.neuroimage.2012.02.055

Kosakowski, H. L., Cohen, M. A., Takahashi, A., Keil, B., Kanwisher, N., & Saxe, R. (2022). Selective responses to faces, scenes, and bodies in the ventral visual pathway of infants. *Current Biology*, *32*, 265–274.e5. https://doi.org/10.1016/j.cub.2021.10.064

Lafer-Sousa, R., Conway, B. R., & Kanwisher, N. G. (2016). Color-biased regions of the ventral visual pathway lie between face- and place-selective regions in humans, as in macaques. *The Journal of Neuroscience*, *36*, 1682–1697. https://doi.org/10.1523/JNEUROSCI.3164-15.2016

Lee, J. J., Scott, T. L., & Perrachione, T. K. (2024). Efficient functional localization of language regions in the brain. *NeuroImage*, *285*, 120489. https://doi.org/10.1016/j.neuroimage.2023.120489

Li, J., Hiersche, K. J., & Saygin, Z. M. (2024). Demystifying visual word form area visual and nonvisual response properties with precision fmri. *iScience*, *27*, 111481. https://doi.org/10.1016/j.isci.2024.111481

Molloy, M. F., Saygin, Z. M., & Osher, D. E. (2024). Predicting high-level visual areas in the absence of task fmri. *Scientific Reports*, *14*, 11376. https://doi.org/10.1038/s41598-024-62098-9

Mur, M., Mitchell, D. J., Brüggemann, S., & Duncan, J. (2025). Stimulus effects dwarf task effects in human visual cortex. *bioRxiv*. https://doi.org/10.1101/2025.06.18.660183

Norman-Haignere, S., Kanwisher, N., & McDermott, J. H. (2013). Cortical pitch regions in humans respond primarily to resolved harmonics and are located in specific tonotopic regions of anterior auditory cortex. *The Journal of Neuroscience*, *33*, 19451–19469. https://doi.org/10.1523/JNEUROSCI.2880-13.2013

Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, *88*, 1281–1296. https://doi.org/10.1016/j.neuron.2015.11.035

O'Craven, K. M., Downing, P. E., & Kanwisher, N. (1999). Fmri evidence for objects as the units of attentional selection. *Nature*, *401*, 584–587. https://doi.org/10.1038/44134

Overath, T., McDermott, J. H., Zarate, J. M., & Poeppel, D. (2015). The cortical analysis of speech-specific temporal structure revealed by responses to sound quilts. *Nature Neuroscience*, *18*, 903–911. https://doi.org/10.1038/nn.4021

Pessoa, L. (2023). The entangled brain. *Journal of Cognitive Neuroscience*, *35*, 349–360. https://doi.org/10.1162/jocn-a-01908

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *NeuroImage*, *56*, 2356–2363. https://doi.org/10.1016/j.neuroimage.2011.03.067

Saxe, R., Schulz, L. E., & Jiang, Y. V. (2006). Reading minds versus following rules: Dissociating theory of mind and executive control in the brain. *Social Neuroscience*, *1*, 284–298. https://doi.org/10.1080/17470910601000446

Saygin, Z. M., Osher, D. E., Norton, E. S., Youssoufian, D. A., Beach, S. D., Feather, J., Gaab, N., Gabrieli, J. D. E., & Kanwisher, N. (2016). Connectivity precedes function in the development of the visual word form area. *Nature Neuroscience*, *19*, 1250–1255. https://doi.org/10.1038/nn.4354

Shain, C., & Fedorenko, E. (2025). A language network in the individualized functional connectomes of over 1,000 human brains doing arbitrary tasks. *bioRxiv*. https://doi.org/10.1101/2025.03.29.646067

Tootell, R. B. H., Reppas, J. B., Dale, A. M., Look, R. B., Sereno, M. I., Malach, R., Brady, T. J., & Rosen, B. R. (1995). Visual motion aftereffect in human cortical area mt revealed by functional magnetic resonance imaging. *Nature*, *375*, 139–141. https://doi.org/10.1038/375139a0

Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science*, *311*, 670–674. https://doi.org/10.1126/science.1119983

Tuckute, G., Lee, E. J., Sathe, A., & Fedorenko, E. (2024). A 3.5-minute-long reading-based fmri localizer for the language network. *bioRxiv*. https://doi.org/10.1101/2024.07.02.601683

Westlin, C., Theriault, J. E., Katsumi, Y., Nieto-Castanon, A., Kucyi, A., Ruf, S. F., Brown, S. M., Pavel, M., Erdogmus, D., Brooks, D. H., Quigley, K. S., Whitfield-Gabrieli, S., & Barrett, L. F. (2023). Improving the study of brain-behavior relationships by revisiting basic assumptions. *Trends in Cognitive Sciences*, *27*, 246–257. https://doi.org/10.1016/j.tics.2022.12.015

Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Covert visual attention modulates face-specific activity in the human fusiform gyrus: Fmri study. *Journal of Neurophysiology*, *79*, 1574–1578. https://doi.org/10.1152/jn.1998.79.3.1574

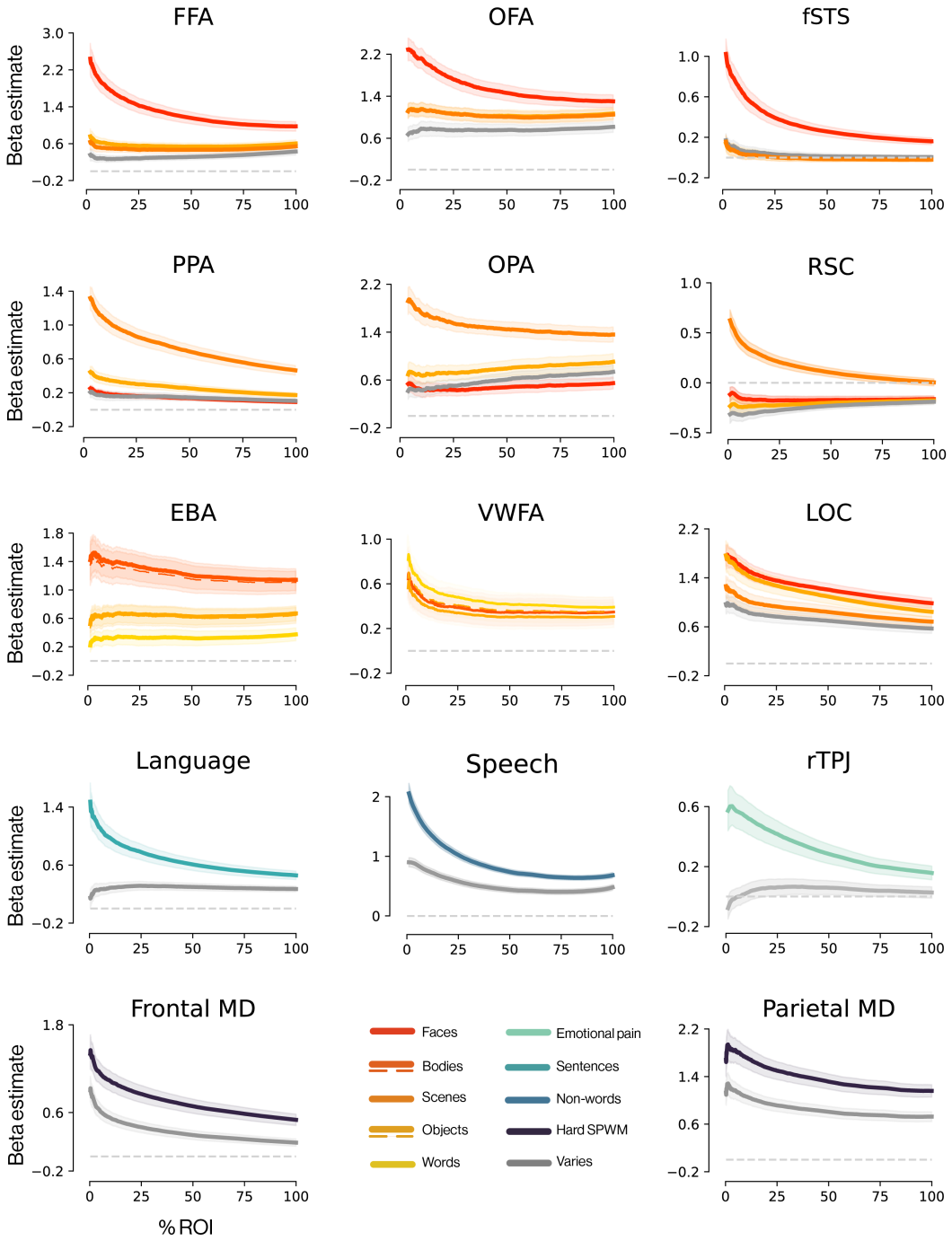
Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The

596 organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal*
597 *of Neurophysiology*, *106*, 1125–1165. https://doi.org/10.1152/jn.00338.2011

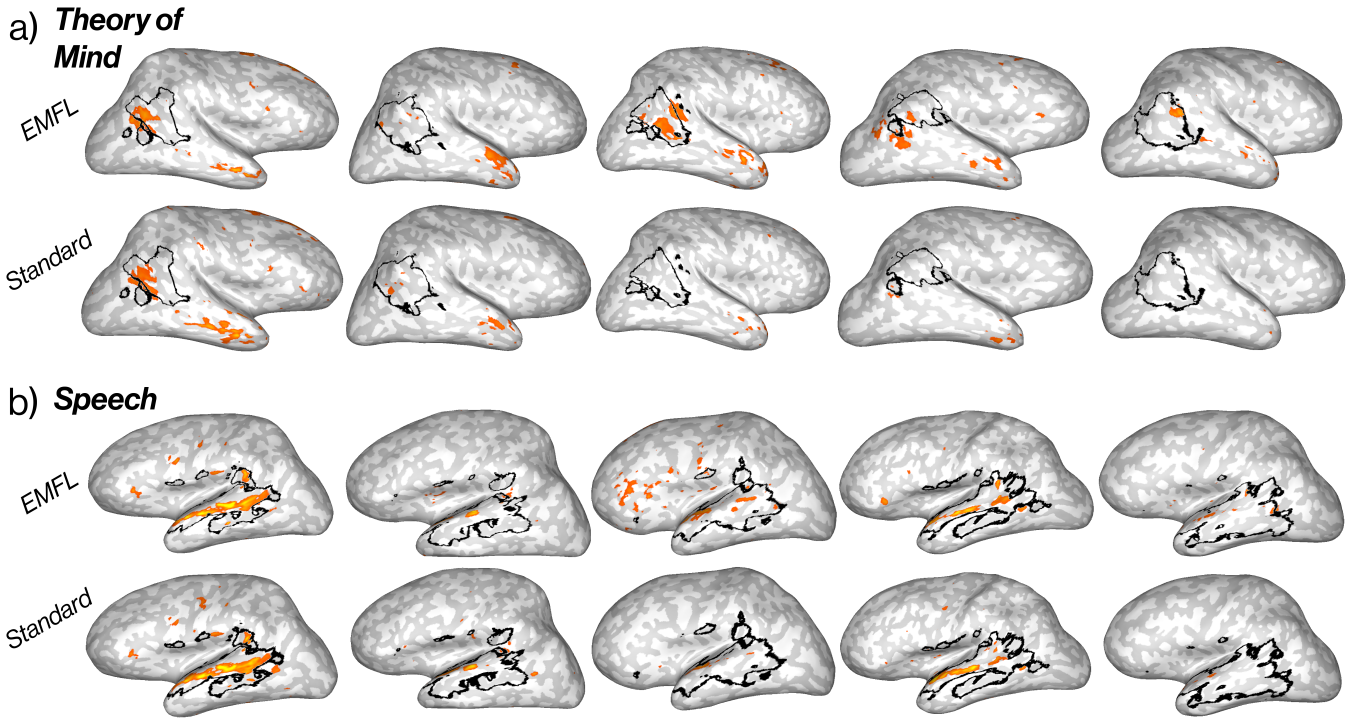
Supplementary content

	Fa	B	S	O	W	FB	FP	NW	QLT	MATH
FFA		3e-07	3e-08	1e-08	1e-07	2e-07	3e-07	5e-08	8e-08	1e-08
OFA		2e-05	1e-07	0.0001	8e-06	3e-06	9e-06	7e-06	2e-05	6e-07
fSTS		0.07	0.0002	0.002	0.0003	0.04	0.005	0.0002	2e-06	0.0006
PPA	1e-10	1e-10		1e-09	4e-10	7e-10	6e-10	2e-10	5e-10	2e-10
OPA	3e-12	6e-13		1e-10	1e-11	7e-12	6e-12	7e-12	2e-11	2e-12
RSC	7e-11	3e-09		2e-08	6e-10	8e-09	1e-09	5e-09	3e-09	4e-09
EBA	2e-08		1e-10	4e-08	5e-10	3e-08	4e-08	9e-10	7e-10	2e-10
VWFA	0.0003	0.0003	2e-06	0.0003		7e-06	0.0001	7e-07	6e-05	0.009
LOC	2e-07	0.4	2e-12		5e-12	2e-10	2e-08	1e-10	2e-09	2e-10
Language	6e-09	1e-09	6e-10	3e-07	2e-09			2e-10	1e-08	2e-08
Speech	2e-05	0.4	0.0006	0.007	4e-05	1.0	0.6		8e-07	2e-07
rTPJ	2e-09	3e-08	4e-07	4e-09	3e-09		4e-09	2e-10	2e-07	3e-09
Frontal MD	1e-08	4e-10	4e-10	2e-09	2e-08	6e-10	6e-09	3e-09	1e-09	
Parietal MD	1e-08	7e-09	2e-07	3e-08	1e-07	5e-09	6e-08	3e-08	2e-08	

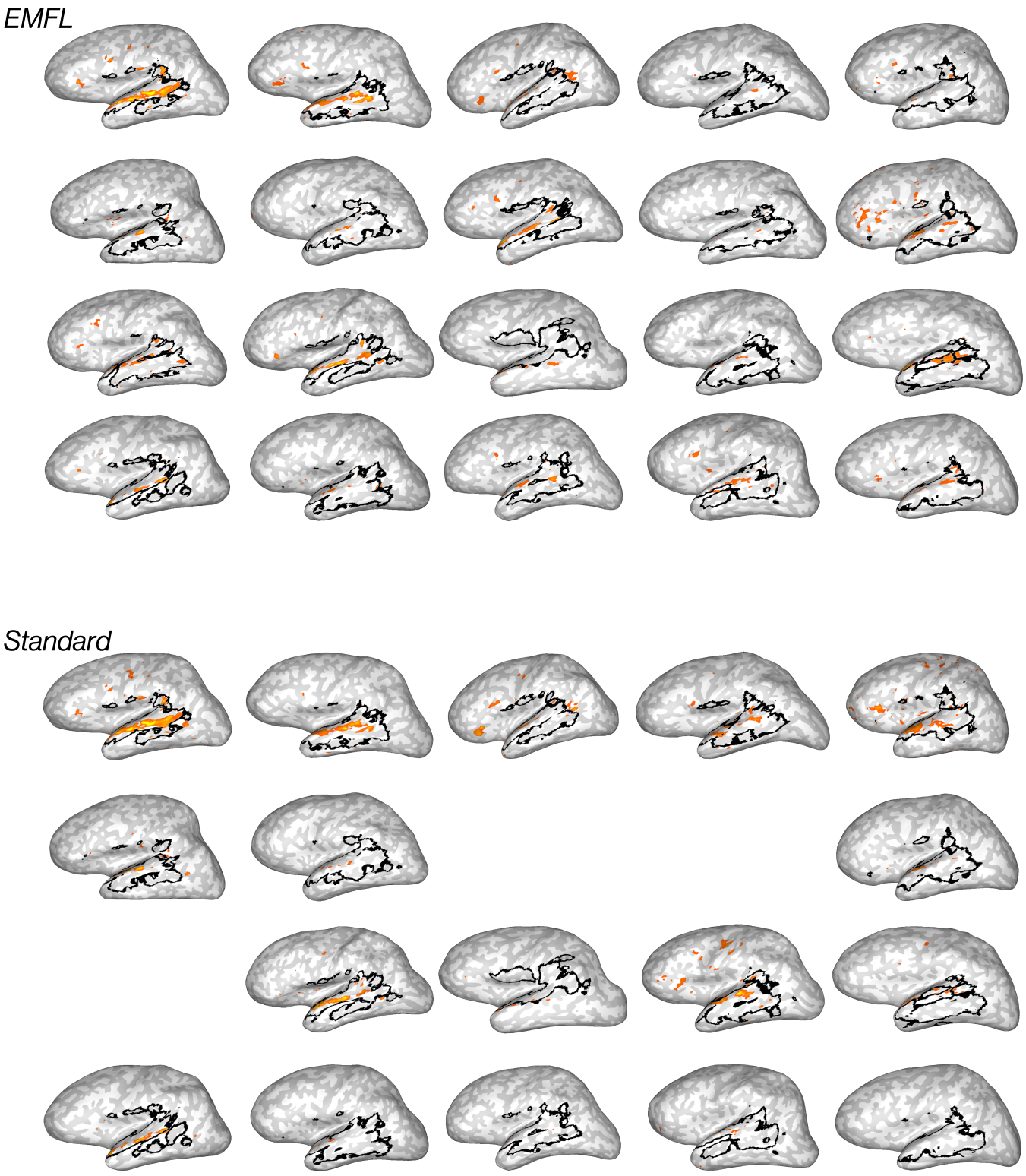
Supp. Table 1: **Contrast significance across all stimulus conditions.** Significance p values of one-tailed paired t-tests performed within each fROI, comparing the preferred condition to each of the non-preferred conditions. Blank cells indicate preferred conditions. *n.b. several contrasts are not predicted to be significant such as Objects > Bodies or Objects > Faces for LOC and Non-words > False Belief or Non-words > False Photo for rTPJ.*



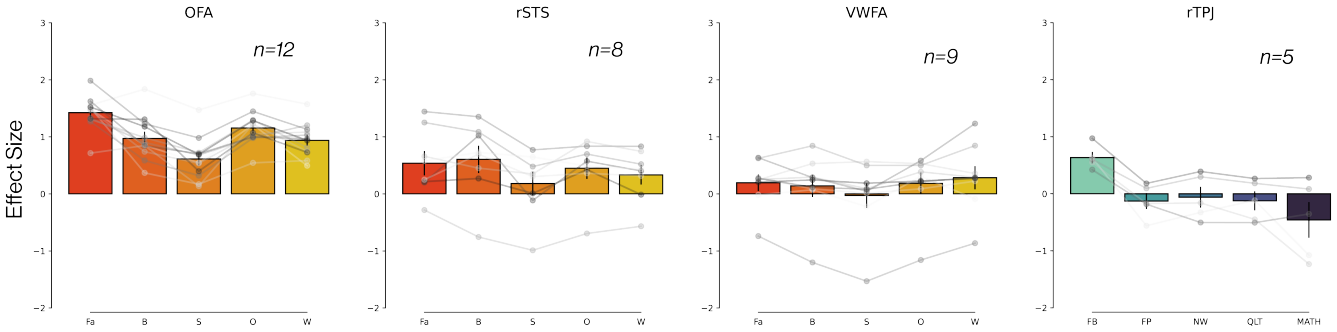
Supp. Figure 1: **Selectivity of EMFL-defined fROIs cross-validates to standard localizers.** Response of each functional region of interest defined by EMFL. We selected voxels by their functional selectivity during a subset of runs and measured their effect size during held-out runs of the relevant standard localizer. Selectivity (y-axis) is plotted as a function of ROI size (x-axis) by averaging the effect size of increasingly larger voxel clusters. *SPWM: spatial working memory, see Tables 2, 3 legend for ROI abbreviations. Dashed curves represent line-drawn versions of the same stimuli, and grey curves represent non-preferred conditions: scrambled objects (visual), non-words (language), physical pain (rTPJ), degraded audio (speech), easy SPWM (MD).*



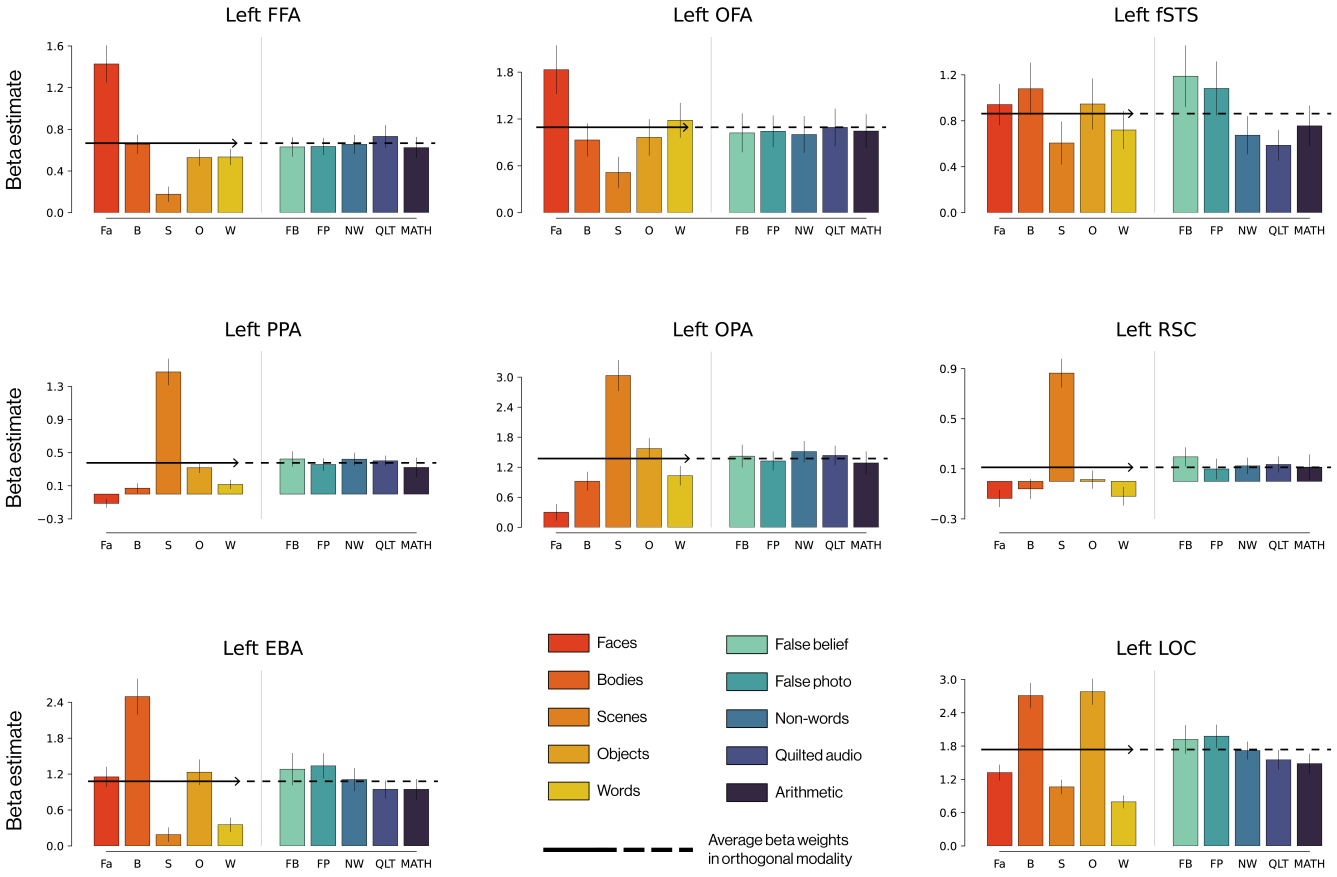
Supp. Figure 2: **Comparison of brain activation maps across EMFL and a standard localizer.** Each brain shows the functional contrast for (a) theory of mind and (b) speech regions as $-log(p\text{-}value) * sgn(t\text{-}test)$, thresholded above $+3$. Surface projections of relevant anatomical parcels are outlined in black.



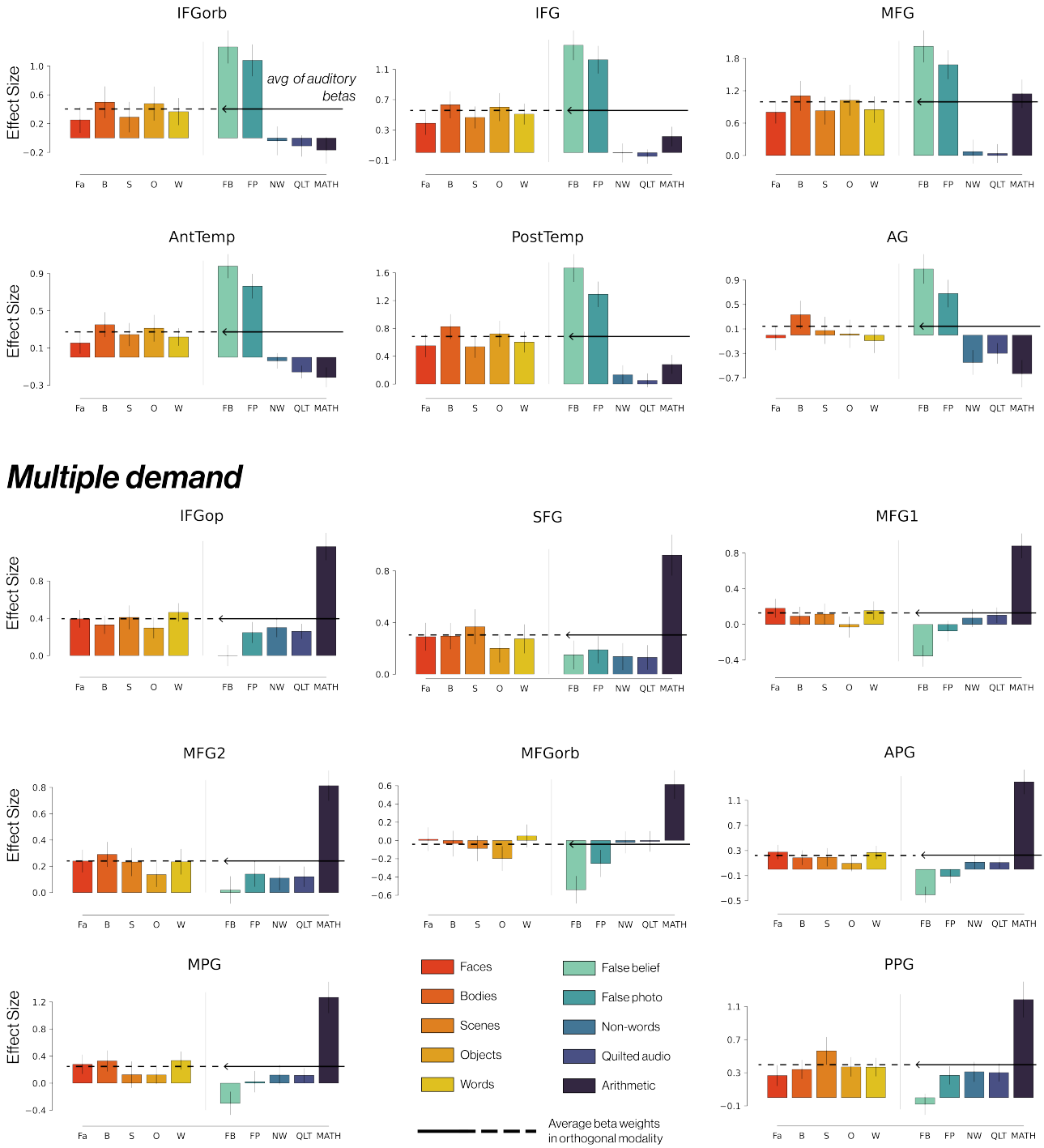
Supp. Figure 3: **Left hemisphere speech activations in all 20 subjects.** Each brain shows the functional contrast for a given condition as $-log(p\text{-}value) * sgn(t\text{-}test)$, thresholded above +3. Surface projections of relevant anatomical parcels are outlined in black. Three subjects did not perform the standard speech localizer.



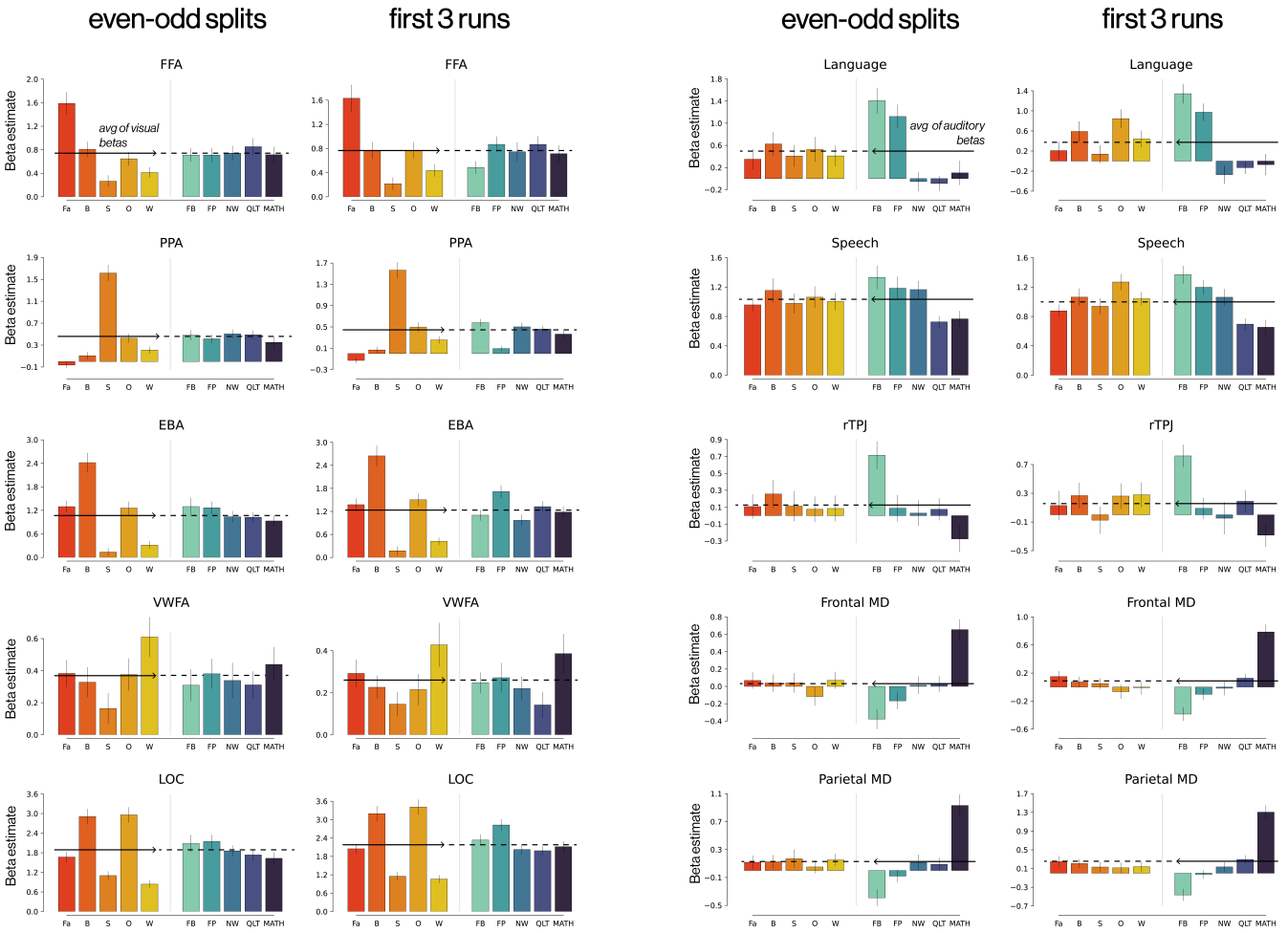
Supp. Figure 4: **Measured selectivity in subjects with <10 significant voxels.** Average effect size of the top 10% of voxels sorted by contrast selectivity. Each plot contains data only from subjects who did not reach the 10-voxel threshold as measured in Table 4. Beta weights from individual subjects are plotted as gray dots and connected with a line.



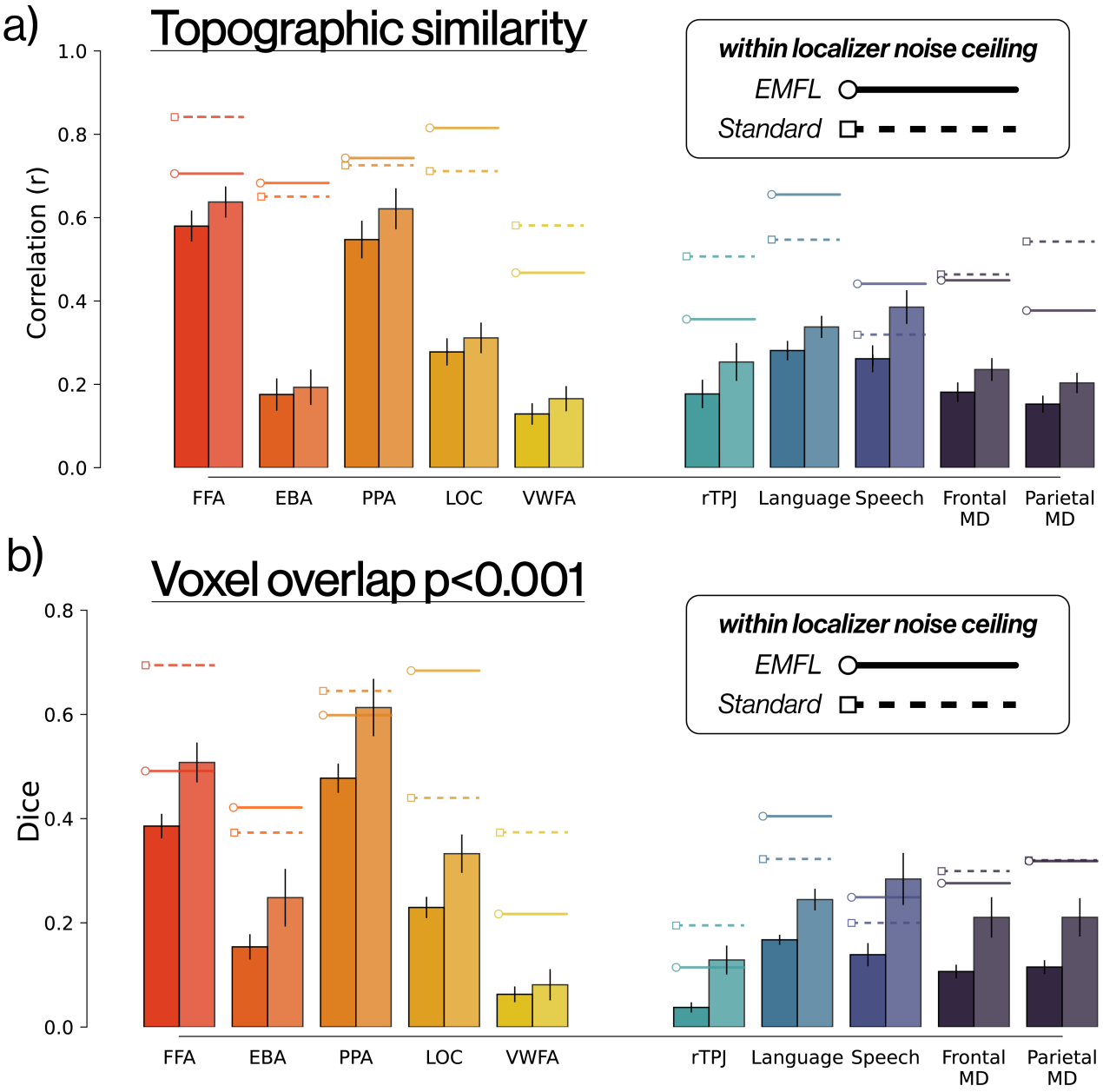
Supp. Figure 5: **Left hemisphere EMFL selectivity.** Beta estimates in held-out runs of eight left hemisphere fROIs in the ventral visual stream.



Supp. Figure 6: **Responses of Individual Language and MD Regions.** Response of individual functional regions of interest defined by EMFL to each of the ten conditions in EMFL, performed within (a) the left hemisphere language network and (b) the bilateral multiple demand network. *See Table 2, 3 for ROI abbreviations.*



Supp. Figure 7: **fROIs identified with first three runs show comparable selectivity.** Beta estimates for each stimulus condition in held-out runs of the EMFL. 'even-odd splits' averages response measurements in held-out data across even- and odd-numbered runs. 'first 3 runs' uses the first three runs (i.e. 14 minutes) to identify fROIs, and measures responses in the final two held-out runs.



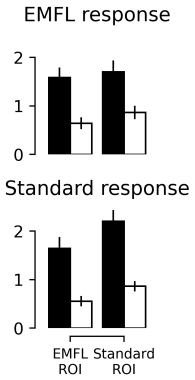
Supp. Figure 8: **Alternative measures of similarity between EMFL and standard localizers within an anatomical parcel.**

(a) Correlation of population-level voxel responses, measured via Pearson's r. (b) Dice coefficient of voxels passing a significance threshold of $p < 0.001$. Dark bars average across four even-odd splits, light bars use all runs available. Solid and dashed lines represent within-localizer split-half noise ceilings.

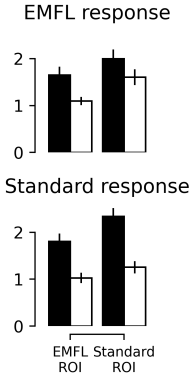
	EMFL				Standard localizers			
Runs	odd		all		odd		all	
Hemisphere	left	right	left	right	left	right	left	right
FFA	30 / 318	57 / 724	39 / 318	77 / 724	49 / 318	120 / 724	71 / 318	157 / 724
OFA	5 / 100	14 / 439	8 / 100	17 / 439	13 / 100	34 / 439	17 / 100	50 / 439
fSTS	8 / 152	44 / 929	3 / 152	50 / 929	2 / 152	38 / 929	4 / 152	69 / 929
PPA	95 / 412	121 / 400	117 / 412	146 / 400	50 / 412	73 / 400	81 / 412	111 / 400
OPA	40 / 369	68 / 417	55 / 369	86 / 417	28 / 369	48 / 417	57 / 369	82 / 417
RSC	48 / 350	79 / 498	63 / 350	103 / 498	23 / 350	45 / 498	45 / 350	78 / 498
EBA	109 / 839	135 / 1135	134 / 839	167 / 1135	77 / 850	192 / 1150	105 / 850	240 / 1150
VWFA	18 / 1238		26 / 1238		43 / 1254		66 / 1254	
LOC	338 / 2644	354 / 2248	588 / 2644	590 / 2248	254 / 2644	222 / 2248	383 / 2644	314 / 2248
rTPJ		90 / 3501		132 / 3501		73 / 3501		156 / 3501
Speech	298 / 12571	298 / 12571	398 / 12571	398 / 12571	270 / 12442	270 / 12442	509 / 12442	509 / 12442
Language (1)	62 / 480		106 / 480		28 / 480		46 / 480	
Language (2)	125 / 895		197 / 895		55 / 895		91 / 895	
Language (3)	118 / 576		157 / 576		59 / 576		83 / 576	
Language (4)	235 / 1821		343 / 1821		158 / 1821		235 / 1821	
Language (5)	447 / 3073		614 / 3073		229 / 3073		332 / 3073	
Language (6)	80 / 767		130 / 767		17 / 767		23 / 767	
Frontal MD	398 / 5192	372 / 4972	655 / 5192	692 / 4972	287 / 5192	352 / 4972	640 / 5192	857 / 4972
Parietal MD	602 / 5935	366 / 5810	912 / 5935	627 / 5810	387 / 5935	447 / 5810	1011 / 5935	1183 / 5810

Supp. Table 2: **Average number of significant voxels in each anatomical parcel.** Number of voxels that reach a significance value of $p < 0.001$ in each anatomical parcel for the relevant functional contrast. p-values were calculated using either all runs or only odd-numbered runs in either the EMFL or a standard localizer. Each cell represents an average across individual subjects, relative to the total size of each parcel. *n.b. the same parcel may differ slightly in total volume across EMFL and standard localizers because these scans were run in different scanning sessions, and alignment of the parcel to anatomicals was conducted separately for each session, resulting in small differences in parcel size.*

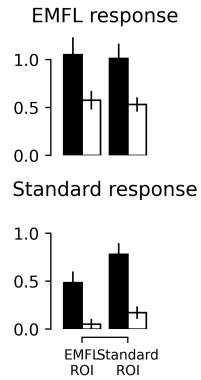
FFA	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	44.0	44.0	80.0	1.1e-15
C(definer)	1.0	3.6	3.6	6.6	0.011
C(measurer)	1.0	0.55	0.55	0.99	0.32
C(contrast):C(definer)	1.0	0.053	0.053	0.097	0.76
C(contrast):C(measurer)	1.0	1.1	1.1	1.9	0.17
C(definer):C(measurer)	1.0	0.69	0.69	1.2	0.27
C(contrast):C(definer):C(measurer)	1.0	0.31	0.31	0.56	0.46
Residual	152.0	84.0	0.55		



OFA	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	20.0	20.0	49.0	6.4e-11
C(definer)	1.0	6.5	6.5	16.0	8.1e-05
C(measurer)	1.0	0.015	0.015	0.037	0.85
C(contrast):C(definer)	1.0	0.049	0.049	0.12	0.73
C(contrast):C(measurer)	1.0	2.1	2.1	5.3	0.023
C(definer):C(measurer)	1.0	0.021	0.021	0.052	0.82
C(contrast):C(definer):C(measurer)	1.0	0.54	0.54	1.4	0.25
Residual	152.0	60.0	0.4		

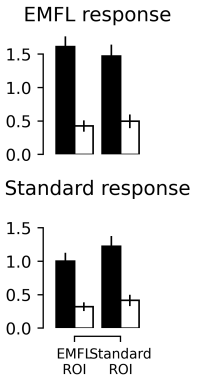


fSTS	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	10.0	10.0	44.0	6e-10
C(definer)	1.0	0.28	0.28	1.2	0.27
C(measurer)	1.0	7.1	7.1	31.0	1.2e-07
C(contrast):C(definer)	1.0	0.079	0.079	0.35	0.56
C(contrast):C(measurer)	1.0	0.019	0.019	0.084	0.77
C(definer):C(measurer)	1.0	0.63	0.63	2.7	0.1
C(contrast):C(definer):C(measurer)	1.0	0.07	0.07	0.3	0.58
Residual	152.0	35.0	0.23		

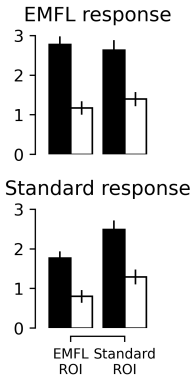


Supp. Figure 9: **ANOVA results and ROI-level plots for individual regions.** Bars represent beta estimates for preferred (black) and non-preferred (white, averaged across all non-preferred) conditions.

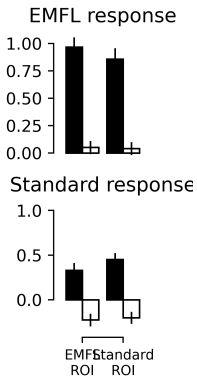
PPA	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	33.0	33.0	130.0	1.2e-22
C(definer)	1.0	0.15	0.15	0.6	0.44
C(measurer)	1.0	2.7	2.7	11.0	0.0012
C(contrast):C(definer)	1.0	0.018	0.018	0.07	0.79
C(contrast):C(measurer)	1.0	1.1	1.1	4.5	0.035
C(definer):C(measurer)	1.0	0.38	0.38	1.5	0.22
C(contrast):C(definer):C(measurer)	1.0	0.29	0.29	1.2	0.28
Residual	152.0	38.0	0.25		



OPA	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	63.0	63.0	100.0	1.2e-18
C(definer)	1.0	4.2	4.2	6.9	0.0097
C(measurer)	1.0	6.6	6.6	11.0	0.0013
C(contrast):C(definer)	1.0	0.045	0.045	0.072	0.79
C(contrast):C(measurer)	1.0	1.1	1.1	1.9	0.17
C(definer):C(measurer)	1.0	3.2	3.2	5.2	0.024
C(contrast):C(definer):C(measurer)	1.0	0.89	0.89	1.4	0.23
Residual	152.0	93.0	0.61		

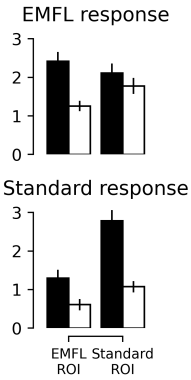


RSC	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	22.0	22.0	230.0	4.6e-32
C(definer)	1.0	0.0018	0.0018	0.019	0.89
C(measurer)	1.0	6.1	6.1	64.0	2.8e-13
C(contrast):C(definer)	1.0	1e-05	1e-05	0.00011	0.99
C(contrast):C(measurer)	1.0	0.68	0.68	7.2	0.0081
C(definer):C(measurer)	1.0	0.18	0.18	1.9	0.17
C(contrast):C(definer):C(measurer)	1.0	0.095	0.095	1.0	0.32
Residual	152.0	14.0	0.095		

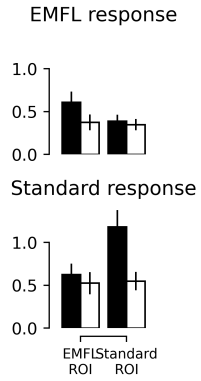


Supp. Figure 9-2: **ANOVA results and ROI-level plots for individual regions.** Bars represent beta estimates for preferred (black) and non-preferred (white, averaged across all non-preferred) conditions.

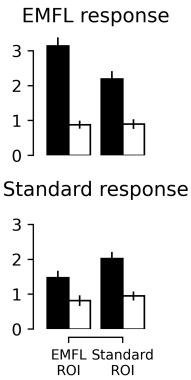
EBA	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	38.0	38.0	52.0	2.1e-11
C(definer)	1.0	12.0	12.0	16.0	8.6e-05
C(measurer)	1.0	8.1	8.1	11.0	0.0011
C(contrast):C(definer)	1.0	0.094	0.094	0.13	0.72
C(contrast):C(measurer)	1.0	2.0	2.0	2.8	0.094
C(definer):C(measurer)	1.0	7.5	7.5	10.0	0.0015
C(contrast):C(definer):C(measurer)	1.0	8.5	8.5	12.0	0.00079
Residual	152.0	110.0	0.72		



VWFA	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	2.5	2.5	10.0	0.0019
C(definer)	1.0	0.27	0.27	1.1	0.3
C(measurer)	1.0	3.4	3.4	13.0	0.00036
C(contrast):C(definer)	1.0	0.29	0.29	1.1	0.29
C(contrast):C(measurer)	1.0	0.52	0.52	2.1	0.15
C(definer):C(measurer)	1.0	1.7	1.7	6.8	0.01
C(contrast):C(definer):C(measurer)	1.0	1.3	1.3	5.2	0.024
Residual	152.0	38.0	0.25		

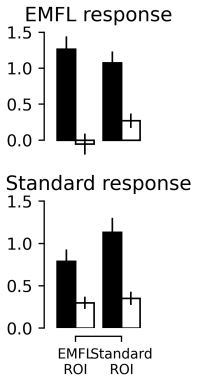


LOC	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	70.0	70.0	140.0	4.8e-23
C(definer)	1.0	0.16	0.16	0.32	0.57
C(measurer)	1.0	8.5	8.5	17.0	6.7e-05
C(contrast):C(definer)	1.0	0.78	0.78	1.5	0.22
C(contrast):C(measurer)	1.0	8.3	8.3	16.0	8.3e-05
C(definer):C(measurer)	1.0	6.4	6.4	13.0	0.00048
C(contrast):C(definer):C(measurer)	1.0	4.7	4.7	9.4	0.0026
Residual	152.0	77.0	0.51		

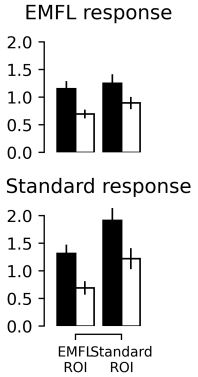


Supp. Figure 9-3: **ANOVA results and ROI-level plots for individual regions.** Bars represent beta estimates for preferred (black) and non-preferred (white, averaged across all non-preferred) conditions.

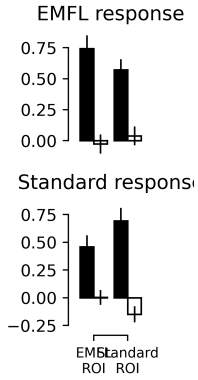
Language	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	29.0	29.0	89.0	5.6e-17
C(definer)	1.0	0.7	0.7	2.2	0.14
C(measurer)	1.0	0.00077	0.00077	0.0024	0.96
C(contrast):C(definer)	1.0	0.13	0.13	0.4	0.53
C(contrast):C(measurer)	1.0	1.8	1.8	5.6	0.019
C(definer):C(measurer)	1.0	0.17	0.17	0.52	0.47
C(contrast):C(definer):C(measurer)	1.0	1.6	1.6	5.1	0.026
Residual	152.0	49.0	0.32		



Speech	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	9.6	9.6	28.0	5.1e-07
C(definer)	1.0	4.3	4.3	13.0	0.00053
C(measurer)	1.0	2.8	2.8	8.1	0.005
C(contrast):C(definer)	1.0	0.0024	0.0024	0.0069	0.93
C(contrast):C(measurer)	1.0	0.54	0.54	1.6	0.21
C(definer):C(measurer)	1.0	1.5	1.5	4.3	0.04
C(contrast):C(definer):C(measurer)	1.0	0.062	0.062	0.18	0.67
Residual	128.0	44.0	0.34		

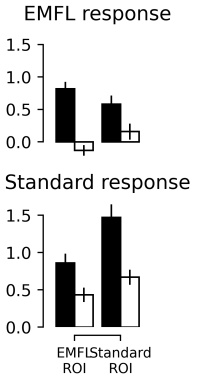


rTPJ	df	sum_sq	mean_sq	F	PR(>F)
C(contrast)	1.0	17.0	17.0	120.0	3.8e-21
C(definer)	1.0	0.0017	0.0017	0.013	0.91
C(measurer)	1.0	0.25	0.25	1.8	0.18
C(contrast):C(definer)	1.0	0.056	0.056	0.4	0.53
C(contrast):C(measurer)	1.0	1.1e-05	1.1e-05	8.2e-05	0.99
C(definer):C(measurer)	1.0	0.09	0.09	0.65	0.42
C(contrast):C(definer):C(measurer)	1.0	0.97	0.97	7.0	0.0089
Residual	152.0	21.0	0.14		

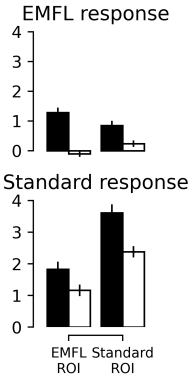


Supp. Figure 9-4: **ANOVA results and ROI-level plots for individual regions.** Bars represent beta estimates for preferred (black) and non-preferred (white, averaged across all non-preferred) conditions.

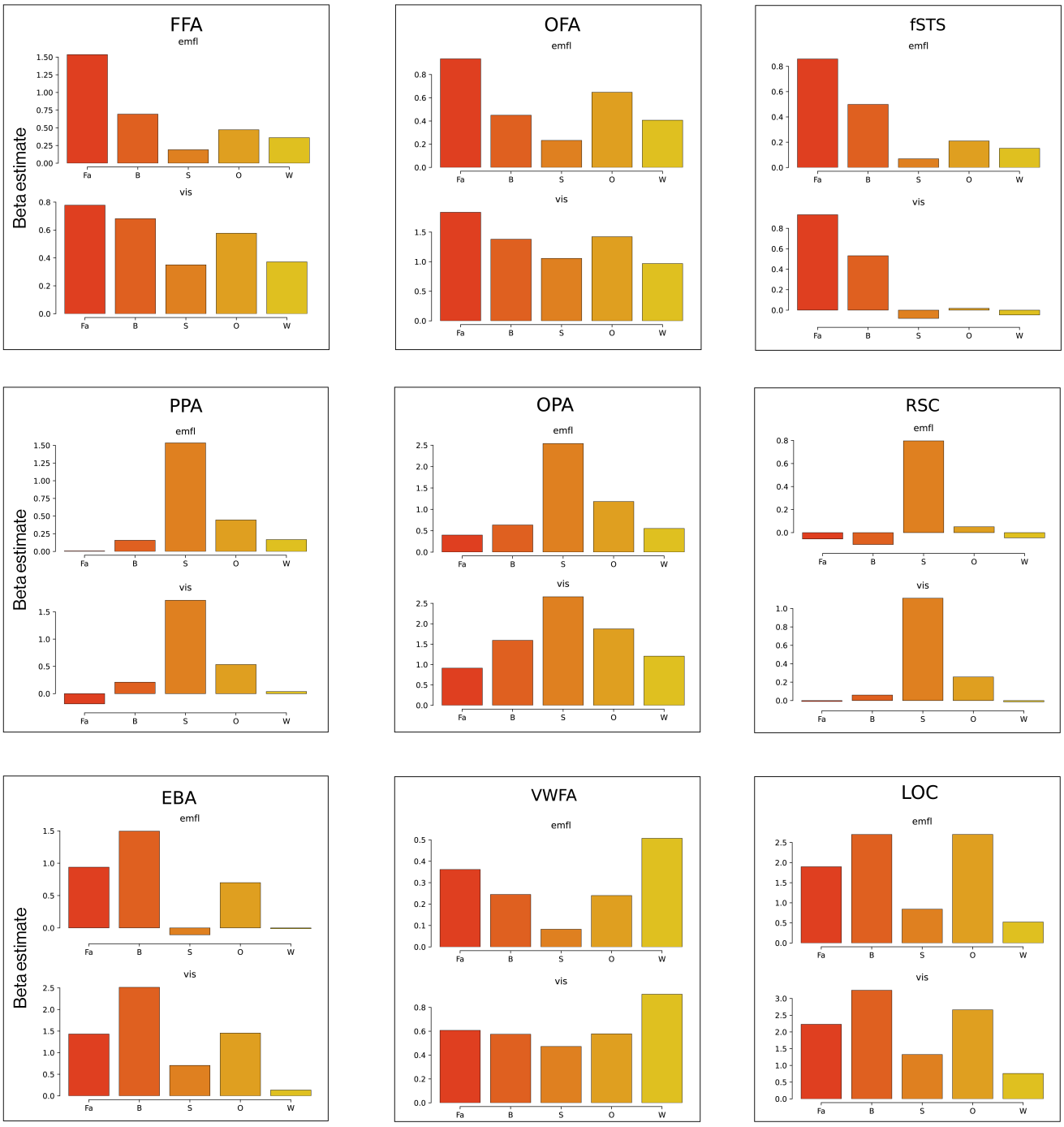
Frontal MD	**df**	**sum_sq**	**mean_sq**	**F**	**PR(>F)**
C(contrast)	1.0	17.0	17.0	70.0	3.3e-14
C(definer)	1.0	2.0	2.0	8.5	0.0042
C(measurer)	1.0	10.0	10.0	42.0	1.4e-09
C(contrast):C(definer)	1.0	0.057	0.057	0.24	0.63
C(contrast):C(measurer)	1.0	0.053	0.053	0.22	0.64
C(definer):C(measurer)	1.0	1.6	1.6	6.6	0.011
C(contrast):C(definer):C(measurer)	1.0	2.0	2.0	8.5	0.0042
Residual	152.0	37.0	0.24		



Parietal MD	**df**	**sum_sq**	**mean_sq**	**F**	**PR(>F)**
C(contrast)	1.0	38.0	38.0	73.0	1.2e-14
C(definer)	1.0	21.0	21.0	41.0	1.9e-09
C(measurer)	1.0	110.0	110.0	220.0	2.6e-31
C(contrast):C(definer)	1.0	0.12	0.12	0.24	0.63
C(contrast):C(measurer)	1.0	0.024	0.024	0.046	0.83
C(definer):C(measurer)	1.0	24.0	24.0	47.0	2e-10
C(contrast):C(definer):C(measurer)	1.0	4.5	4.5	8.7	0.0036
Residual	152.0	78.0	0.52		



Supp. Figure 9-5: **ANOVA results and ROI-level plots for individual regions.** Bars represent beta estimates for preferred (black) and non-preferred (white, averaged across all non-preferred) conditions.



Supp. Figure 10: **Visual fROI response magnitudes to EMFL stimuli.** Five subjects were recruited for a return scan in which they viewed the EMFL stimuli without the simultaneous audio. Data from both the original EMFL with auditory task (top) and visual-only (bottom) are shown for comparison.