Subject-specific segregation of functional territories based on deep phenotyping

Ana Luísa Pinho, Alexis Amadon, Murielle Fabre, Elvis Dohmatob, Isabelle Denghien, Juan Jesús Torres, Chantal Ginisty, Séverine Becuwe-Desmidt, Séverine Roger, Laurence Laurier, et al.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 Framework Program for Research and Innovation under Grant Agreement No 945539 (Human Brain Project SGA3) and Grant Agreement No. 785907 (Human Brain Project SGA2).

A battery of tasks developed by the *Human Connectome Project* (HCP) were reproduced in this study. The authors are thus thankful to the coordination facility of HCP for having made publicly available the corresponding behavioral protocols.

The authors are also thankful to the Center for Magnetic Resonance Research, University of Minnesota for having kindly provided the Multi-Band Accelerated EPI Pulse Sequence and Reconstruction Algorithms.

At last, the authors especially thank all volunteers who have accepted to be part of this challenging study, with many MRI-repeated scans over a long period of time.

Subject-specific segregation of functional territories based on deep phenotyping

Ana Luísa Pinho[1], Alexis Amadon[2], Murielle Fabre[3], Elvis Dohmatob[1,7], Isabelle Denghien[3], Juan Jesús Torre[1], Chantal Ginisty[4], Séverine Becuwe-Desmidt[4], Séverine Roger[4], Laurence Laurier[4], Véronique Joly-Testault[4], Gaëlle Médiouni-Cloarec[4], Christine Doublé[4], Bernadette Martins[4], Philippe Pinel[3], Evelyn Eger[3], Gaël Varoquaux[1], Christophe Pallier[3], Stanislas Dehaene[3,6], Lucie Hertz-Pannier[4,5], Bertrand Thirion[1]

Functional Magnetic Resonance Imaging (fMRI) has opened the possibility to investigate how brain activity is modulated by behavior. Most studies so far are bound to one single task, in which functional responses to a handful of contrasts are analyzed and reported as a group average brain map. Contrariwise, recent data-collection efforts have started to target a systematic spatial representation of multiple mental functions. In this paper, we leverage the *Individual Brain Charting* (IBC) dataset—a high-resolution task-fMRI dataset acquired in a fixed environment—in order to study the feasibility of individual mapping. First, we verify that the IBC brain maps reproduce those obtained from previous, large-scale datasets using the same tasks. Second, we confirm that the elementary spatial components, inferred across all tasks, are consistently mapped within and, to a lesser extent, across participants. Third, we demonstrate the relevance of the topographic information of the individual contrast maps, showing that contrasts from one task can be predicted by contrasts from other tasks. At last, we showcase the benefit of contrast accumulation for the fine functional characterization of brain regions within a pre-specified network. To this end, we analyze the cognitive profile of functional territories pertaining to the language network and prove that these profiles generalize across participants.

KEYWORDS

Functional Magnetic Resonance Imaging, Cognitive Function, Brain Imaging, Atlases, Data Set

[1]Université Paris-Saclay, Inria, CEA, Palaiseau, 91120, France

[2]Université Paris-Saclay, CEA, CNRS, BAOBAB, NeuroSpin, 91191, Gif-sur-Yvette, France

[3]Cognitive Neuroimaging Unit, INSERM, CEA, Université Paris-Saclay, NeuroSpin center, 91191 Gif/Yvette, France

[4]CEA Saclay/DRF/IFJ/NeuroSpin/UNIACT, France

[5]UMR 1141, NeuroDiderot, Université de Paris, France

[6]Collège de France, Paris, France

[7]Criteo AI Lab, France

Correspondence

Ana Luísa Pinho, NeuroSpin, CEA Saclay, Bâtiment 145, point courier 156 91191 Gif-sur-Yvette, France.
Email: ana.pinho@inria.fr / anagpinho@gmail.com

Funding information

Human Brain Project (HBP) SGA3, Grant Agreement No. 945539 / HBP SGA2: Grant Agreement No. 785907

1 | INTRODUCTION

Functional Magnetic Resonance Imaging (fMRI) is a non-invasive neuroimaging technique widely used to study the neural correlates of mental processes in the human brain. As such, it provides a means to characterize functional responses of brain regions. Yet, to date, fMRI research in cognitive neuroscience has focused mostly on group-level effects in the performance of isolated tasks[1]. On the other hand, the extraction of elementary cognitive components—from neuroimaging data in general—further requires the establishment of a link between tasks and cognitive functions (Posner et al., 1988; Poldrack & Yarkoni, 2016; Varoquaux et al., 2018). This can only be achieved by combining responses to many tasks, which hinges on pooling data or results from different studies. So far, data pooling relies on either meta-analytic or mega-analytic methods applied to fMRI data, allowing knowledge on brain systems to be accumulated across studies. However, as it is directly impacted by intersubject and intersite variability, this approach hinders the fine demarcation of brain regions. Contrariwise, individual mapping is free from this variability (Hanke et al., 2014; Huth, de Heer, et al., 2016), but the resulting functional topographies have not yet been integrated into brain function templates.

We thus investigate herein the feasibility of performing individual functional atlasing free from intersubject and intersite variability, as an effort to establish an univocal relationship between functional segregation of brain regions and human cognition. To this end, we leverage a collection of task-fMRI brain images from the *Individual Brain Charting* (IBC) dataset, acquired at high spatial resolution—i.e. 1.5 mm—in a fixed cohort and environment. We make use of the first IBC-dataset release (Pinho et al., 2018) composed by twelve tasks that cover already a wide variety of cognitive systems. We describe an individual-brain-modeling approach, that benefits from high-resolution data, to better account for subject-specific functional organization. The main goal of this analytic approach is to reliably localize and extract regional-specific signatures from the tasks probed, at a fine-grained scale, and delineate neural territories based on such functional fingerprints.

To understand the expected benefits of this approach, we first review the current practices for data integration and knowledge accumulation in cognitive neuroscience. Particularly, knowledge accumulation in the field relies not only on task-specific studies but also, and increasingly, on meta-analyses and mega-analyses as well as individual brain mapping.

Task-specific group studies probe the neural correlates underlying the performance of a few experimental conditions. Effects-of-interest are then obtained from linear combinations—*aka* contrasts—of these conditions. Ideally, activation patterns elicited by the contrasts should provide a mapping of the targeted mental process. Nevertheless, effects-of-interest may incorporate an undetermined level of inaccuracy, since contrasts are typically bound to idiosyncrasies of the task implementation. Some specific issues can be outlined: *(a)* arbitrary features of the elements composing the stimuli; *(b)* cues provided during performance and/or introduced informally by the experimenter during training, that can induce different cognitive strategies (Miller et al., 2012; Kirchhoff & Buckner, 2006); *(c)* specifications of the experimental equipment interacting with the presentation of the stimuli; and *(d)* temporal structure of the task, that influences perception and reaction times. These constraints may hamper the identification of brain cognitive systems from functional signatures.

Meta-analysis is a way to generalize across different implementations of tasks (Wager et al., 2007; Costafreda, 2011; Müller et al., 2018; Gurevitch et al., 2018). This approach typically relies on pooling published results from different task-specific studies, to assess which functional regions are consistently linked to some mental functions. However, combining results from different studies is subject to loss of information due to sparse peak-coordinate representation and the difficulty of consistently annotating cognitive activity. The variability in spatial location of activation across studies is caused mostly by the impact of: *(a)* the diversity of experimental settings involved in data

collection; *(b)* the use of different data processing routines (Carp, 2012); and *(c)* between-subject variability, especially in the small sample-size regime encountered in the majority of neuroimaging studies (Button et al., 2013).

As an alternative to meta-analysis, *mega-analysis* relies instead on pooled analysis of brain maps (Costafreda, 2009, 2011). This approach solves the problems linked to the usage of post-processed data taken from different sources. Recent studies, both in cognitive mapping (Schwartz et al., 2012, 2013; Varoquaux et al., 2013, 2018) and in physiological phenotyping (Wager et al., 2013), have used mega-analytic methods in order to examine large-scale neuroimaging data. As other approaches, mega-analyses also suffer from the lack of nomenclature in the annotations of task-related data (Poldrack & Yarkoni, 2016). In this respect, large-scale public repositories of task-fMRI data, namely OpenNeuro (Poldrack et al., 2013) and NeuroVault (Gorgolewski et al., 2015; Gorgolewski, Varoquaux, et al., 2016) are opening novel perspectives regarding system-level cognitive studies. Nevertheless, it is still a challenge to mitigate heterogeneity in the data related to: *(a)* different acquisition settings; and *(b)* functional and anatomical intersubject variability.

Recent neuroimaging studies have started to adopt *individual analysis*, in order to overcome both functional and anatomical intersubject variability (Fedorenko et al., 2011; Haxby et al., 2011; Nieto-Castañón & Fedorenko, 2012; Frost & Goebel, 2012; Hanke et al., 2014; Laumann et al., 2015; Huth, de Heer, et al., 2016; Huth, Lee, et al., 2016; Braga & Buckner, 2017; Gordon et al., 2017; Chang et al., 2019). Specifically, previous neuroimaging-data initiatives have already triggered the development of large-scale datasets, laying the groundwork for a more comprehensive and systematic analysis of the organization of neural structures. Many of these projects consist of resting-state fMRI data (Biswal et al., 2010; S. Smith et al., 2013; Jovicich et al., 2016; Taylor et al., 2017; van Essen et al., 2017) and some of them are specifically dedicated to research on clinical neuroscience (Book et al., 2016; Jack et al., 2015). Yet, resting-state fMRI analyses, while providing fine delineations of brain structures, do not clarify the organization of cognitive functions across territories. The task-fMRI tenet of the *Human Connectome Project* (HCP) sought to delineate and characterize representative functional territories, according to their implication in task performance (Barch et al., 2013; Glasser et al., 2016). The battery of tasks developed for the study comprises only seven tasks defining maps of twenty *main contrasts*[2]—arguably not enough for a comprehensive cognitive mapping. Much emphasis of the project was also given to characterizing population variability, in order to both conceive a common contrast atlas of the cohort (Glasser et al., 2016) and investigate correlations with behavior (S. Smith et al., 2015) or genetics (Kochunov et al., 2015). As a precursor of the IBC dataset, Pinel et al. (2007) reported a standardized large-scale acquisition set (eighty one participants) of a fast event-related task—*ARCHI Standard*—that is also included in the IBC set of experiments. However, this study was dedicated to between-subject comparison rather than fine cognitive mapping. Part of the data are openly available in the Brainomics/Localizer database (Orfanos et al., 2017). This large-scale acquisition has now been extended to four *localizers*, all of which are used in the IBC dataset (i.e. *ARCHI Standard*, *ARCHI Spatial*, *ARCHI Social* and *ARCHI Emotional*), and individual functional data, from seventy eight participants, make up the "CONNECT/Archi" Database (Pinel et al., 2019). Yet, this dataset only accounts for four different tasks, comprising twenty three main-contrast maps. Worthy of note is also the *studyforrest* initiative. It stands for a group of openly available multimodal datasets stored in the OpenNeuro repository, featuring fMRI data on the continuous presentation of scenes included in the "Forrest Gump" movie. This project has thus launched several studies dedicated to investigating the neurocognitive encoding of complex auditory and visual information, like the ability to perceive language, music or social interplay, by modeling specific audio and visual properties of the stimuli (Hanke et al., 2014, 2015, 2016; Sengupta et al., 2016). Nevertheless, the tasks employed across these studies were restricted to naturalistic stimuli. Hence, the ensuing results can not be easily integrated into current brain-function knowledge bases.

The IBC dataset yields a comprehensive collection of individual contrasts that aims at characterizing the cognitive components underlying the tasks. The first release of this dataset (Pinho et al., 2018) pertains mostly to data acquired

from localizers, using a fast and randomized design, whose conditions range from perception to higher-order thinking skills (Pinel et al., 2007; Barch et al., 2013). Besides, data from a rapid-serial-visual-presentation paradigm on language comprehension were also included in this first release (Humphries et al., 2006; Pinho et al., 2018). Given that language constitutes a primary form of social behavior among humans and its manipulation relies on higher-order cognitive processes, the corresponding functional networks share neural substrates with other cognitive-control mechanisms (Fedorenko, 2014). Plus, an individualized data integration approach is herein developed, keeping in mind that an average brain may not be similar to an individual brain (Fedorenko et al., 2011, 2012).

In this article, we provide a series of results that validate functional atlasing on the basis of individual contrast maps. With respect to previous work, we focus on quantitative arguments to precisely assess the amount of information captured by individual atlasing as well as the predictive value of the individualized model. To this end, we make use of the fifty-one main-contrast maps that can be extracted from the twelve tasks featuring the first release of the IBC dataset. These contrast maps are linearly independent and each contrast map refers to two contrasts, i.e. to the labeling contrast and its reverse. Concretely, these contrast maps pertain to unthresholded *z*-maps that are available in NeuroVault, collection #4438 (`https://identifiers.org/neurovault.collection:4438`).

First, we performed a quality check of the IBC data to verify the reliability of the functional signatures measured in the contrast maps. We thus assessed whether contrast maps of those tasks taken from previous studies were successfully reproduced. Second, we extended this quality-checking investigation and studied the amount of variability across runs (with different phase-encoding directions) and across participants in these contrast maps. Third, we synthesized the individual contrast maps into latent components, that reflect subject-specific representations, using dictionary learning; to describe the cognitive counterpart of these components, each of them was labeled according to the most contributive contrast. Similarly to the evaluation performed for the contrast maps, we also investigated the intrasubject and intersubject stability of these individual components. Fourth, to quantify the contribution of each task to these common representations, we tested whether shared variability between contrast maps can be learnt, by estimating the prediction accuracy of contrasts belonging to one task from contrasts belonging to all remaining tasks. Importantly, we also investigated how much our predictions were affected by subject-specific organization. Fifth, we demonstrate how cognitive mapping can benefit from contrasts accumulation, by analyzing functional fingerprints of regions within the language network. We thus selected and individualized six regions-of-interest from this network and draw their cognitive profile according to a subset of language-related contrast maps from different tasks, in order to disambiguate their functional role.

2 | MATERIAL AND METHODS

To prevent any ambiguity on the interpretation of MRI-related terms, definitions follow the *Brain-Imaging-Data-Structure* (BIDS) Specification version 1.2.1 (Gorgolewski, Auer, et al., 2016).

2.1 | Participants

The IBC dataset refers to neuroimaging data collected in thirteen individuals[3] (11 males, 2 females; age mean/SD=34±5 years) with no previous history of psychiatric, neurological or any other medical disorders that can change brain function. Handedness was determined with the *Edinburgh Handedness Inventory* (Oldfield, 1971); results of the survey indicate that members of the cohort are predominantly right handed (range: [0.3, 1]; mean∼0.8).

The experiments were carried out with the understanding and formal consent of the participants, in accordance

with the Helsinki declaration and the french public health regulation.

For more information about demographic data of the cohort, consult Table A.1 of the Appendix section A.

2.2 | Materials

2.2.1 | Stimuli

Stimuli consisted of both visual and auditory material presented to the participant during BOLD runs. For all tasks, they were delivered through custom-made scripts that allowed for a fully automated environment and computer-controlled collection of the behavioral data. Both visual and auditory stimuli of the protocols obtained from the HCP consortium (Barch et al., 2013) were translated to french.

All protocols are available in a public repository: `https://github.com/hbp-brain-charting/public_protocols`.

2.2.2 | MRI Equipment

The fMRI data were acquired using an MRI scanner Siemens 3T Magnetom Prisma^{fit} and a Siemens Head/Neck 64-channel coil.

Behavioral responses of the participants were registered during the MRI sessions with one of the two MR-compatible, optic-fiber response devices, depending on the protocol: *(a)* a five-button ergonomic pad (Current Designs, Package 932 with Pyka HHSC-1x5-N4); and *(b)* a pair of in-house custom-made sticks featuring one-top button. MR-Confon package was used as audio system in the MRI environment.

All sessions were conducted at the NeuroSpin platform of the CEA Research Institute, Saclay, France.

2.3 | Experimental Procedure

The task-fMRI data of the IBC-dataset first release were collected per participant in four different MRI sessions. Upon arrival to the research center, participants were instructed about the execution and timing of the tasks referring to the upcoming session. Every session was always composed of several runs dedicated to one or more tasks and each task-related run was always repeated in multiples of two, alternating its phase-encoding direction (see Section 2.5 for details).

Full information on the structure of the MRI sessions, namely which tasks were acquired in every session, can be found in Sections *2.3 Experimental Procedure* and *2.5 Data Acquisition* and Tables 2 and 3 of Pinho et al. (2018) as well as in the Section *2 Acquisition parameters* of the IBC documentation (see Section 2.6 of this article for more details about the IBC documentation).

Additionally, a detailed description about the procedures that were undertaken toward handling and training the participant before each MRI session are also provided in Section 2.3 of Pinho et al. (2018).

2.4 | Experimental Paradigms

The first release of the IBC dataset comprises twelve tasks. Their majority pertains to already-existing paradigms developed and validated in previous studies. They were thus combined in four different MRI sessions according to this criterion. For further details about the organization of the MRI sessions, consult Table 2 of Pinho et al. (2018). Concretely, these twelve tasks refer to the reproduction of the ARCHI and HCP batteries, with minor adaptations,

as well as to a new task addressing language comprehension and based on a *Rapid-Serial-Visual-Presentation* (RSVP) paradigm. A summary of the experimental paradigm of each task is given next. Each design is categorical, i.e. it relies on the principle of "cognitive subtraction" as means to isolate effects-of-interest across experimental conditions within every task. All conditions per task are listed and described in Appendix section B. More information about adaptation and implementation of the corresponding software protocols can be found in Section 2.4 of Pinho et al. (2018).

2.4.1 | ARCHI Tasks

The *ARCHI* battery was developed at NeuroSpin Research Center and validated in the context of numerous neuroimaging projects, namely Pinel et al. (2007, 2019). It is composed of four localizers—ARCHI Standard, ARCHI Spatial, ARCHI Social and ARCHI Emotional—addressing different psychological domains. For further details about these localizers, consult the Appendix section B.1; the list of experimental conditions composing each localizer, together with their descriptions, can be found in Tables B.1, B.2, B.3 and B.4, respectively.

2.4.2 | HCP tasks

The HCP battery refer to a subset of tasks—*HCP Emotion*, *HCP Gambling*, *HCP Motor*, *HCP Language*, *HCP Relational*, *HCP Social* and *HCP Working Memory*—that were originally developed in the task-fMRI tenet of HCP, as described in Barch et al. (2013). They are organized in block-design paradigms, covering a variety of motor, sensory, emotional and high-order cognitive mechanisms. For further details about the HCP tasks, consult the Appendix section B.2; the list of experimental conditions composing each task, together with their descriptions, can be found in Tables B.5, B.6, B.7, B.8, B.9, B.10 and B.11, respectively.

2.4.3 | RSVP Language task

Finally, the *RSVP Language* task consists in the rapid-serial-visual-presentation of linguistic constituents, such as sentences of different complexity, *jabberwocky*, words, pseudowords and nonwords. This novel task is a follow-up of the one presented in Humphries et al. (2006) and it aims foremost at dissociating syntactic and semantic processes related to language comprehension. Main conditions of this task are listed and described in Table B.12 in the Appendix section B.2.8.

2.5 | Imaging-data acquisition

FMRI data were collected using a *Gradient-Echo* (GE) pulse, whole-brain *Multi-Band* (MB) accelerated (Moeller et al., 2010; Feinberg et al., 2010) *Echo-Planar Imaging* (EPI) T2*-weighted sequence with *Blood-Oxygenation-Level-Dependent* (BOLD) contrasts, using the following parameters: the repetition time (TR) is 2000 ms; the echo time (TE) is 27 ms; the flip angle is 74°; the field-of-view (FOV) is $192 \times 192 \times 140$ mm^3; voxel size is $1.5 \times 1.5 \times 1.5$ mm^3; the slice orientation is axial; slices are acquired in interleaved fashion; in-plane acquisitions were accelerated by a factor (GRAPPA) of 2; and across slices, a multi-band factor of 3 was used. Two different acquisitions for the same task were always performed using two opposite phase-encoding directions: one from *Posterior to Anterior* (PA) and the other from *Anterior to Posterior* (AP). The main purpose was to mitigate geometrical distortions while assuring built-in, within-subject replication of the same tasks.

Spin-Echo (SE) EPI-2D image volumes were acquired in order to correct for spatial distortions, using the following

parameters: a TR of 7680 ms; a TE of 46 ms; a FOV of $192 \times 192 \times 140$ mm^3; a voxel size of $1.5 \times 1.5 \times 1.5$ mm^3; axial slice orientation; and acceleration factor (GRAPPA) = 2. Similarly to the GE-EPI sequences, two different acquisitions were also performed using PA and AP phase-encoding directions.

In addition, a *3D magnetization-prepared rapid gradient-echo* (MP-RAGE) T1-weighted anatomical-image volume, covering the whole brain, was acquired with the following parameters: voxel size of $1 \times 1 \times 1$ mm^3; sagittal slice orientation; flip angle of 9°; and FOV of $256 \times 256 \times 160$ mm.

2.6 | Data Storage

The online access of all data is assured by the *Human Brain Project* (HBP) EBRAINS platform as well as the OpenNeuro public repository (Poldrack et al., 2013) under the accession number ds000244 (dataset DOI: 10.18112/openneuro.ds000244.v1.0.0).

The individual and unthresholded *z*-maps, obtained from all contrast maps of the aforementioned experimental conditions (see Section Experimental Paradigms and Appendix section B), can be found in the *NeuroVault* repository (Gorgolewski et al., 2015), under the collection with the id=4438: `https://identifiers.org/neurovault.collection:4438`.

The scripts used for data analysis are publicly available under the Simplified BSD license: `https://github.com/hbp-brain-charting/public_analysis_code`.

The documentation of the dataset, containing a full description of the experimental designs, acquisition parameters and analysis pipeline, is accessible on the IBC website: `https://project.inria.fr/IBC/data/`.

For further details, consult the *Data and Code Availability Statement* provided as Supplementary Material.

2.7 | Data Analysis

2.7.1 | Preprocessing

Raw data were preprocessed using *PyPreprocess* (`https://github.com/neurospin/pypreprocess`). This framework stands for a collection of Python scripts oriented toward a common workflow of fMRI-data preprocessing analysis. It is built upon the *Nipype* interface (Gorgolewski et al., 2011) v0.12.1, thus allowing for the stand-alone use of precompiled modules in *SPM12* software package (Wellcome Department of Imaging Neuroscience, London, UK) v6685, and *FSL* library (Analysis Group, FMRIB, Oxford, UK) v5.0, used for the preprocessing of neuroimaging data.

All fMRI images, i.e. GE-EPI volumes, were collected twice with reversed phase-encoding directions, resulting in pairs of images with distortions going in opposite directions. Susceptibility-induced off-resonance field was estimated from the two Spin-Echo EPI volumes in reversed phase-encoding directions. The images were corrected based on the estimated deformation model. Details about the method and its implementation in FSL can be found in Andersson et al. (2003) and S. Smith et al. (2004), respectively.

Further, the GE-EPI volumes were aligned to each other within every participant. A rigid body transformation was employed, in which the average volume of all images was used as reference (Friston et al., 1995).

All analysis were carried out in both volume and cortical surface. For the volume-based analysis, the corresponding T1-weighted MPRAGE (anatomical) volume was co-registered onto the mean EPI volume for every participant (Ashburner & Friston, 1997). All anatomical volumes were then segmented to finally allow for the normalization of both functional and anatomical data (Ashburner & Friston, 2005). Concretely, the segmented volumes were used to compute the deformation field for normalization into the MNI152 space. The deformation field was then applied to

the EPI data. In the end, all volumes were resampled to their original resolution, i.e. 1 mm isotropic for the MPRAGE T1-weighted anotomical images and 1.5 mm for the EPI images. For the cortical-surface analysis, the anatomical and motion-corrected fMRI images were given as input to *FreeSurfer* v6.0.0, in order to extract meshes of the tissue interfaces and a sampling of functional activation on these meshes, as described in van Essen et al. (2012). The corresponding maps were then resampled to the fsaverage7 template of FreeSurfer (Fischl et al., 1999).

2.7.2 | FMRI-Model Specification

FMRI data were analyzed using the *General Linear Model* (GLM). Regressors of the model were designed to capture variations in BOLD response strictly following stimulus timing specifications. They were estimated through the convolution of boxcar functions, that represent per-condition stimulus occurrences, with the canonical *Hemodynamic Response Function* (HRF), defined according to Friston, Fletcher, et al. (1998) and Friston, Josephs, et al. (1998).

To build such models, paradigm descriptors grouped in triplets (i.e. onset time, duration and trial type) according to BIDS Specification were determined from the log files' registries generated by the stimulus-delivery software.

To account for small fluctuations in the latency of the HRF peak response, additional regressors were computed based on the convolution of the same task-conditions profile with the time derivative of the HRF.

Nuisance regressors were also added to the design matrix in order to minimize the final residual error. To remove signal variance associated with spurious effects arising from movements, six temporal regressors were defined for the motion parameters. Further, the first five principal components of the signal, extracted from voxels showing the 5% highest variance, were also regressed to capture physiological noise (Behzadi et al., 2007).

In addition, a discrete-cosine basis was included for high-pass filtering (*cutoff*= $\frac{1}{128}$ Hz). Model specification was implemented using *Nistats* library v0.0.1b, a Python module devoted to statistical analysis of fMRI data (`http://nistats.github.io`), that leverages *Nilearn* (Abraham et al., 2014), a Python library for statistical learning on neuroimaging data (`https://nilearn.github.io`).

2.7.3 | Model Estimation

To restrict GLM-parameters estimation of volume data to gray-matter voxels, a group-level gray-matter mask was initially generated by averaging all individual, gray-matter "density" maps obtained upon the preprocessing step; it was then thresholded at a liberal level, i.e. 0.25, which corresponds to an average probability of 25% of finding gray matter in a certain voxel across subjects, in order to obtain a comprehensive mask. An illustration of the group-level, gray-matter mask is provided as Supplementary Material.

Regarding noise modeling, a first-order autoregressive model was used in the maximum likelihood estimation procedure.

A mass-univariate GLM fit was applied separately to the preprocessed GE-EPI data of each run with respect to a specific task. Parameter estimates pertaining to the experimental conditions were thus computed, along with the respective covariance at every voxel. Various contrasts—linear combinations of the effects—were then defined, referring only to differences in evoked responses between either *(a)* two conditions-of-interest or *(b)* one condition-of-interest and baseline. GLM estimation and subsequent statistical analyses were also implemented using Nistats v0.0.1b. FMRI-data analysis was firstly employed with no regularization and, afterwards, with a smoothing parameter set to 5 mm full-width-at-half-maximum. Such procedure allows for an increase of the *functional Signal-to-Noise Ratio* (fSNR) and it facilitates between-subject comparison. The images used in Section 3 are based on smoothed data. The same analytical steps were carried out on the data sampled on the cortical surface.

2.7.4 | Summary Statistics

Because data from each task were collected, at least in two acquisitions, with opposite phase-encoding directions (see Section Imaging-data acquisition for details), statistics of their joint effects were calculated for each subject with a fixed-effects (FFX) model. Specifically, *t*-statistics were computed at every voxel for every contrast map.

After completing the within-subject level analysis, a group-conjunction analysis (Heller et al., 2007) was performed on the main contrasts; significance was assessed against the null hypothesis asserting that less than a fourth of the subjects display activity in a given voxel. The motivation to employ conjunction analysis instead of the random-effects (RFX) linear model relates to the small size of the cohort, i.e. thirteen subjects. Due to this small size, the variance term of the random-effects model is estimated with few degrees of freedom, leading to unreliable and conservative estimates. Moreover, as variability tends to co-localize with large effects (Thirion et al., 2007), a form of inference relying on the consensus between subjects is more appropriate for small samples. Map consistency using either conjunction analysis or RFX-model approach was investigated on post-processed data, with and without smoothing. Concretely, the distributions of the Jaccard index between pairs of group-level *z*-maps from the original set and bootstrap resampling were estimated using either conjunction analysis or RFX-model, to assess how much they are impacted by sampling variability in the population. Results are provided as Supplementary Material. We observe that the median of the indices is higher for the conjunction analysis while its interquartile range is smaller, indicating that group-level conjunction contrasts are more consistent across different samples of the population. We also denote that there is a likely difference between RFX results obtained for smoothed and unsmoothed data, whereas there is no difference between the conjunction counterparts; this evidence also hints at a greater consistency of the second-level results obtained from conjunction, regardless different post-processing choices.

A *z*-transform of the *t*-maps was always performed both on individual and group-level statistical maps. Such procedure allows for standardized results, which are not anymore dependent on the number of degrees of freedom. This is particularly useful when comparing results from different datasets (see Sections 2.7.5 and 3.1).

2.7.5 | Assessing Reproducibility and Functional Variability

In order to assess the reproducibility of the main results, group-level contrast *z*-maps of the IBC dataset were compared to the corresponding ones obtained from the original ARCHI dataset, n=78 (Pinel et al., 2019) as well as to those obtained from a subset of the HCPS900 dataset, n=786.

To evaluate how consistent the task-specific functional signatures are among participants, the correlation of the contrast maps was computed both between and within subjects, i.e. both between all subjects' pairs of FFX individual maps and between pairs of "PA" and "AP" maps of the same task per subject. The mean and the 95% Confidence Interval (CI) were then estimated from the distribution of this measure, as a way to quantify the functional variability of the contrasts within and across individuals.

2.7.6 | Dictionary Learning

To obtain individual topographies that summarize the spatial representation of the effects-of-interest, dictionary learning was applied to statistical maps featuring the main contrasts of each experiment, in each subject. This sparse decomposition method extracts spatially distinct components from the joint distribution of contrast values. It thus provides a low-dimension representation of the overlapping neural support of the contrasts probed in the dataset.

We used a multiple-subject dictionary learning model that captures functional correspondence across datasets,

as described in Varoquaux et al. (2013). Such hypothesis allows the factorization of the individual contrast maps into a dictionary of cognitive profiles (common to all subjects) plus subject-specific spatial maps (*aka* loadings). The sparsity was enforced with an ℓ_1–norm penalty on the loadings of the components, together with a nonnegative constraint.

Formally, consider the set of brain maps $\mathbf{X}^s = (\mathbf{X}^s_j), j \in [c]$ obtained for $c = 51$ contrasts in a subject $s \in [n]$. By enumerating the values across a mesh of vertices, each $\mathbf{X}^s_j$ is a p–dimensional vector, where p is the number of vertices; $\mathbf{X}^s$ is thus a matrix of size $p \times c$. Functional-correspondence dictionary learning solves the following minimization problem for $\lambda > 0$:

$$\min_{(\mathbf{U}^s)s=1\ldots n, \mathbf{V}\in C} \sum_{s=1}^{n} \left(\|\mathbf{X}^s - \mathbf{U}^s\mathbf{V}\|^2 + \lambda \|\mathbf{U}^s\|_1 \right),$$

where $\mathbf{U}^s \geq 0\,,\ \forall s \in [n]$. Here, C denotes the set of matrices with row norm smaller than 1. $\mathbf{U}_s$ matrices have shape $p \times k$, whereas the functional-loading matrix $\mathbf{V}$ has shape $k \times c$, k being the number of components. Herein, we used $k = 20$. On the other hand, the $\mathbf{V}$ matrix defines the functional characteristics of the components. The estimated subject-specific spatial components $(\mathbf{U}^s),\ s \in [n]$ can be interpreted as individual topographies; these components may overlap, although their values are zero in most regions. The λ parameter was calibrated in order to yield a sparsity of around 75%. As the estimation problem is non-convex, initialization matters; here, we created an initial $\mathbf{V}$ matrix by clustering the voxels across subjects into $k = 20$ clusters and took the normalized average of the cluster signal. The implementation relies on the *mini-batch k-means* and the *dictionary-learning* methods of *scikit-learn* v0.21.3 (Pedregosa et al., 2011), a Python machine-learning library (`https://scikit-learn.org/stable/`).

For display purposes, a hard assignment of vertices to components was undertaken by ascribing each vertex to the component that gets the largest loading (note that all $\mathbf{U}^s$ are nonnegative):

$$\text{label}^s_i = \begin{cases} 0 \text{ if } \mathbf{U}^s_{i,j} = 0\,,\ \forall j \in [c] \\ \text{argmax}_{j\in[c]}\mathbf{U}^s_{i,j}\,,\ \text{otherwise} \end{cases}$$

Finally, group maps $\bar{\mathbf{U}}$ were obtained by first computing the median of the components, as follows:

$$\bar{\mathbf{U}}_{i,j} = \text{median}_{s\in[n]}\mathbf{U}^s_{i,j}$$

Then, the same labeling strategy, as the one described above, was employed at every vertex. Note that by construction, a label was assigned to a given vertex if, at least, half of the subjects obtained a nonzero loading for the corresponding component at that vertex. Because components were labeled to outline their functional definition, we chose the contrast label with the highest value in the dictionary—row of $\mathbf{V}$ matrix—for that component.

The stability of this decomposition was thus tested by re-estimating the dictionary, using only AP- and PA-based activation maps; spatial correspondence of the components was measured through their intrasubject and intersubject correlations. *A priori*, it was hypothesized that such correlations would be higher within than between subjects.

2.7.7 | Prediction of Contrast Maps across Tasks

Because functional activations are overlapping between the IBC tasks due to common underlying cognitive components, we hypothesized that such similarities allow for predicting these contrast maps from the contrast maps of other tasks.

Keeping the notation of the previous section, let us consider the problem of predicting one contrast of one subject, which refers to the column $\mathbf{X}_j^s$ of $\mathbf{X}^s$, $j \in [c]$. We shall also consider the remaining columns $\mathbf{X}_{-j}^s$ that represent brain activity from the remaining contrasts, as well as all data from the other subjects $\mathbf{X}^i$, $i \neq s$. To this end, we used Ridge regression:

$$\widehat{\mathbf{w}}^{s,\lambda,j} = \text{argmin}_{\mathbf{w}\in\mathbb{R}^{c-1}} \sum_{i\neq s} \|\mathbf{X}_j^i - \mathbf{X}_{-j}^i \mathbf{w}\|^2 + \lambda \|\mathbf{w}\|^2,$$

in which λ was chosen by generalized cross-validation. The prediction output was defined as follows:

$$\widehat{\mathbf{X}}_j^s = \mathbf{X}_{-j}^s \widehat{\mathbf{w}}^{s,\lambda,j}.$$

The success of the prediction is quantified by the cross-validated coefficient of determination, that is computed at each location i:

$$\mathbf{R}_i^2(j) = 1 - \text{mean}_{s\in[n]} \frac{\|\widehat{\mathbf{X}}_{i,j}^s - \mathbf{X}_{i,j}^s\|^2}{\|\mathbf{X}_{i,j}^s\|^2}$$

In practice, a leave-three-subjects-out cross-validation experiment was employed, i.e. training sets were composed of the contrasts from ten subjects and three were left for the test set. Contrasts from the same task were predicted jointly, as they are correlated; hence, predicting one from the other would lead to optimistic bias. *Ward's* clustering method was used for parceling the brain volume into 100 sub-regions. The aforementioned analysis was performed in these sub-regions, similarly to what was done in Tavor et al. (2016). Prediction accuracy at every voxel was measured in terms of $\mathbf{R}^2$–scores obtained from each data partition. The maximum of this statistic across tasks was reported afterward. The reason for taking the maximum relates to the fact that, at a given location, it is unlikely to obtain robust responses for more than one task; this means that the baseline of the predictive score is zero[4], as confirmed by the analogous computation performed with a dummy classifier. It thus becomes more appropriate to select the maximum rather than the average or the median of the accuracy predictions across tasks, since the maximum outlines the existence of, at least, one latent cognitive dimension explaining activity in a certain brain location.

Finally, the same experiment was repeated but scrambling the subject correspondence between train and test through a random permutation. This aims at quantifying the subject specificity of these functional relationships.

2.7.8 | Region-of-Interest Analysis

The functional-response profiles of a set of six brain regions, previously reported to participate in the language network, were obtained from a subset of contrasts pertaining to language processing. Six regions-of-interest, based on the statistical maps obtained from Pallier et al. (2011), were selected *a priori*.

To better account for the differences inherently associated with the precise location and shape of these regions in every participant, subject-specific counterparts of these region-of-interest were computed by applying a dual-regression approach adapted from Nickerson et al. (2017). Formally, consider the set of six regions-of-interest's represented as binary vectors (masks) with p voxels aggregated in a matrix $\mathbf{R}$ of size $6 \times p$. For a subject $s \in [n]$, its individual projection $\mathbf{R}^{(s)}$ is a matrix with the same size as $\mathbf{R}$, computed as follows:

$$\mathbf{R}^{(s)} = \mathbf{R}\ pinv(\mathbf{X}^{(s)})\ \mathbf{X}^{(s)},$$

in which matrix $\mathbf{X}^{(s)}$, with size $j \times p$, represents the set of j individual contrast maps, where $j \in [c]$ obtained for $c = 43$; $pinv(\mathbf{X}^{(s)})$ denotes the *pseudoinverse matrix* of $\mathbf{X}^{(s)}$. One should note that $pinv(\mathbf{X}^{(s)})\,\mathbf{X}^{(s)}$ is a linear projection operator that transforms $\mathbf{R}$ into $\mathbf{R}^{(s)}$, according to the subject-specific functional information embedded in $\mathbf{X}^{(s)}$. This approach is an adaptation of the two stages featuring the dual-regression approach (consult Figure 1 and equations *(1)* and *(3)* of Nickerson et al. (2017)) and its notation is adapted to the type of data used in the present context. The number of contrasts maps c amounts to 43 because these are the 51 main IBC contrast maps from the first release minus those 8 selected to draw their functional profiles (see Section 3.5). In order to generate the corresponding subject-specific region masks, the maps obtained in $\mathbf{R}^{(s)}$ were then transformed into binary maps according to a prespecified threshold; this threshold was set in order to yield the same number of voxels in individual regions as in the group region.

Afterwards, for each region-of-interest, the mean of the z-scores at the voxels of the selected contrasts, belonging to the corresponding subject-specific region mask, was computed in every subject. Then, for each contrast, the average across subjects and its 95% CI was estimated in order to create the functional fingerprint specific of that region.

Moreover, a leave-one-group-out cross-validation experiment was set in order to verify whether voxels can be correctly assigned to one of these six regions-of-interest, according to the profile of their functional activations. This experiment was run on data from all participants; voxels belonging to pairs of regions-of-interest were classified to their respective region—based on the functional contrasts—using a linear support vector classifier. Training sets were composed of data from twelve participants and prediction was performed in the data from the remaining participant; each group was thus referring to the data from one participant. Predictions were made between all possible combinations of two regions from the original collection of six regions-of-interest. Accuracy prediction at every voxel was measured in terms of classification accuracy per region pair. This analysis was performed using scikit-learn. To compute the chance level, the same procedure was employed using the scikit-learn's dummy classifier.

3 | RESULTS

All results herein described were obtained from the individual and unthresholded z-maps of the main contrasts extracted from each task. For a cognitive description of the effects-of-interest depicted by these contrasts, refer to the IBC documentation available on the website of the project (`https://project.inria.fr/IBC/data/`).

3.1 | Reproduction of ARCHI and HCP Tasks

Figure 1 displays the correlation of group-average z-maps between corresponding contrasts obtained from: *(a)* IBC versus ARCHI datasets and *(b)* IBC versus HCP datasets. The diagonal of the matrix refers to homologous contrasts and, therefore, it shows how successfully the maps of ARCHI and HCP were reproduced in IBC. The high correlations present on the diagonal confirm that the functional signatures of the same contrast maps were preserved across different datasets, providing clear evidence that the original results obtained in these large datasets are reproduced using the IBC dataset. One exception stands for the *punishement-reward* contrast map of the HCP Gambling task; nonetheless, results for this contrast map were also very variable within and between subjects for the original dataset (see Figure C.1 in Appendix section C). Additionally, correlations between different contrast maps provide a quantitative measure of the similarity of their neural correlates. For instance, *left hand* versus *right hand* contrast map from the ARCHI Standard task and *left hand* versus *any motion* contrast map from the HCP Motor task are highly correlated because both functional signatures mostly refer to movements of the left hand. As another example, *social-interaction*

motion versus *random motion* contrast maps from both ARCHI Social and HCP Social tasks are also highly correlated since their functional signatures both pertain to the social judgment of motion-specific properties of a triangle-shape clip art.

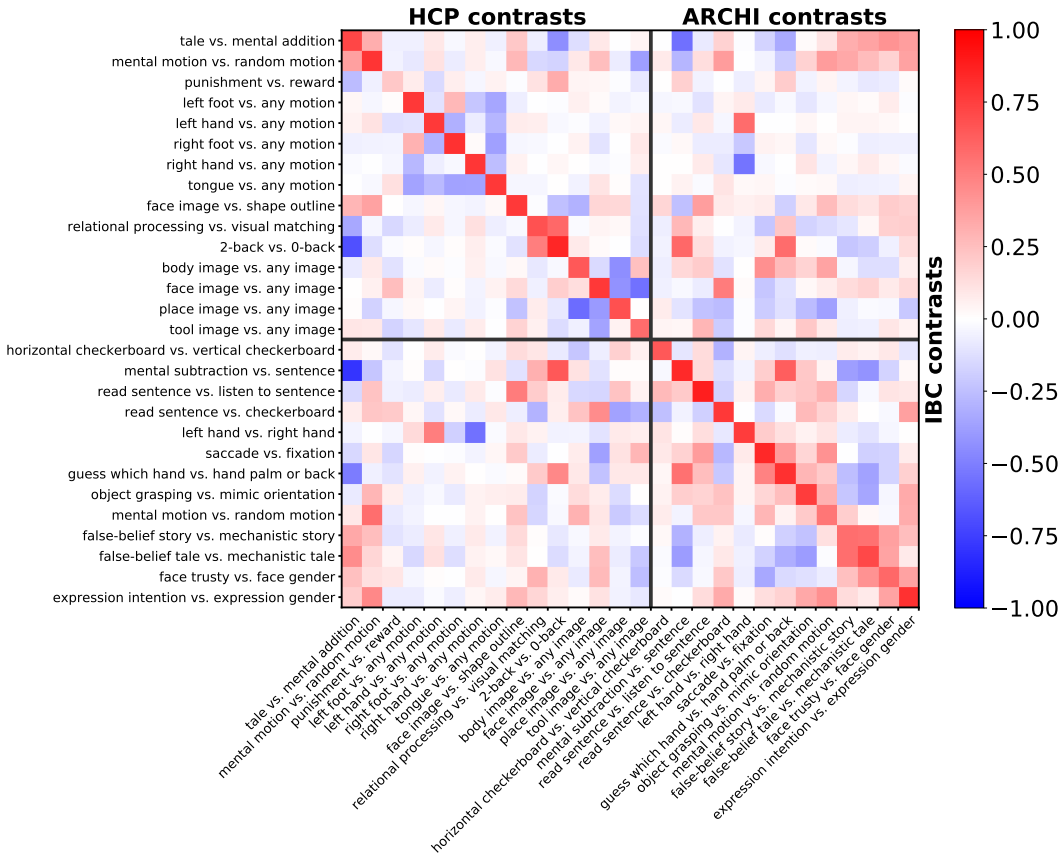

FIGURE 1 **Reproduction of previous large-cohort neuroimaging results**. The matrix displays the correlation between group-average *z*-maps of the main contrasts from different datasets, thus providing a quantitative assessment of the similarity between ARCHI and *Individual Brain Charting* IBC results as well as HCP and IBC results. With one exception, i.e. *punishement-reward* of the *HCP Gambling* task, the dominant diagonal terms hint at good reproduction.

3.2 | Variability of Functional Signatures within and between Participants

Figure 2 shows the individual functional signatures obtained for the Working Memory task adapted from the HCP protocol. It displays a qualitative visual description of the variability of such functional signatures across participants. Although the main features of functional responses are preserved at a coarse scale across individuals, there are clear differences in the precise locations of the active regions. For instance, this is the case for the *place image* versus *any*

images contrast that displays an active focus in the vicinity of the transverse occipital sulcus for all participants; yet, the size of the cluster and its MNI-space position vary conspicuously among them. Similarly, the *body image* versus *any image* and *face image* versus *any image* contrasts display activity along the fusiform gyrus, as well as *face image* versus *any image* and *tool image* versus *any image* in the lateral occipital complex. However, their expressions differ across subjects, resulting into different functional preferences at the corresponding locations in MNI space. It remains to be checked whether such variations represent intrinsic between-subjects variability or simply sampling noise.

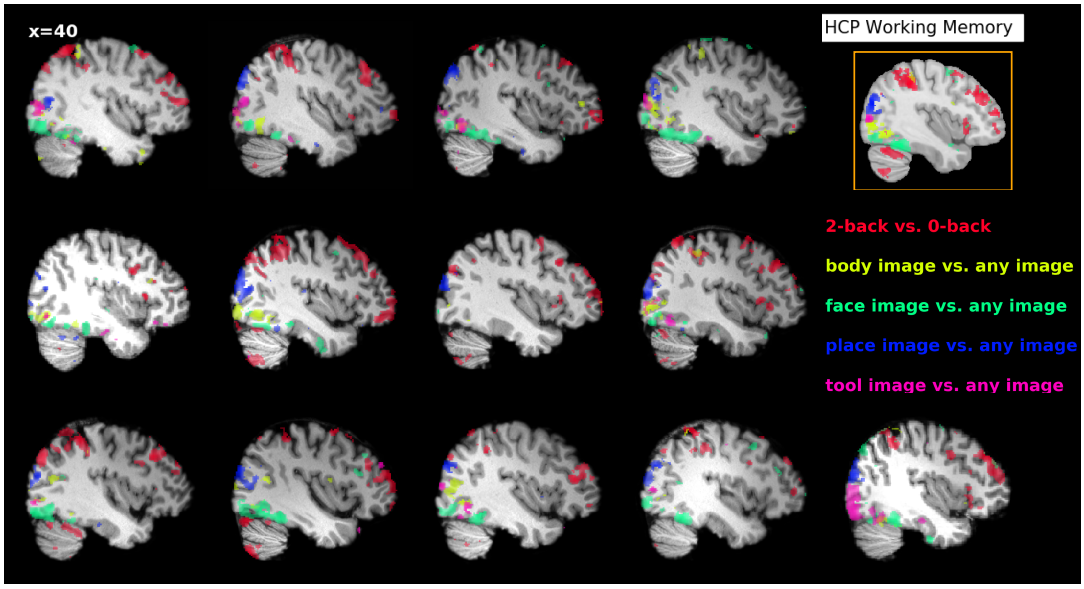

FIGURE 2 Activation maps of the contrasts estimated from the conditions in the HCP Working Memory task. Individual maps for fixed effects are displayed for every participant, using an FDR-corrected threshold $q = 0.05$. The group-level conjunction map of these individual maps is shown inside the orange frame. All maps correspond to the slice $x = 40mm$ in the sagittal view.

In turn, Figure 3 provides a quantitative characterization of the spatial consistency, within and between participants, of the main contrasts from each task. Consistency is measured by a similarity metric, namely Pearson correlation. The mean of the correlations between participants stands always above zero, revealing some level of spatial consistency among individual maps. Nevertheless, they are always below 0.5, which quantifies the level of noise of these maps. Importantly, the contrasts show different levels of overlap because of either the low fSNR or the large variability present in each contrast. For instance, contrasts from both ARCHI Social and HCP Social tasks, i.e. *false-belief story* versus *mechanistic story*, *false-belief tale* versus *mechanistic tale* and *social-interaction motion* versus *random motion*, as well as *punishment* versus *reward* from the HCP Gambling task exhibit lower consistency than tasks involved in language or perception.

On the other hand, spatial consistency is higher at the individual level in 27/33 contrasts displayed in Figure 3. Note that, since per-session data are used to compute the individual-maps consistency, the fSNR in every map is lower for the intrasubject consistency analysis than for the intersubject consistency. These correlations also reflect

imperfectly corrected PA/AP spatial distortions. Overall, the fact that maps are generally more consistent within subjects than across subjects hints at subject-specific topographies.

We have also compared intersubject consistency of the main-contrast maps (except those from the RSVP Language task) estimated from the IBC dataset and the original datasets, i.e. ARCHI and HCP datasets (see Figures C.1 and C.2 in Appendix section C). Results show that, in the large majority of the cases, considerably higher correlations were achieved using the IBC data than the ARCHI data as well as the HCP data. Intrasubject consistency was also evaluated between the IBC and HCP data and correlations are also overall higher among contrasts from the IBC dataset (see Appendix section C). They thus confirm that the IBC dataset hints at good levels of fSNR—in agreement with the results stated in Pinho et al. (2018)—as well as an overall high data quality, including an adequate parametrization of the MRI acquisitions.

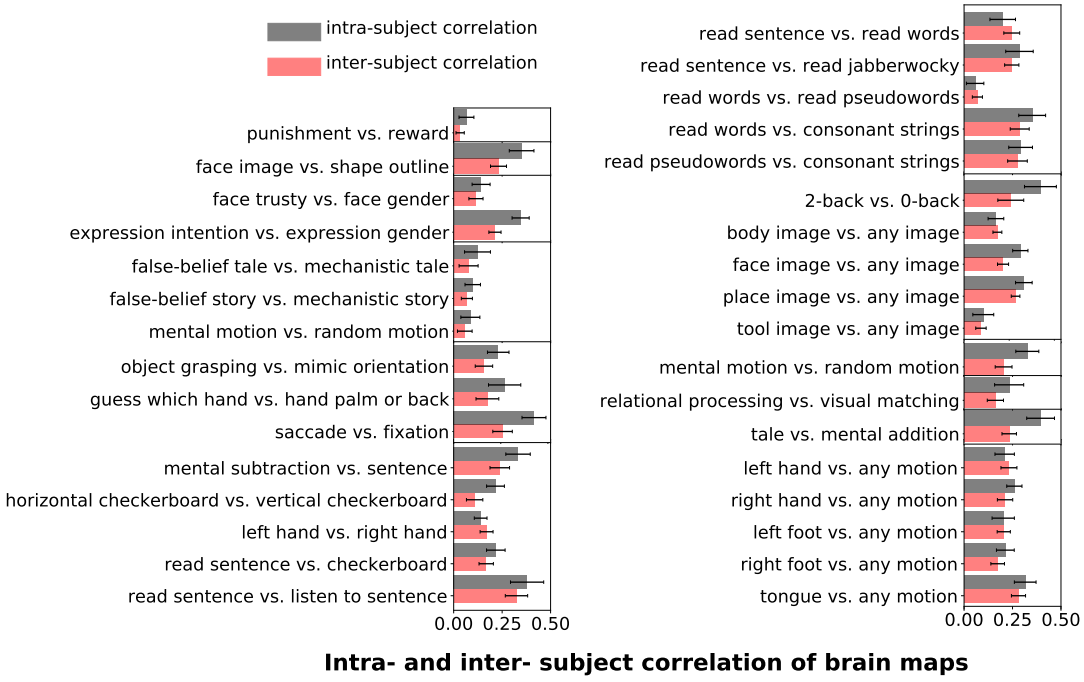

FIGURE 3 Comparison of intrasubject and intersubject variability across contrasts. The bar chart represents the means of the distributions of the correlations of *z*-maps per contrast within and between subjects. Bars in salmon correspond to correlations of *z*-maps per contrast, for all subject pairs; bars in gray correspond to correlations estimated from all possible pairs of "PA" and "AP" runs (see Section 2.5 for details). Error bars represent the 95% CI.

3.3 | A Dictionary of Cognitive Components

We then analyzed the topographies of twenty cognitive components obtained by decomposing, per subject, the original data of all tasks—i.e. 51 main-contrast maps—using sparse dictionary learning. These individual topographies are displayed as spatial maps (see top section of Figure 4), wherein each map represents the functional profile of a component in each subject and hemisphere (i.e. maps of the top-left and top-right panels refer to the left and right

hemispheres, respectively). According to S. M. Smith et al. (2009) and given the total number of IBC contrasts, $k = 20$ represents a consensual number of components to feature this study. A larger number of components was deliberately not chosen for the sake of readability of the present results. These components synthesize the cognitive information across all contrasts, making between-subject comparison easier. They were named according to the main conditions of the contrasts that obtained the largest value in their functional fingerprint. Such labeling was only possible due to the use of task data. The functional fingerprints of the components are provided as Supplementary Material.

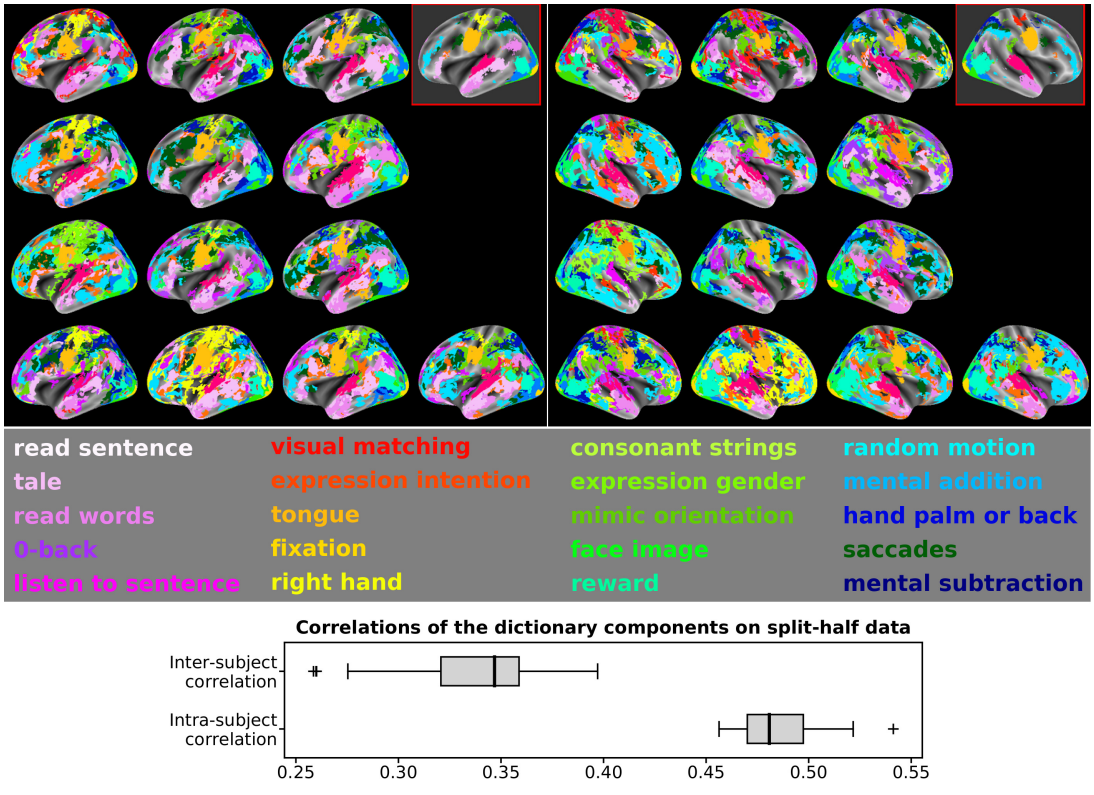

FIGURE 4 Dictionary of twenty cognitive components summarizing thirteen individual topographies in fsaverage space. *(Top left/right)* Labeling of left/right hemispheric cortical regions, according to the strongest dictionary loading in that region. The *top-right brain maps outlined in red* of each image display a median map obtained at group level, i.e. a label is assigned to a component if, at least, half of the participants have that label at that location. *(Middle)* The twenty cognitive components are labeled according to the contrast *z*-map that gets the maximum loading for that component. *(Bottom)* The boxplots represent the distribution of the intrasubject and intersubject stability of the components and whiskers show the 95% CI. Both intrasubject and intersubject data were split into two halves of the dataset, according to the phase-encoding direction parameter of acquisition. Dictionaries were estimated separately for each dataset. Split-half reproducibility of the spatial loadings was obtained, as their correlation, within and between subjects. While intrasubject correlations are relatively high, intersubject ones are low, thus hinting again at a strong subject effect. Black horizontal lines, going through the boxes, represent the median in both distributions. Overall, this figure illustrates the consistency of some components, while it outlines local and large-scale differences across the *Individual Brain Charting* (IBC) participants.

Since most of the activations depicted in the original contrast maps are located in the cortex, surface-based analysis was herein employed in order to enable a fine functional parcellation of the cortical structures from the individual topographies (see Section 2.7.1 for details).

The two red-framed brain maps, displayed at the right-top corner of each panel in Figure 4, show the median map of the components across individuals, i.e. each vertex has the label corresponding to the strongest component observed in—at least—half of the participants. These median maps thus provide a robust consensus model of the topographies obtained across subjects. Overall, tasks consistently map brain networks across participants. Nevertheless, one can still observe differences between subjects that are worthy of notice. For instance, components tagged as *math*, *saccades* and *consonant strings* are observed in similar brain regions across subjects, but one typically dominates the others. In other words, their spatial representation *compete* across subjects. While their outline is broadly similar, they reveal in detail notable differences at high-resolution.

To assess the stability of the dictionary, the dataset was split into two halves, according to the phase-encoding direction parameter. The split-half reproducibility of the twenty components was measured as correlations of the individual topographies both within and between subjects. The topographies were, in this case, obtained through the contrast *z*-maps estimated from the "PA" and "AP" runs of each task. Intrasubject stability was evaluated through the distribution of the correlations between PA- and AP-related topographies in every subject. On the other hand, intersubject stability was estimated through the distribution of correlations between PA- and AP-related topographies within pairs of subjects. The results, displayed at the bottom section of Figure 4, show higher correlations estimated from intrasubject data than intersubject data, highlighting that the variability of such spatial representations is linked to individual differences.

Finally, it is also worthy of note that these aggregated components are more stable than the original contrast maps (see Figure C.1 of Appendix section C). Overall, means of the distributions for both intrasubject and intersubject correlations of dictionary components (see bottom section of Figure 4) are higher than means of the distributions of *z*-maps correlations (see Figure 3).

3.4 | Reconstructing Functional Contrasts from different Tasks

Because different conditions across tasks share the same cognitive properties, the contrasts are expected to capture commonalities resulting from the corresponding functional activations. Such commonalities can be learnt and used to predict other contrasts from tasks supposedly sharing the same cognitive properties (see e.g. Tavor et al. (2016)). Thus, contrast maps from a target task (one contrast map at a time) are predicted from contrast maps of the training tasks. The $\mathbf{R}^2$–statistic was used to quantify the discrepancy between the predicted and actual maps, and the maximum of this statistic was computed across tasks.

Figure 5 - *top* shows that a good accuracy prediction was attained throughout all regions in the brain, except in the hippocampus, superior temporal asymmetrical pit (Leroy et al., 2015; Glasel et al., 2011), precuneus and inferior temporal gyrus. The few exceptions might be explained by either the lack of tasks targeting these regions or a reduced fSNR at these regions.

Several tasks were well predicted by other tasks (see Figure 5 - *bottom*). Indeed, tasks whose contrasts share cognitive properties with many contrasts from other tasks—like ARCHI Standard, RSVP Language and HCP Working Memory—received higher scores. Contrariwise, a task like HCP Motor, that yields not only on hand movements (common to other tasks) but also on foot and tongue movements—which are unique to this task—could not be as well predicted by the other tasks. At last, this analysis was repeated, but with subjects permuted between train and test, potentially breaking the fine-grained structure of brain responses. Results show a clear decrease of the proportion

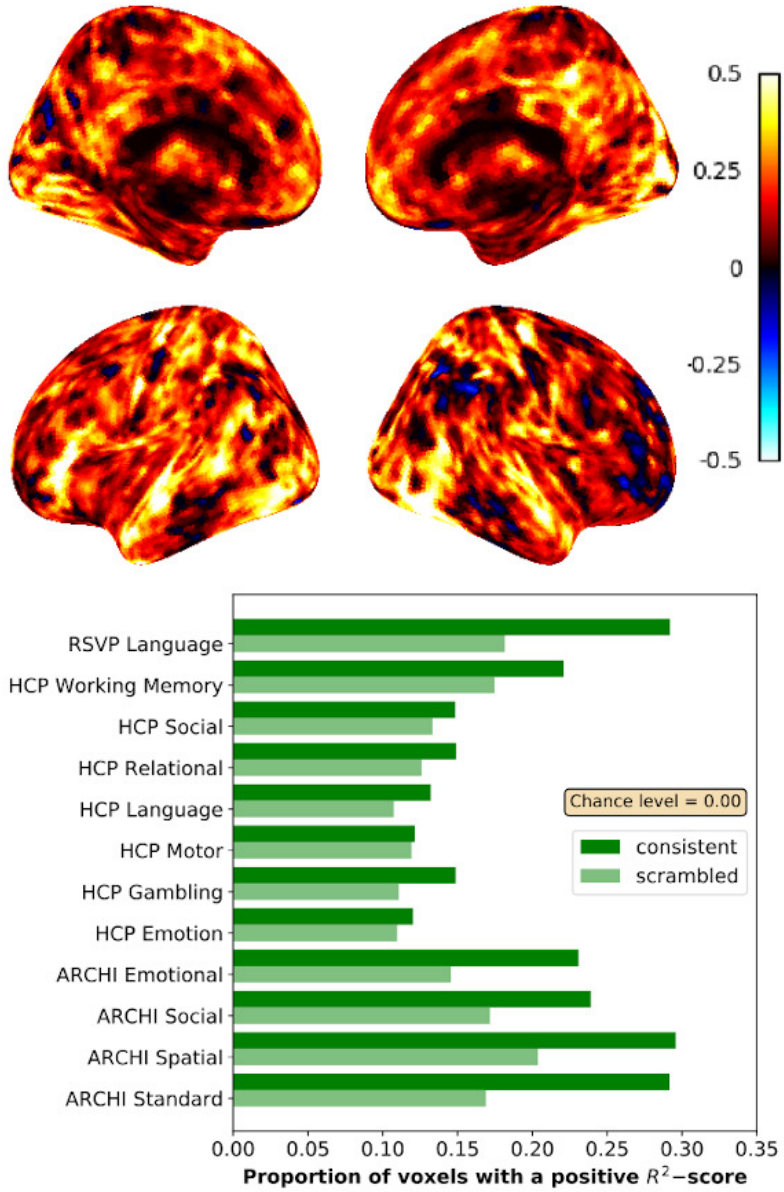

FIGURE 5 **Within-subject accuracy prediction of contrast maps.** *(Top)* Accuracy prediction (maximum R^2−score) obtained from a leave-three-subjects-out cross-validation experiment across tasks, in which contrast maps from a target task were predicted from the contrasts of the training tasks. These results quantify the amount of functional activity at every voxel that can be predicted, conditional on other contrasts. Most regions of the brain are covered by the predicted functional signatures, with a few exceptions: hippocampus, superior temporal asymmetrical pit, precuneus and inferior temporal gyrus. *(Bottom)* Proportion of voxels with a positive R^2−score per task indicates the size of brain regions functionally characterized by each task. Permutations of the subjects in the same analysis decrease this proportion in all tasks, showing that the captured topographies are subject-specific. Chance level is 0.00 in all tasks, for both "consistent" and "scrambled" schemes.

of the predicted voxels, confirming that topographies are subject-specific. Interestingly, such decrease was not very evident for the HCP Motor task, highlighting that low-level cognitive processes are less sensitive to spatial variability across individuals. Chance level was computed for all tasks with and without permutations of subjects between train and test. In all cases, its value was zero.

Overall, results from Figure 5 highlight two important findings. First, brain responses in different contrasts underlie a latent structure, allowing for the prediction of a response to a given task in a new subject; this structure is, in turn, conditional to other brain responses. Second, the topographic organization of these components is subject-specific.

3.5 | Functional Mapping of the Language Network

As several contrasts cover the same psychological domains, the IBC dataset provides the appropriate framework to derive a fine and unambiguous characterization of brain regions with respect to cognition. To showcase the benefits resulting from contrast-maps accumulation, we thus investigated the cognitive profile of a set of brain regions known to be linked to language mechanisms. These regions-of-interest refer to: *(a)* the left Inferior Frontal Gyrus (IFG) *pars orbitalis* and *(b) pars triangularis*, *(c)* the left Temporoparietal Junction, *(d)* the left Temporal Pole, *(e)* the left anterior and *(f* posterior Superior Temporal Sulcus (STS). They were subsequently made subject-specific using the dual-regression approach described in Section 2.7.8. The maps of their projections onto individual templates are provided in Figure D.1 of the Appendix section D.1. To illustrate the specific function of these regions, Figure 6 shows the activations elicited by—linearly independent—contrasts selected from the ARCHI Standard and RSVP Language tasks, whose effects-of-interest are related to language.

The two portions of the left IFG reveal greater responses to *read sentence* than *read words*, as highlighted in the selected contrast from the RSVP Language task, although effects are significant in both cases. Worthy of notice are the effects obtained in all regions for the contrast *sentence*[5] versus *mental subtraction* (reverse contrast of *mental subtraction* versus *sentence* in Figure 6), except in the left IFG *pars triangularis*. These results indicate higher level of activation in these regions, but in the left IFG *pars triangularis* for the general-domain semantic condition (i.e. *sentence*) than the math-specific semantic condition (i.e. *mental subtraction*). Indeed, Amalric & Dehaene (2016) reports several regions of the language network activating less for math statements than non-math statements. On the other hand, the left IFG *pars triangularis* was shown to be specifically implicated in arithmetic processes (Andin et al., 2015).

The left Temporoparietal Junction and left Temporal Pole show greater responses in the *read sentence* versus *read words* and *read sentence* versus *read jabberwocky* contrasts, revealing their involvement in the modulation of combinatorial semantics. In fact, the absence of activations in the left Temporoparietal Junction for the remaining contrasts highlights that this region is specifically responsive to sentence comprehension. Opposite effects in the *read words* versus *read pseudowords*, *read words* versus *consonant strings* and *read pseudowords* versus *consonant strings* contrasts indicate low responsiveness to lists of words. Additionally, the left Temporal Pole also exhibits a high level of activation in the *listen sentence* versus *read sentence* contrast, which reflects some sensitivity to auditory stimuli.

At last, the left STS exhibits overall strong effects to different levels of specificity in language mechanisms, given its high level of activation in most of the contrasts. Yet, its posterior portion is more responsive to list of words than the anterior, as shown in the contrasts *read words* versus *read pseudowords*, *read words* versus *consonant strings* and *read pseudowords* versus *consonant strings*. Contrariwise, its anterior portion is more responsive to combinatorial semantics than the posterior, as shown by the contrasts *read sentence* versus *read words* and *read sentence* versus *read jabberwocky*. Similarly to the left Temporal Pole, both portions of the left STS are responsive to auditory stimuli.

All regions were equally sensitive to the *read sentence* versus *checkerboard* contrast, highlighting the low level of specificity of this contrast for language-related mechanisms.

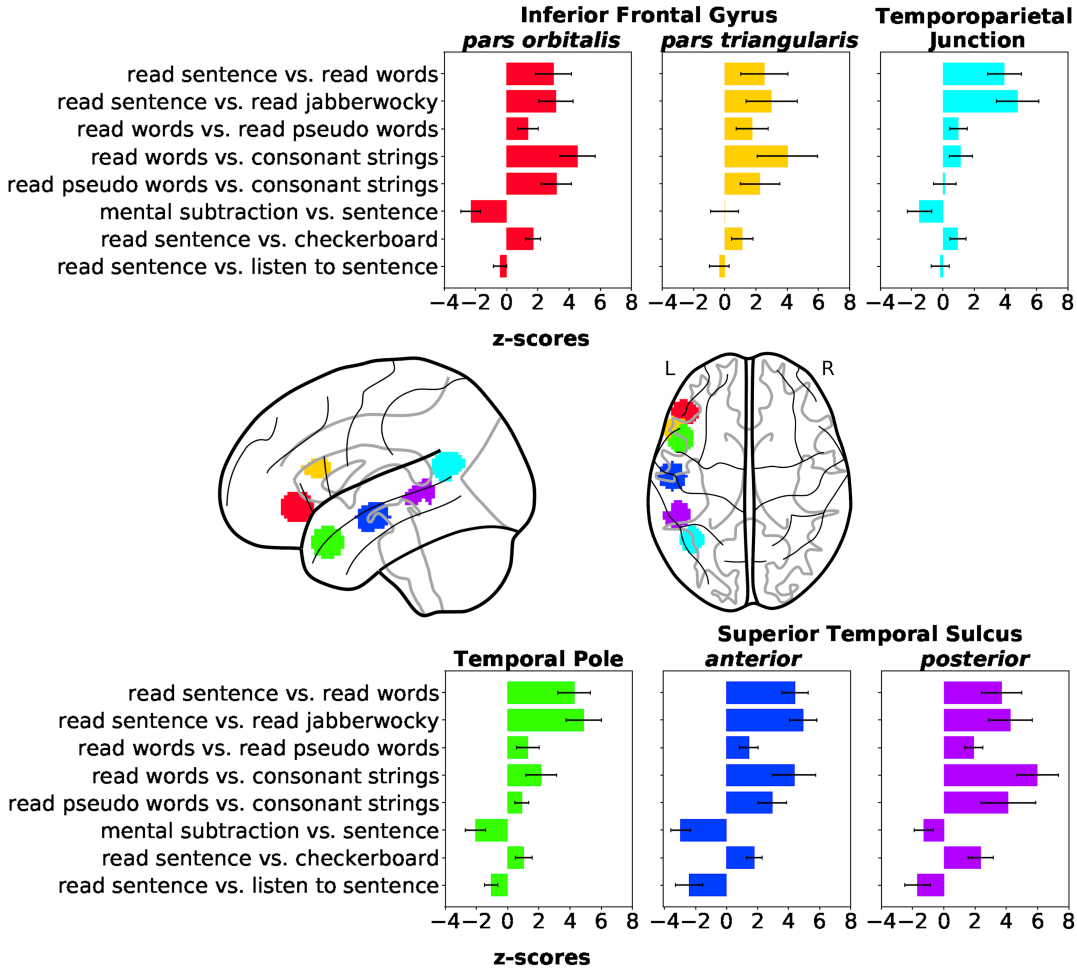

FIGURE 6 **Comparison of the cognitive profile for a set of regions-of-interest belonging to the language network**. This figure illustrates how cognitive profiles of functional regions can be derived from their quantitative contribution in a set of independent contrasts pertaining to language mechanisms. The direct comparison of such contributions highlights the dissociation of the role of these regions in lower-order and higher-order semantic processing, given respectively their contribution either in the contrasts "read words versus read pseudo words", "read words versus read consonant strings" plus "read pseudo words versus consonant strings" or in the contrasts "read sentence versus read words", "read sentence versus read *jabberwocky*" plus "sentence versus mental subtraction". Each bar plot represents, for a set of contrast *z*-maps, the means of the distributions across subjects of the average of *z*-scores in a set of voxels inside a regions-of-interest. Error bars represent the 95% CI. Bar colors identify each of the six regions-of-interest, placed in the left hemisphere, that are in evidence on the glass brain: *(red)* Inferior Frontal Gyrus (IFG) *pars orbitalis*, *(yellow)* IFG *pars triangularis*, *(cyan)* Temporoparietal Junction, *(green)* Temporal Pole, *(dark blue)* anterior Superior Temporal Sulcus (STS), and *(purple)* posterior STS.

Overall, the variety of contrasts referring to language highlights differences in the functional specialization of the aforementioned regions-of-interest. To assess the statistical significance of these differences, we tested whether voxels could be reliably classified as belonging to one of these regions, based on their functional activation in the given contrasts. Figure 7 presents the prediction scores obtained between pairs of regions against their baselines, i.e. the corresponding chance levels of such predictions. All scores outperformed the chance level, although the classifier between IPS *pars orbitalis/pars triangularis* was close to chance. In order to generalize our conclusions to other cognitive domains and demonstrate the benefits of accumulated task data, the same experiment was also performed with data extracted from all main-contrast maps obtained from the entire dataset. The decoding performance was slightly higher (see Figure D.2 in Appendix section D.2, which shows that including a wider range of cognitive domains improved the functional characterization of the regions-of-interest.

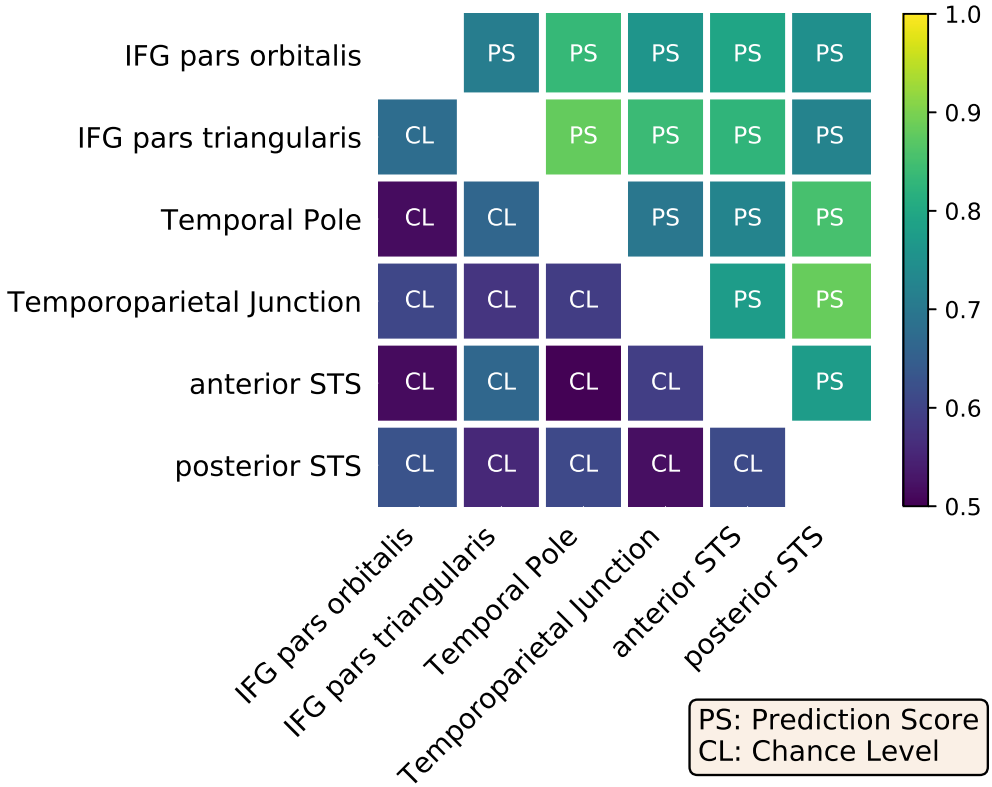

FIGURE 7 Accuracy obtained for the classification of voxels into pairs of regions-of-interest, against chance level. The scores were estimated based on the functional activation of voxels from different pairs of regions-of-interest pertaining to the language network. Combinations of two brain regions from a total amount of six regions from the left hemisphere—Inferior Frontal Gyrus (IFG) *pars orbitalis*, IFG *pars triangularis*, Temporal Pole, Temporoparietal Junction, anterior Superior Temporal Sulcus (STS) and posterior STS—were employed in a leave-one-subject-out cross-validation experiment to calculate classification accuracy. The functional profiles of the voxels were extracted from contrasts of the *Individual Brain Charting* (IBC) dataset that relate to language mechanisms (see Figure 6).

To further evaluate the generalizability of these results, we conducted the same analysis in another set of regions-of-interest, also pertaining to the language network but extracted from functional signatures obtained from a larger cohort. Concretely, the group-level contrast *story* versus *math* of the HCP Language task (for more information about this task, consult Section 2.4.2), referring to a subset of the HCP900 dataset (n=786), was used to define functional regions participating in language-related mechanisms. A complete description of the procedures concerning regions' extraction can be found in the Appendix section D.3. This set of regions-of-interest partially overlaps with the previous one. The left Temporoparietal Junction was clearly isolated. Besides, it was possible to extract the left anterior STS as well as the left Temporal Pole but we were not able to dissociate both into two distinct regions; similarly, the results allowed the extraction of the Inferior Frontal Gyrus, yet comprising both *pars orbitalis* and *pars triangularis*. We further identified the Ventromedial Prefrontal Cortex, usually reported to be involved in semantic integration. In summary, the cognitive profiles obtained are consistent with the previous ones (see Figure D.3 in Appendix section D.3) and differences between them are statistically significant, since all prediction scores between pairs of regions were higher than chance (see Figure D.4 in Appendix section D.3.

The taskwise structure of the IBC dataset thus opens a new type of analysis, whereby brain-region profiles can be compared in order to highlight their functional specificity.

4 | DISCUSSION

High-resolution brain images allow for fine-grained contrast mapping and detailed characterization of brain networks at the individual level. However, these maps only overlap partially at the group level, making it hard to build a functional template from all of them. On the other hand, system-level analysis of cognitive functions requires pooling data from multiple tasks.

This article lays the ground for individual functional atlasing, a novel analytic approach that requires a specific strategy to avoid circular reasoning, i.e. *delineating different topographies from functional contrasts is not sufficient to state that subjects differ regarding these contrasts*. Because it aims at providing quantitative insights about individual differences of elementary processes in cognition, *deep phenotyping* of behavioral responses becomes crucial in such approach. To this end, it is thus important to ensure that individual data are comparable to existing resources, i.e. that the contrasts used herein reproduce those of previous large-scale studies. Thence, an essential step is to analyze the consistency of functional signatures across regions. After a qualitative inspection of some contrasts in a given task, that provided propitious results, we scaled up the analysis; a relatively coarse, network-level structure was extracted from all tasks using dictionary learning, as means to obtain a global picture of the between-subject consistency of the functional patterns. Third, we controlled the predictive power of these individual topographies by reconstructing contrast maps organized by tasks, using cross-validation across individuals. Successful predictions, in particular within subjects, strictly confirms the above intuitions. We then illustrated the practical significance of this individualized mapping approach, by showing how to adapt population-level regions-of-interest to individuals. This framework allows for probing unique functional profiles associated with specialized brain regions, giving the flexibility of adapting to an individual configuration and assessing functional specialization across individuals, while avoiding circular reasoning. We review all these steps in further detail over the next sections.

A small-n, yet reliable functional mapping

Fifty one main-contrast maps were investigated in this cross-task analysis, yielding a unique high-coverage cognitive mapping. Despite the low number of subjects, group-average results are consistent with those from previous larger-

scale datasets, as highlighted by the results presented on Section 3.1.

Spatial variability of functional signatures

High-resolution, individual contrast maps provide evidence of the intersubject spatial variability inherent to functional signatures associated with cognitive processes. In Figure 2, one can see e.g. the differences across participants in the cluster extent placed on a subsection of the transversal occipital sulcus, present in the contrast *place image* versus *any images*. This variability could be mitigated by normalization techniques relying on macro-anatomical landmarks (Frost & Goebel, 2012) rather than MNI coordinates; this matter is subject of further investigation. Importantly, the number and diversity of contrasts allow us to quantify such variability within and between participants (see Figure 3). For instance, it shows which cognitive domains are more stable between participants. Contrasts isolating vision processes, motor performance and low-level of semantic comprehension in language (e.g. manipulation of words) are more consistent across individuals than contrasts pertaining to social cognition or theory-of-mind. Besides, spatial consistency can also be studied within subjects for every task, showing which brain networks display a low fSNR, due to e.g. geometric distortions inherently associated with MRI acquisition parameters. Overall, intrasubject consistency is higher than intersubject consistency, which hints at subject-specific topographies.

Topographies of cognitive components

The taskwise structure of the IBC dataset can support the investigation of common functional profiles or fingerprints and link them to experimental tasks. Here, we use dictionary learning to extract these functional profiles in an automatic fashion. Note that, other approaches would also be possible, such as independent components analysis or clustering. Our data-driven analysis, exploiting individual topographies across tasks, highlight the intersubject variability inherent to cognitive mapping (see Figure 4). Similarly to the results in Section 3.2, intrasubject consistency was higher than intersubject consistency. Interestingly, the mean of intrasubject correlation for all components (see bottom section of Figure 4) is about twice the intrasubject consistency registered for the contrast maps themselves (see Figure 3). Obviously, more contrasts lead to less noisy, hence more reproducible components.

Note that we chose here to arbitrarily estimate twenty components to represent the data. There is no reason to believe that this is an optimal number. We defer a more thorough discussion on this topic to future work, bearing in mind that the level of description always depends on the scientific question posed. We also note some alternative modeling choices that have been proposed in the literature, e.g. relying on group structure (Varoquaux et al., 2013) or spatial regularization (Abraham et al., 2013). These algorithms contain some hyperparameters that are hard to set in practice, and their convergence is computationally challenging. Hence, we have decided to rely on a simpler, data-driven approach, expecting that the accumulation of contrasts obviates the need for regularization terms.

At last, we opted to map the individual topographies on the surface, since most of the activations in the original volume-based contrasts are located in the cortex. The importance of surface-based mapping comes to the forefront when one aims at individual-specific localization in the upper cortical areas of the brain and, thus, surface-based cortical registration becomes relevant toward setting a common system of topographical representation on the surface. Recent methods based on geometric features derived from functional and diffusion imaging, such as Spherical Daemons (Yeo et al., 2010) or Multimodal Surface Matching (Robinson et al., 2014), are promising in regards to deliver a better surface alignment because of a greater sensitivity to a wider variety of surface descriptors. Therefore, we acknowledge the possibility to integrate, in the future, such steps in the preprocessing pipeline.

Prediction of contrasts within the dataset

Since many contrasts share the same cognitive components, they display overlapping functional patterns. This shared variability can be learnt and used to predict other contrasts. Such a successful prediction reveals the existence of a latent structure underlying them (see Figure 5 - bottom). Additionally, predictions are higher when subject identity is not randomly shuffled, highlighting once again that functional anatomy varies across subjects.

Interestingly, this experiment, together with the reproduction of contrast maps from the ARCHI and HCP datasets (see Figure 1), implies that some brain maps could be *transferred* from IBC to these aforementioned datasets. For instance, we hypothesize the possibility of predicting the topographies of the HCP participants from the individual contrast maps of the IBC-ARCHI or RSVP-language tasks.

Cognitive regional profiles of the language network

The taskwise organization of the dataset also opens the possibility to investigate the functional profiles of brain regions linked to cognitive domains covered by the contrasts (see Section 3.5). For instance, the results shown in Section 3.5 can identify which regions elicit effects modulated by word-level (e.g. left posterior STS) and sentence-level semantics (e.g. left Temporoparietal Junction and left Temporal Pole). On the other hand, effects modulated by syntax are difficult to isolate in the context of brain mapping. Some studies have suggested that syntactic processing is distributed throughout an ensemble of brain regions that support high-level linguistic processing. Therefore, they cannot be separated from other aspects regarding language comprehension, like lexico-semantic processing (Blank et al., 2016), overall suggesting that the language network might be more strongly concerned with meaning than structure (Siegelman et al., 2019). Yet, as an attempt to address this issue, we plan to integrate, in a future release, a naturalistic language-comprehension paradigm, dedicated to syntactic-composition modulation (Bhattasali et al., 2019). The more tasks are included in the dataset, the greater the richness of the contrasts that can be used to not only disambiguate the cognitive role of functional regions but also delineate finer demarcations of their anatomical boundaries.

Limitations

According to the structure of the IBC dataset, the analysis is limited to thirteen subjects. This is not enough if one were to report populations effects, yet it leaves the possibility of conducting parallel analysis on thirteen unique brains and highlight commonalities among them. The question of deriving a sensible group-level model for this dataset remains thus open.

Toward more individualized models of brain function

One possible extension of the present work relates to the development and testing of feature-based alignment approaches (Sabuncu et al., 2010; Haxby et al., 2011) that can match specific properties of functional brain regions, such as cognitive responses or connectivity. They can thus be used to improve the estimation of functional templates in the presence of conspicuous between-subject variability, as demonstrated in the recent work by Bazeille et al. (2019). Additionally, Bijsterbosch et al. (2018) has shown that most of the relevant cross-subject differences, identified through brain imaging, are rooted in the topography of their individual maps. These findings suggest that future investigations on the latent structure of individual contrast maps may contribute to improving the analysis of between-subject characteristics in brain imaging. Nevertheless, an important question still remains: what is the most appropriate method to achieving these correspondences?

Ongoing taskwise development of the Individual Brain Charting dataset

The collection of new data in IBC is ongoing until year 2022 and more releases are planned over the upcoming years. The second and third releases have recently been made publicly available—source data are available in OpenNeuro under the data accession *ds002685*[6] and derivatives in NeuroVault with the *id* collection 6618[7]—and they comprise a variety of cognitive tasks. Like the first release, the second release covers tasks on both lower-order and higher-order cognitive functions, such as mental-time travel (Gauthier & van Wassenhove, 2016a,b), positive-incentive value (Lebreton et al., 2015), theory-of-mind and pain matrices (Dodell-Feder et al., 2011; Jacoby et al., 2016; Richardson et al., 2018), numerosity (Knops et al., 2014), self-reference effect (Genon et al., 2014), and speech recognition (Campbell et al., 2015). The third release is dedicated instead to an extensive data collection tackling the visual system; it is concerned with tasks on visualization of naturalistic scenes (Huth et al., 2012), classic retinotopy, and movie watching (Haxby et al., 2011). The fourth release is currently in preparation and it includes a tonotopy task (Santoro et al., 2017) as well as two batteries of localizers (Perrone-Bertolotti et al., 2012; Saignavongs et al., 2017; Ossandón et al., 2012; Hamamé et al., 2012; Vidal et al., 2010; Fan et al., 2002; Eriksen & Eriksen, 1974; Figner et al., 2009; Kirby & Maraković, 1996; Otto et al., 2014; Bissett & Logan, 2011; Aron et al., 2007; Crump et al., 2013; Schneider & Logan, 2011; Kaller et al., 2011; Ward & Allport, 1997; Shallice, 1982; Stroop, 1935) addressing many other cognitive modules, such as: stimulus salience, working memory, visual object categorization, audio perception, risk-associated decision making, motor inhibition, planning and vigilance. Our ultimate goal is to achieve a comprehensive brain coverage of functional signatures, associated with a large variety of mental functions, by the end of the project.

5 | CONCLUSION

Our results show that the application of rich taskwise datasets is necessary in studies concerned with cognitive mapping and individual modeling of the human brain.

Tasks adapted from former projects are in overall agreement with the ones reported by the original studies. Therefore, this study advocates the principles of data-sharing and reproducibility in neuroscience, as means to achieving transparency in research practice and consistency of results across time. Raw data and data derivatives of the first release of the IBC dataset have been made available in the public repositories of OpenNeuro and NeuroVault, respectively. In particular, individual and unthresholded *z*-maps in NeuroVault were extensively used to deliver the results of the present article. They are thus intended to serve as a valuable tool for the community and complement mega-analytic and meta-analytic studies dedicated to the functional examination of cognition in the human brain.

References

Abraham, A., Dohmatob, E., Thirion, B., Samaras, D., & Varoquaux, G. (2013, September). Extracting brain regions from rest fMRI with Total-Variation constrained dictionary learning. In *MICCAI - 16th International Conference on Medical Image Computing and Computer Assisted Intervention - 2013*. Nagoya, Japan: Springer. Retrieved from https://www.doi.org/10.1007/978-3-642-40763-5_75

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., … Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Front Neuroinform*, *8*, 14. Retrieved from https://doi.org/10.3389/fninf.2014.00014

Amalric, M., & Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. *Proc Natl Acad Sci U S A*, *113*(18), 4909–4917. Retrieved from https://doi.org/10.1073/pnas.1603205113

Andersson, J. L., 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. Retrieved from http://doi.org/10.1016/S1053-8119(03)00336-7

Andin, J., Fransson, P., Rönnberg, J., & Rudner, M. (2015). Phonology and arithmetic in the language–calculation network. *Brain Lang*, *143*, 97 - 105. Retrieved from http://doi.org/10.1016/j.bandl.2015.02.004

Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., & Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *J Neurosci*, *27*(14), 3743–3752. Retrieved from https://doi.org/10.1523/JNEUROSCI.0519-07.2007

Ashburner, J., & Friston, K. (1997, October). Multimodal Image Coregistration and Partitioning - A Unified Framework. *Neuroimage*, *6*(3), 209-217. Retrieved from https://doi.org/10.1006/nimg.1997.0290

Ashburner, J., & Friston, K. J. (2005, July). Unified segmentation. *Neuroimage*, *26*(3), 839-851. Retrieved from https://doi.org/10.1016/j.neuroimage.2005.02.018

Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M., … Van Essen, D. C. (2013, October). Function in the human connectome: Task-fMRI and individual differences in behavior. *Neuroimage*, *80*, 169-89. Retrieved from https://doi.org/10.1016/j.neuroimage.2013.05.033

Bazeille, T., Richard, H., Janati, H., & Thirion, B. (2019). Local Optimal Transport for Functional Brain Template Estimation. In A. C. S. Chung, J. C. Gee, P. A. Yushkevich, & S. Bao (Eds.), *Information Processing in Medical Imaging* (Vol. 11492, pp. 237–248). Cham: Springer International Publishing. Retrieved from https://doi.org/10.1007/978-3-030-20351-1_18

Behzadi, Y., Restom, K., Liau, J., & Liu, T. T. (2007). A component based noise correction method (compcor) for {BOLD} and perfusion based fMRI. *Neuroimage*, *37*(1), 90 - 101. Retrieved from https://doi.org/10.1016/j.neuroimage.2007.04.042

Bhattasali, S., Fabre, M., Luh, W.-M., Saied, H. A., Constant, M., Pallier, C., … Hale, J. (2019). Localising memory retrieval and syntactic composition: an fMRI study of naturalistic language comprehension. *Lang Cogn Neurosci*, *34*(4), 491-510. Retrieved from https://doi.org/10.1080/23273798.2018.1518533

Bijsterbosch, J. D., Woolrich, M. W., Glasser, M. F., Robinson, E. C., Beckmann, C. F., Van Essen, D. C., … Smith, S. M. (2018, feb). The relationship between spatial configuration and functional connectivity of brain regions. *ELife*, *7*, e32992. Retrieved from https://doi.org/10.7554/eLife.32992

Bissett, P. G., & Logan, G. D. (2011). Balancing cognitive demands: control adjustments in the stop-signal paradigm. *J Exp Psychol Learn Mem Cogn*, *37*(2), 392. Retrieved from https://doi.org/10.1037/a0021800

Biswal, B., Mennes, M., Zuo, X., Gohel, S., Kelly, C., Smith, S., … Milham, M. (2010, 3). Toward discovery science of human brain function. *Proc Natl Acad Sci U S A*, *107*(10), 4734–4739. Retrieved from https://doi.org/10.1073/pnas.0911855107

Blank, I., Balewski, Z., Mahowald, K., & Fedorenko, E. (2016). Syntactic processing is distributed across the language system. *Neuroimage*, *127*, 307 - 323. Retrieved from https://doi.org/10.1016/j.neuroimage.2015.11.069

Book, G. A., Stevens, M. C., Assaf, M., Glahn, D. C., & Pearlson, G. D. (2016). Neuroimaging data sharing on the neuroinformatics database platform. *Neuroimage*, *124*, *Part B*, 1089 - 1092. Retrieved from https://doi.org/10.1016/j.neuroimage.2015.04.022 (Sharing the wealth: Brain Imaging Repositories in 2015)

Braga, R. M., & Buckner, R. L. (2017, July). Parallel Interdigitated Distributed Networks within the Individual Estimated by Intrinsic Functional Connectivity. *Neuron*, *95*(2), 457-471.e5. Retrieved from https://doi.org/10.1016/j.neuron.2017.06.038

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*, *14*(5), 365–376. Retrieved from https://doi.org/10.1038/nrn3475

Campbell, K. L., Shafto, M. A., Wright, P., Tsvetanov, K. A., Geerligs, L., Cusack, R., … Tyler, L. K. (2015). Idiosyncratic responding during movie-watching predicted by age differences in attentional control. *Neurobiol Aging*, *36*(11), 3045 - 3055. Retrieved from https://doi.org/10.1016/j.neurobiolaging.2015.07.028

Carp, J. (2012). On the Plurality of (Methodological) Worlds: Estimating the Analytic Flexibility of fMRI Experiments. *Front Neurosci*, *6*, 149. Retrieved from https://doi.org/10.3389/fnins.2012.00149

Chang, N., Pyles, J., Marcus, A., Gupta, A., Tarr, M., & Aminoff, E. (2019, 5). BOLD5000, a public fMRI dataset while viewing 5000 visual images. *Sci Data*, *6*, 49. Retrieved from https://doi.org/10.1038/s41597-019-0052-3

Costafreda, S. (2009). Pooling fMRI data: meta-analysis, mega-analysis and multi-center studies. *Front Neuroinform*, *3*, 33. Retrieved from https://doi.org/10.3389/neuro.11.033.2009

Costafreda, S. (2011, 12). Meta-Analysis, Mega-Analysis, and Task Analysis in fMRI Research. *Philosophy, Psychiatry, & Psychology*, *18*, 275-277. Retrieved from https://doi.org/10.1353/ppp.2011.0049

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS One*, *8*(3). Retrieved from https://doi.org/10.1371/journal.pone.0057410

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., … Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, *31*(3), 968 - 980. Retrieved from http://doi.org/10.1016/j.neuroimage.2006.01.021

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a theory of mind task. *Neuroimage*, *55*(2), 705 - 712. Retrieved from https://doi.org/10.1016/j.neuroimage.2010.12.040

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Percept Psychophys*, *16*(1), 143–149. Retrieved from https://doi.org/10.3758/BF03203267

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *J Cogn Neurosci*, *14*(3), 340–347. Retrieved from https://doi.org/10.1162/089892902317361886

Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Front Psychol*, *5*, 335. Retrieved from https://doi.org/10.3389/fpsyg.2014.00335

Fedorenko, E., Behr, M. K., & Kanwishera, N. (2011, December). Functional specificity for high-level linguistic processing in the human brain. *Proc Natl Acad Sci U S A*, *108*(39), 16428-33. Retrieved from https://doi.org/10.1073/pnas.1112937108

Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-Selective and Domain-General Regions Lie Side by Side within Broca's Area. In *Curr biol* (Vol. 22, p. 2059-62). Retrieved from https://doi.org/10.1016/j.cub.2012.09.011

Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Glasser, M. F., … Yacoub, E. (2010, 12). Multiplexed Echo Planar Imaging for Sub-Second Whole Brain fMRI and Fast Diffusion Imaging. *PLoS One*, *5*(12), 1-11. Retrieved from http://doi.org/10.1371/journal.pone.0015710

Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberative processes in risky choice: age differences in risk taking in the Columbia Card Task. *J Exp Psychol Learn Mem Cogn*, *35*(3), 709. Retrieved from https://doi.org/10.1037/a0014983

Fischl, B., Sereno, M. I., Tootell, R. B., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Hum Brain Mapp*, *8*(4), 272-284. Retrieved from https://doi.org/10.1002/(SICI)1097-0193(1999)8:4<272::AID-HBM10>3.0.CO;2-4

Frazier, J. A., Chiu, S., Breeze, J. L., Makris, N., Lange, N., Kennedy, D. N., … Biederman, J. (2005, 2005 Jul). Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder. *Am J Psychiatry*, *162*(7), 1256-65. Retrieved from http://doi.org/10.1176/appi.ajp.162.7.1256

Friston, K., Fletcher, P., Josephs, O., Holmes, A., Rugg, M., & Turner, R. (1998). Event-related fMRI: Characterizing differential responses. *Neuroimage*, *7*(1), 30 - 40. Retrieved from http://doi.org/10.1006/nimg.1997.0306

Friston, K., Frith, C., Frackowiak, R., & Turner, R. (1995). Characterizing Dynamic Brain Responses with fMRI: a Multivariate Approach. *Neuroimage*, *2*(2), 166-172. Retrieved from https://doi.org/10.1006/nimg.1995.1019

Friston, K., Josephs, O., Rees, G., & Turner, R. (1998, January). Nonlinear event-related responses in fMRI. *Magn Reson Med*, *39*(1), 41-52. Retrieved from https://doi.org/10.1002/mrm.1910390109

Frost, M. A., & Goebel, R. (2012). Measuring structural–functional correspondence: Spatial variability of specialised brain regions after macro-anatomical alignment. *Neuroimage*, *59*(2), 1369 - 1381. Retrieved from https://doi.org/10.1016/j.neuroimage.2011.08.035

Gauthier, B., & van Wassenhove, V. (2016a). Cognitive mapping in mental time travel and mental space navigation. *Cognition*, *154*, 55 - 68. Retrieved from https://doi.org/10.1016/j.cognition.2016.05.015

Gauthier, B., & van Wassenhove, V. (2016b). Time Is Not Space: Core Computations and Domain-Specific Networks for Mental Travels. *J Neurosci*, *36*(47), 11891–11903. Retrieved from https://doi.org/10.1523/JNEUROSCI.1400-16.2016

Genon, S., Bahri, M. A., Collette, F., Angel, L., d'Argembeau, A., Clarys, D., … Bastin, C. (2014). Cognitive and neuroimaging evidence of impaired interaction between self and memory in Alzheimer's disease. *Cortex*, *51*, 11 - 24. Retrieved from https://doi.org/10.1016/j.cortex.2013.06.009

Glasel, H., Leroy, F., Dubois, J., Hertz-Pannier, L., Mangin, J., & Dehaene-Lambertz, G. (2011). A robust cerebral asymmetry in the infant brain: The rightward superior temporal sulcus. *Neuroimage*, *58*(3), 716 - 723. Retrieved from https://doi.org/10.1016/j.neuroimage.2011.06.016

Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., … van Essen, D. C. (2016, July 20). A multi-modal parcellation of human cerebral cortex. *Nature*, *536*(7615), 171-178. Retrieved from http://doi.org/10.1038/nature18933

Goldstein, J. M., Seidman, L. J., Makris, N., Ahern, T., O'Brien, L. M., Jr., V. S. C., … Tsuang, M. T. (2007). Hypothalamic Abnormalities in Schizophrenia: Sex Effects and Genetic Vulnerability. *Biol Psychiatry*, *61*(8), 935 - 945. Retrieved from http://doi.org/10.1016/j.biopsych.2006.06.027

Gordon, E. M., Laumann, T. O., Gilmore, A. W., Newbold, D. J., Greene, D. J., Berg, J. J., … Dosenbach, N. U. (2017). Precision Functional Mapping of Individual Human Brains. *Neuron*, *95*(4), 791 - 807.e7. Retrieved from https://doi.org/10.1016/j.neuron.2017.07.011

Gorgolewski, K., Auer, T., Calhoun, V., Craddock, C., Das, S., Duff, E., … Poldrack, R. (2016, June). The brain imaging data structure: a standard for organizing and describing outputs of neuroimaging experiments. *Sci Data*, *3*, 160044. Retrieved from http://doi.org/10.1038/sdata.2016.44

Gorgolewski, K., Burns, C. D., Madison, C., Clark, D., Halchenko, Y. O., Waskom, M. L., & Ghosh, S. S. (2011). Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Front Neuroinform*, *5*, 13. Retrieved from http://doi.org/10.3389/fninf.2011.00013

Gorgolewski, K., Varoquaux, G., Rivera, G., Schwartz, Y., Sochat, V., Ghosh, S., … Poldrack, R. (2016). NeuroVault.org: A repository for sharing unthresholded statistical maps, parcellations, and atlases of the human brain. *Neuroimage*, *124*, *Part B*, 1242 - 1244. Retrieved from https://doi.org/10.1016/j.neuroimage.2015.04.016 (Sharing the wealth: Brain Imaging Repositories in 2015)

Gorgolewski, K., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S., Maumet, C., … Margulies, D. (2015). NeuroVault.org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Front Neuroinform*, *9*, 8. Retrieved from https://doi.org/10.3389/fninf.2015.00008

Gurevitch, J., Koricheva, J., Nakagawa, S., & Stewart, G. (2018). Meta-analysis and the science of research synthesis. *Nature*, *555*, 175-182. Retrieved from http://doi.org/10.1038/nature25753

Hamamé, C. M., Vidal, J. R., Ossandón, T., Jerbi, K., Dalal, S. S., Minotti, L., … Lachaux, J.-P. (2012). Reading the mind's eye: online detection of visuo-spatial working memory and visual imagery in the inferior temporal lobe. *Neuroimage*, *59*(1), 872–879. Retrieved from https://doi.org/10.1016/j.neuroimage.2011.07.087

Hanke, M., Adelhöfer, N., Kottke, D., Iacovella, V., Sengupta, A., Kaule, F. R., … Stadler, J. (2016). A studyforrest extension, simultaneous fMRI and eye gaze recordings during prolonged natural stimulation. *Sci Data*, *3*. Retrieved from https://doi.org/10.1038/sdata.2016.92

Hanke, M., Baumgartner, F. J., Ibe, P., Kaule, F. R., Pollmann, S., Speck, O., … Stadler, J. A. (2014, May 27). A high-resolution 7-Tesla fMRI dataset from complex natural stimulation with an audio movie. *Sci Data*, *1*. Retrieved from http://doi.org/10.1038/sdata.2014.3

Hanke, M., Dinga, R., Häusler, C., Guntupalli, J. S., Casey, M., Kaule, F. R., & Stadler, J. (2015). High-resolution 7-tesla fmri data on the perception of musical genres–an extension to the studyforrest dataset. *F1000Res*, *4*, 174. Retrieved from https://doi.org/10.12688/f1000research.6679.1

Haxby, J., Guntupalli, J., Connolly, A., Halchenko, Y., Conroy, B., Gobbini, M., … Ramadge, P. (2011). A Common, High-Dimensional Model of the Representational Space in Human Ventral Temporal Cortex. *Neuron*, *72*(2), 404 - 416. Retrieved from http://doi.org/10.1016/j.neuron.2011.08.026

Heller, R., Golland, Y., Malach, R., & Benjamini, Y. (2007). Conjunction group analysis: An alternative to mixed/random effect analysis. *Neuroimage*, *37*(4), 1178 - 1185. Retrieved from https://doi.org/10.1016/j.neuroimage.2007.05.051

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006, April). Syntactic and Semantic Modulation of Neural Activity During Auditory Sentence Comprehension. *J Cogn Neurosci*, *18*(4), 665–679. Retrieved from http://doi.org/10.1162/jocn.2006.18.4.665

Huth, A., de Heer, W., Griffiths, T., Theunissen, F., & Gallant, J. (2016, April). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature*, *532*(7600), 453-8. Retrieved from https://doi.org/10.1038/nature17637

Huth, A., Lee, T., Nishimoto, S., Bilenko, N., Vu, A., & Gallant, J. (2016). Decoding the Semantic Content of Natural Movies from Human Brain Activity. *Front Syst Neurosci*, *10*, 81. Retrieved from https://doi.org/10.3389/fnsys.2016.00081

Huth, A., Nishimoto, S., Vu, A., & Gallant, J. (2012). A continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain. *Neuron*, *76*(6), 1210 - 1224. Retrieved from https://doi.org/10.1016/j.neuron.2012.10.014

Jack, C. R., Barnes, J., Bernstein, M. A., Borowski, B. J., Brewer, J., Clegg, S., … Weiner, M. (2015). Magnetic resonance imaging in Alzheimer's Disease Neuroimaging Initiative 2. *Alzheimers Dement*, *11*(7), 740 - 756. Retrieved from http://doi.org/10.1016/j.jalz.2015.05.002

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. Retrieved from https://doi.org/10.1016/j.neuroimage.2015.11.025

Jovicich, J., Minati, L., Marizzoni, M., Marchitelli, R., Sala-Llonch, R., Bartrés-Faz, D., … Frisoni, G. B. (2016). Longitudinal reproducibility of default-mode network connectivity in healthy elderly participants: A multicentric resting-state fMRI study. *Neuroimage*, *124*, *Part A*, 442 - 454. Retrieved from https://doi.org/10.1016/j.neuroimage.2015.07.010

Kaller, C. P., Rahm, B., Spreer, J., Weiller, C., & Unterrainer, J. M. (2011). Dissociable contributions of left and right dorsolateral prefrontal cortex in planning. *Cereb Cortex*, *21*(2), 307–317. Retrieved from https://doi.org/10.1093/cercor/bhq096

Kirby, K. N., & Maraković, N. N. (1996). Delay-discounting probabilistic rewards: Rates decrease as amounts increase. *Psychon Bull Rev*, *3*(1), 100–104. Retrieved from https://doi.org/10.3758/BF03210748

Kirchhoff, B. A., & Buckner, R. L. (2006, July). Functional-anatomic correlates of individual differences in memory. *Neuron*, *51*(2), 263–274. Retrieved from https://doi.org/10.1016/j.neuron.2006.06.006

Knops, A., Piazza, M., Sengupta, R., Eger, E., & Melcher, D. (2014). A Shared, Flexible Neural Map Architecture Reflects Capacity Limits in Both Visual Short-Term Memory and Enumeration. *J Neurosci*, *34*(30), 9857–9866. Retrieved from https://doi.org/10.1523/JNEUROSCI.2758-13.2014

Kochunov, P., Jahanshad, N., Marcus, D., Winkler, A., Sprooten, E., Nichols, T. E., … van Essen, D. C. (2015). Heritability of fractional anisotropy in human white matter: A comparison of Human Connectome Project and ENIGMA-DTI data. *Neuroimage*, *111*, 300 - 311. Retrieved from http://doi.org/10.1016/j.neuroimage.2015.02.050

Laumann, T., Gordon, E., Adeyemo, B., Snyder, A., Joo, S., Chen, M.-Y., … Petersen, S. (2015). Functional System and Areal Organization of a Highly Sampled Individual Human Brain. *Neuron*, *87*(3), 657 - 670. Retrieved from https://doi.org/10.1016/j.neuron.2015.06.037

Lebreton, M., Abitbol, R., Daunizeau, J., & Pessiglione, M. (2015). Automatic integration of confidence in the brain valuation signal. *Nat Neurosci*, *18*(8), 1159-67. Retrieved from https://doi.org/10.1038/nn.4064

Leroy, F., Cai, Q., Bogart, S. L., Dubois, J., Coulon, O., Monzalvo, K., … Dehaene-Lambertz, G. (2015). New human-specific brain landmark: The depth asymmetry of superior temporal sulcus. *Proc Natl Acad Sci U S A*, *112*(4), 1208–1213. Retrieved from https://doi.org/10.1073/pnas.1412389112

Makris, N., Goldstein, J. M., Kennedy, D., Hodge, S. M., Caviness, V. S., Faraone, S. V., … Seidman, L. J. (2006). Decreased volume of left and total anterior insular lobule in schizophrenia. *Schizophr Res*, *83*(2–3), 155 - 171. Retrieved from http://doi.org/10.1016/j.schres.2005.11.020

Miller, M. B., Donovan, C.-L., Bennett, C. M., Aminoff, E. M., & Mayer, R. E. (2012). Individual differences in cognitive style and strategy predict similarities in the patterns of brain activity between individuals. *Neuroimage*, *59*(1), 83 - 93. Retrieved from https://doi.org/10.1016/j.neuroimage.2011.05.060

Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Ugurbil, K. (2010, May). Multiband multislice GE-EPI at 7 Tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn Reson Med*, *63*(5), 1144-53. Retrieved from https://doi.org/10.1002/mrm.22361

Müller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., … Eickhoff, S. B. (2018). Ten simple rules for neuroimaging meta-analysis. *Neurosci Biobehav Rev*, *84*, 151 - 161. Retrieved from https://doi.org/10.1016/j.neubiorev.2017.11.012

Nickerson, L. D., Smith, S. M., Öngür, D., & Beckmann, C. F. (2017). Using Dual Regression to Investigate Network Shape and Amplitude in Functional Connectivity Analyses. *Front Neurosci*, *11*, 115. Retrieved from https://doi.org/10.3389/fnins.2017.00115

Nieto-Castañón, A., & Fedorenko, E. (2012). Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *Neuroimage*, *63*(3), 1646 - 1669. Retrieved from https://doi.org/10.1016/j.neuroimage.2012.06.065

Oldfield, R. C. (1971, March). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97-113. Retrieved from https://doi.org/10.1016/0028-3932(71)90067-4

Orfanos, D. P., Michel, V., Schwartz, Y., Pinel, P., Moreno, A., Bihan, D. L., & Frouin, V. (2017). The Brainomics/Localizer database. *Neuroimage*, *144*, *Part B*, 309 - 314. Retrieved from https://doi.org/10.1016/j.neuroimage.2015.09.052 (Data Sharing Part {II})

Ossandón, T., Vidal, J. R., Ciumas, C., Jerbi, K., Hamamé, C. M., Dalal, S. S., … Lachaux, J.-P. (2012). Efficient "pop-out" visual search elicits sustained broadband gamma activity in the dorsal attention network. *J Neurosci*, *32*(10), 3414–3421. Retrieved from https://doi.org/10.1523/JNEUROSCI.6048-11.2012

Otto, A. R., Skatova, A., Madlon-Kay, S., & Daw, N. D. (2014). Cognitive control predicts use of model-based reinforcement learning. *J Cogn Neurosci*, *27*(2), 319–333. Retrieved from https://doi.org/10.1162/jocn_a_00709

Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011, February). Cortical representation of the constituent structure of sentences. *Proc Natl Acad Sci U S A*, *108*(6), 2522-7. Retrieved from https://doi.org/10.1073/pnas.1018711108

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011, November). Scikit-Learn: Machine Learning in Python. *J Mach Learn Res*, *12*, 2825–2830. Retrieved from https://dl.acm.org/doi/10.5555/1953048.2078195

Perrone-Bertolotti, M., Kujala, J., Vidal, J. R., Hamame, C. M., Ossandon, T., Bertrand, O., … Lachaux, J.-P. (2012). How silent is silent reading? Intracerebral evidence for top-down activation of temporal voice areas during reading. *J Neurosci*, *32*(49), 17554–17562. Retrieved from https://doi.org/10.1523/JNEUROSCI.2982-12.2012

Pinel, P., d'Arc, B. F., Dehaene, S., Bourgeron, T., Thirion, B., Bihan, D. L., & Poupon, C. (2019). The functional database of the ARCHI project: Potential and perspectives. *Neuroimage*, *197*, 527 - 543. Retrieved from https://doi.org/10.1016/j.neuroimage.2019.04.056

Pinel, P., Thirion, B., Meriaux, S., Jobert, A., Serres, J., Bihan, D. L., … Dehaene, S. (2007). Fast reproducible identification and large-scale databasing of individual functional cognitive networks. *BMC Neurosci*, *8*, 91. Retrieved from https://doi.org/10.1186/1471-2202-8-91

Pinho, A. L., Amadon, A., Ruest, T., Fabre, M., Dohmatob, E., Denghien, I., … Thirion, B. (2018). Individual Brain Charting, a high-resolution fMRI dataset for cognitive mapping. *Sci Data*, *5*, 180105. Retrieved from https://doi.org/10.1038/sdata.2018.105

Poldrack, R., Barch, D., Mitchell, J., Wager, T., Wagner, A., Devlin, J., … Milham, M. (2013). Toward open sharing of task-based fMRI data: the openfMRI project. *Front Neuroinform*, *7*, 12. Retrieved from https://doi.org/10.3389/fninf.2013.00012

Poldrack, R., & Yarkoni, T. (2016, Jan). From Brain Maps to Cognitive Ontologies: Informatics and the Search for Mental Structure. *Annu Rev Psychol*, *67*, 587–612. Retrieved from http://doi.org/10.1146/annurev-psych-122414-033729

Posner, M., Petersen, S., Fox, P., & Raichle, M. (1988). Localization of cognitive operations in the human brain. *Science*, *240*(4859), 1627–1631. Retrieved from http://doi.org/10.1126/science.3289116

Richardson, H., Lisandrelli, G., Riobueno-Naylor, A., & Saxe, R. (2018). Development of the social brain from age three to twelve years. *Nat Commun*, *9*(1027). Retrieved from https://doi.org/10.1038/s41467-018-03399-2

Robinson, E. C., Jbabdi, S., Glasser, M. F., Andersson, J., Burgess, G. C., Harms, M. P., … Jenkinson, M. (2014). MSM: A new flexible framework for Multimodal Surface Matching. *NeuroImage*, *100*, 414 - 426. Retrieved from https://doi.org/10.1016/j.neuroimage.2014.05.069

Sabuncu, M. R., Singer, B. D., Conroy, B., Bryan, R. E., Ramadge, P. J., & Haxby, J. V. (2010). Function-based Intersubject Alignment of Human Cortical Anatomy. *Cereb Cortex*, *20*(1), 130. Retrieved from http://doi.org/10.1093/cercor/bhp085

Saignavongs, M., Ciumas, C., Petton, M., Bouet, R., Boulogne, S., Rheims, S., … Ryvlin, P. (2017). Neural activity elicited by a cognitive task can be detected in single-trials with simultaneous intracerebral EEG-fMRI recordings. *Int J Neural Syst*, *27*(01), 1750001. Retrieved from https://doi.org/10.1142/S0129065717500010

Santoro, R., Moerel, M., De Martino, F., Valente, G., Ugurbil, K., Yacoub, E., & Formisano, E. (2017). Reconstructing the spectrotemporal modulations of real-life sounds from fMRI response patterns. *Proc Natl Acad Sci U S A*, *114*(18), 4799–4804. Retrieved from https://doi.org/10.1073/pnas.1617622114

Schneider, D. W., & Logan, G. D. (2011). Task-switching performance with 1: 1 and 2: 1 cue–task mappings: Not so different after all. *J Exp Psychol Learn Mem Cogn*, *37*(2), 405. Retrieved from https://doi.org/10.1037/a0021967

Schwartz, Y., Thirion, B., & Varoquaux, G. (2013). Mapping cognitive ontologies to and from the brain. In *Nips'13: Proceedings of the 26th international conference on neural information processing systems* (Vol. 2, pp. 1673–1681). Red Hook, NY, USA: Curran Associates Inc. Retrieved from https://arxiv.org/abs/1311.3859

Schwartz, Y., Varoquaux, G., Pallier, C., Pinel, P., Poline, J.-B., & Thirion, B. (2012). Improving Accuracy and Power with Transfer Learning Using a Meta-analytic Database. In N. Ayache, H. Delingette, P. Golland, & K. Mori (Eds.), *Med Image Comput Comput Assist Interv. 2012* (Vol. 15, pp. 248–255). Berlin, Heidelberg: Springer. Retrieved from https://doi.org/10.1007/978-3-642-33454-2_31

Sengupta, A., Kaule, F. R., Guntupalli, J. S., Hoffmann, M. B., Häusler, C., Stadler, J., & Hanke, M. (2016, October). A studyforrest extension, retinotopic mapping and localization of higher visual areas. *Sci Data*, *3*. Retrieved from https://doi.org/10.1038/sdata.2016.93

Shallice, T. (1982). Specific impairments of planning. *Philos Trans R Soc Lond B Biol Sci*, *298*(1089), 199–209. Retrieved from https://doi.org/10.1098/rstb.1982.0082

Siegelman, M., Blank, I. A., Mineroff, Z., & Fedorenko, E. (2019). An Attempt to Conceptually Replicate the Dissociation between Syntax and Semantics during Sentence Comprehension. *Neuroscience*, *413*, 219 - 229. Retrieved from https://doi.org/10.1016/j.neuroscience.2019.06.003

Smith, S., Beckmann, C., Andersson, J., Auerbach, E., Bijsterbosch, J., Douaud, G., … Glasser, M. (2013). Resting-state fMRI in the Human Connectome Project. *Neuroimage*, *80*, 144 - 168. Retrieved from https://doi.org/10.1016/j.neuroimage.2013.05.039 (Mapping the Connectome)

Smith, S., Jenkinson, M., Woolrich, M., Beckmann, C., Behrens, T. E., Johansen-Berg, H., … Matthews, P. M. (2004). Advances in functional and structural {MR} image analysis and implementation as {FSL}. *Neuroimage*, *23*, *Supplement 1*, S208 - S219. Retrieved from http://doi.org/10.1016/j.neuroimage.2004.07.051

Smith, S., Nichols, T., Vidaurre, D., Winkler, A., Behrens, T., Glasser, M., … Miller, K. (2015, November). A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat Neurosci*, *18*(11), 1565–1567. Retrieved from https://doi.org/10.1038/nn.4125

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., … Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc Natl Acad Sci U S A*, *106 31*, 13040-5. Retrieved from https://doi.org/10.1073/pnas.0905267106

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J Exp Psychol*, *18*(6), 643. Retrieved from https://doi.org/10.1037/h0054651

Tavor, I., Jones, O. P., Mars, R. B., Smith, S. M., Behrens, T. E., & Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science*, *352*(6282), 216–220. Retrieved from https://doi.org/10.1126/science.aad8127

Taylor, J. R., Williams, N., Cusack, R., Auer, T., Shafto, M. A., Dixon, M., … Henson, R. N. (2017). The Cambridge Centre for Ageing and Neuroscience (cam-CAN) data repository: Structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample. *Neuroimage*, *144*, *Part B*, 262 - 269. Retrieved from https://doi.org/10.1016/j.neuroimage.2015.09.018 (Data Sharing Part {II})

Thirion, B., Pinel, P., Mériaux, S., Roche, A., Dehaene, S., & Poline, J.-B. (2007). Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. *Neuroimage*, *35*(1), 105 - 120. Retrieved from https://doi.org/10.1016/j.neuroimage.2006.11.054

van Essen, D. C., Glasser, M. F., Dierker, D. L., Harwell, J., & Coalson, T. (2012). Parcellations and Hemispheric Asymmetries of Human Cerebral Cortex Analyzed on Surface-Based Atlases. *Cereb Cortex*, *22*(10), 2241. Retrieved from http://doi.org/10.1093/cercor/bhr291

van Essen, D. C., Smith, J., Glasser, M. F., Elam, J., Donahue, C. J., Dierker, D. L., … Harwell, J. (2017). The Brain Analysis Library of Spatial maps and Atlases (BALSA) database. *Neuroimage*, *144*, *Part B*, 270 - 274. Retrieved from https://doi.org/10.1016/j.neuroimage.2016.04.002 (Data Sharing Part {II})

Varoquaux, G., Schwartz, Y., Pinel, P., & Thirion, B. (2013). Cohort-Level Brain Mapping: Learning Cognitive Atoms to Single Out Specialized Regions. In J. C. Gee, S. Joshi, K. M. Pohl, W. M. Wells, & L. Zöllei (Eds.), *Inf Process Med Imaging* (Vol. 23, pp. 438–449). Berlin, Heidelberg: Springer. Retrieved from https://doi.org/10.1007/978-3-642-38868-2_37

Varoquaux, G., Schwartz, Y., Poldrack, R., Gauthier, B., Bzdok, D., Poline, J.-B., & Thirion, B. (2018). Atlases of cognition with large-scale brain mapping. *PLoS Comput Biol*, *14*(11). Retrieved from https://doi.org/10.1371/journal.pcbi.1006565

Vidal, J. R., Ossandón, T., Jerbi, K., Dalal, S. S., Minotti, L., Ryvlin, P., … Lachaux, J.-P. (2010). Category-specific visual responses: an intracranial study comparing gamma, beta, alpha, and ERP response selectivity. *Front Hum Neurosci*, *4*, 195. Retrieved from https://doi.org/10.3389/fnhum.2010.00195

Wager, T. D., Atlas, L. Y., Lindquist, M. A., Roy, M., Woo, C.-W., & Kross, E. (2013). An fMRI-Based Neurologic Signature of Physical Pain. *N Engl J Med*, *368*(15), 1388-1397. Retrieved from https://doi.org/10.1056/NEJMoa1204471

Wager, T. D., Lindquist, M., & Kaplan, L. (2007, 06). Meta-analysis of functional neuroimaging data: current and future directions. *Soc Cogn Affect Neurosci*, *2*(2), 150-158. Retrieved from https://doi.org/10.1093/scan/nsm015

Ward, G., & Allport, A. (1997). Planning and problem solving using the five disc Tower of London task. *Q J Exp Psychol Section A*, *50*(1), 49–78. Retrieved from https://doi.org/10.1080/713755681

Yeo, B. T. T., Sabuncu, M. R., Vercauteren, T., Ayache, N., Fischl, B., & Golland, P. (2010, Mar). Spherical demons: fast diffeomorphic landmark-free surface registration. *IEEE transactions on medical imaging*, *29*(3), 650-668. Retrieved from https://doi.org/10.1109/TMI.2009.2030797

A | DEMOGRAPHIC DATA OF THE PARTICIPANTS

Table A.1 summarizes the demographic profile of the extended version of the IBC dataset, which accounts for thirteen participants in total.

Subject ID	Age	Sex	Handedness score
sub-01	39.5	M	0.3
sub-02	32.8	M	1
sub-04	26.9	M	0.8
sub-05	27.4	M	0.6
sub-06	33.1	M	0.7
sub-07	38.8	M	1
sub-08	36.5	F	1
sub-09	38.5	F	1
sub-11	35.8	M	1
sub-12	40.8	M	1
sub-13	28.2	M	0.6
sub-14	28.3	M	0.7
sub-15	30.3	M	0.9

TABLE A.1 Demographic data of the participants. Age stands for the participants' age upon recruitment.

For more information about the profile of the participants and procedures undertaken during recruitment, consult Pinho et al. (2018).

B | DESCRIPTION OF THE BEHAVIORAL CONDITIONS INCLUDED IN THE TASKS

This section provides a summarized description of the tasks featuring the first release of the IBC dataset. For a full description of the tasks, consult Pinho et al. (2018)

B.1 | ARCHI Tasks

The ARCHI tasks consist of a serie of localizers targeting several cognitive processes of interest.

B.1.1 | ARCHI Standard

This multi-functional localizer—dedicated to fast brain-mapping—isolates a variety of elementary mental functions, ranging from low-level to high-level cognition. They are concerned with visual perception, motor actions, reading, language comprehension and mental calculation. The task is organized as a fast event-related paradigm, composed of trials including ten different experimental conditions. These conditions are listed and described in Table B.1.

Condition	Description
Right hand	Right-hand three-times button press[§]
Left hand	Left-hand three-times button press[§]
Listen to sentence	Listen to narrative sentences
Read sentence	Read narrative sentences
Horizontal checkerboard	Viewing of flashing horizontal checkerboard
Vertical checkerboard	Viewing of flashing vertical checkerboard
Mental subtraction	Subtraction[§]

TABLE B.1 Description of conditions from ARCHI Standard task.

[§] Conditions indicated by either visual or auditory instruction

B.1.2 | ARCHI Spatial

This task isolates neurocognitive mechanisms concerned with visual orientation and body representation in space. Its paradigm is organized in terms of five conditions; events per condition are combined and displayed in blocks. Conditions are listed and described in Table B.2.

Condition	Description
Saccades	Saccade, in which ocular movements were performed according to the displacement of a fixation cross from the center toward peripheral locations in the image displayed
Object grasping	Mimicry of object grasping with right hand, in which the corresponding object was displayed on the screen[§]
Mimic orientation	Mimic orientation of rhombus, displayed as image background on the screen[§]
Guess which hand	Mental judgment on the left-right orientation of a hand displayed as visual stimulus
Hand palm or back	Mental judgment on the palmar-dorsal direction of a hand displayed as visual stimulus

TABLE B.2 Description of conditions from ARCHI Spatial task.

[§] Conditions featured by the same visual stimuli, in order to capture grasping-specific activity.

B.1.3 | ARCHI Social

This task examines mental functions about social representations and social identity, such as *animacy perception*, *narrative comprehension*, *mentalization* and *theory-of-mind*. Its paradigm is organized in terms of eight conditions; events from subsets of different conditions are combined and displayed in blocks. Conditions are listed and described in Table B.3.

Condition	Description
Mental motion	Watch short movies of triangles, exhibiting a putative social interaction
Random motion	Watch short movies of triangles, displaying random movements
False-belief story	Interpret silently short stories[§], presented in text on the screen, featuring a *false-belief* plot
Mechanistic story	Interpret silently short stories[§], presented in text on the screen, featuring a cause-consequence mechanistic plot
False-belief tale	Interpret silently short stories[§], presented in an audio clip, featuring a *false-belief* plot
Mechanistic tale	Interpret silently short stories[§], presented in an audio clip, featuring a cause-consequence mechanistic plot
Voice sound	Listen passively to short samples of human voices
Natural sound	Listen passively to short samples of natural sounds

TABLE B.3 Description of conditions from ARCHI Social task.

[§] No active response was involved.

B.1.4 | ARCHI Emotional

This task isolates mental functions involved in the perception of emotions, from facial to eye expressions, required for judging temper or subjective trustworthiness. Its paradigm is organized in terms of five conditions; events from subsets of different conditions are combined and displayed in blocks. Conditions are listed and described in Table B.4.

Condition	Description
Face gender	Assess gender through visualization of pictures of human faces
Face trusty	Assess trustfulness through visualization of pictures of human faces
Expression gender	Assess gender through visualization of pictures with eyes' expression
Expression intention	Assess temper through visualization of pictures with eyes' expression
Scrambled image	Mental assessment on the slope of a gray-scale grid image (obtained from scrambling an eyes' image) that may be tilted or not

TABLE B.4 Description of conditions from ARCHI Emotional task. The baseline condition for this task consisted on the display of a gray-scale grid image that should be judged as either being tilted or not.

B.2 | HCP Tasks

These set of tasks were adapted from the task-fMRI paradigms used in HCP, yet with some timing differences and more trials for a subset of them (consult Pinho et al. (2018) for further details).

B.2.1 | HCP Emotion

The main purpose of this task is to isolate neurocognitive mechanisms involved in the perception of fear and anger, from the visual comparison between generic geometric shapes and affective facial expressions. The paradigm is organized in terms of two conditions; events of these conditions are combined and displayed in blocks. They are listed and described in Table B.5.

Condition	Description
Face image	Match one of the two faces, displayed on the bottom of the screen, with a face, displayed on the top of the screen, according to emotional expression
Shape outline	Match one of the two shapes, displayed on the bottom of the screen, with a shape displayed on the top of the screen

TABLE B.5 Description of conditions from HCP Emotion task.

B.2.2 | HCP Gambling

The aim of this task is to capture effects concerned with incentive as well as aversive salience. The paradigm is organized in eight blocks, each of them composed of eight events. Every event consists of playing a gambling game, whose goal is to guess whether the next number to be displayed—ranging from one to nine—will be smaller or larger than five. Feedback is given afterwards and conditions are defined according to the *reward* or *loss* experienced after a given outcome (see Table B.6).

Condition	Description
Reward	Gambling task with reward as outcome
Punishment	Gambling task with loss as outcome

TABLE B.6 Description of conditions from HCP Gambling task.

B.2.3 | HCP Motor

The aim of this task is to isolate contralateral and ipsilateral motor representations of the foot, hand and togue in the motor cortex and cerebellum, respectively. The paradigm is organized in terms of five conditions concerning left foot, right foot, left hand, right hand and tongue; events of sets of different conditions are combined and displayed in blocks. Conditions are listed and described in more detail in Table B.7.

Condition	Description
Tongue	Movement of tongue in sync with the flashes of the visual stimulus
Right hand	Movement of right hand in sync with the flashes of the visual stimulus
Left hand	Movement of left hand in sync with the flashes of the visual stimulus
Right foot	Movement of right foot in sync with the flashes of the visual stimulus
Left foot	Movement of left foot in sync with the flashes of the visual stimulus

TABLE B.7 Description of conditions from HCP Motor task.

B.2.4 | HCP Language

This task aims at localizing effects involved in narrative comprehension and auditory sentence recognition. The paradigm comprises two conditions, wherein each of them is presented in a full block, one at the time. Conditions are listed and described in Table B.8.

Condition	Description
Tale	Listen and interprete auditory narratives
Mental addition	Listen and solve arithmetic problems

TABLE B.8 Description of conditions from HCP Language task.

B.2.5 | HCP Relational

This task is intended to isolate neurocognitive mechanisms pertaining to relational and feature comparison as well as visual form recognition. It employs a relational match-to-sample paradigm, featuring a second-order comparison of relations between two pairs of objects. The paradigm comprises two conditions, wherein each of them is presented in a full block, one at the time. Conditions are listed and described in Table B.9.

Condition	Description
Relational processing	Evaluate whether a pair of objects, displayed on the bottom of the screen, change along the same dimension (i.e. either shape or texture) as the pair displayed on the top of the screen
Visual matching	Evaluate whether the object, displayed on the bottom of the screen, match one of the two objects displayed on the top of the screen, according to the dimension specified (i.e. shape or texture)

TABLE B.9 Description of conditions from HCP Relational task.

B.2.6 | HCP Social

This task aims at providing evidence for task-specific brain activation implicated in social cognition. Similarly to the ARCHI Social task (see Appendix section B.1.3), this task is intended to tackle functional mechanisms related to social cognition, namely *animacy perception*, *narrative comprehension*, *mentalization* and *theory-of-mind*. The paradigm comprises two conditions, wherein each of them is presented in a full block, one at the time. Conditions are listed and described in Table B.10.

Condition	Description
Mental motion	Watch the movement of geometric shapes depicting putative social interactions
Random motion	Watch the random movement of geometric shapes

TABLE B.10 Description of conditions from HCP Social task.

B.2.7 | HCP Working Memory

This task was adapted from the classical n-back task to serve as a functional localizer of not only working-memory mechanisms but also those involved in visual-form discrimination. The paradigm comprises two conditions, wherein each of them is presented in a full block, one at the time. Conditions are listed and described in Table B.11.

Condition	Description
0-Back face image	Recall whether the picture under display match the one displayed as a target, i.e. at the beginning of the consecutive presentation of a set of pictures depicting faces
2-Back face image	Recall whether the picture under display match the one displayed two events before, during the consecutive presentation of a set of pictures depicting faces
0-Back place image	Recall whether the picture under display match the one displayed as a target, i.e. at the beginning of the consecutive presentation of a set of pictures depicting places
2-Back place image	Recall whether the picture under display match the one displayed two events before, during the consecutive presentation of a set of pictures depicting places
0-Back tool image	Recall whether the picture under display match the one displayed as a target, i.e. at the beginning of the consecutive presentation of a set of pictures depicting tools
2-Back tool image	Recall whether the picture under display match the one displayed two events before, during the consecutive presentation of a set of pictures depicting tools
0-Back body image	Recall whether the picture under display match the one displayed as a target, i.e. at the beginning of the consecutive presentation of a set of pictures depicting different parts of the human body
2-Back body image	Recall whether the picture under display match the one displayed two events before, during the consecutive presentation of a set of pictures depicting different parts of the human body

TABLE B.11 Description of conditions from HCP Working-Memory task.

B.2.8 | RSVP Language

To localize the areas implicated in language comprehension, participants were presented with stimuli consisting of sequences of words, pseudowords or nonwords. For some conditions, these sequences were composed by well-formed sentences. By the end of each trial, a probe condition was presented with a word. The participant had to indicate afterwards whether that word had been presented or not in the beginning of the trial. Experimental conditions are described in Table B.12.

Condition	Description
Complex sentence	Presentation of complex meaningful sentences, containing at least two clauses
Simple sentence	Presentation of simple meaningful sentences, with only one main clause
Read *jabberwocky*	Presentation of syntactically well-formed sentences whose content consist of meaningless, yet pronounceable, words
Read words	Presentation of list of words
Read pseudowords	Presentation of list of pseudowords
Consonant strings	Presentation of list of nonwords (i.e. consonant strings)

TABLE B.12 Description of conditions from RSVP Language task.

C | INTERSUBJECT VARIABILITY ACROSS CONTRASTS FROM THE IBC, ARCHI AND HCP DATASETS

Intersubject variability of the ARCHI and HCP contrast maps was tested between different datasets. Figures C.1 and C.2 display the average of individual *z*-maps' correlations of a given contrast between every subject pair, within datasets. They summarize how similar individual maps are per contrasts for each dataset, thus highlighting different levels of consistency of the functional signatures across individuals for these different datasets. Overall, both figures show that better levels of consistency were obtained for the IBC contrasts than for the ARCHI and HCP counterparts.

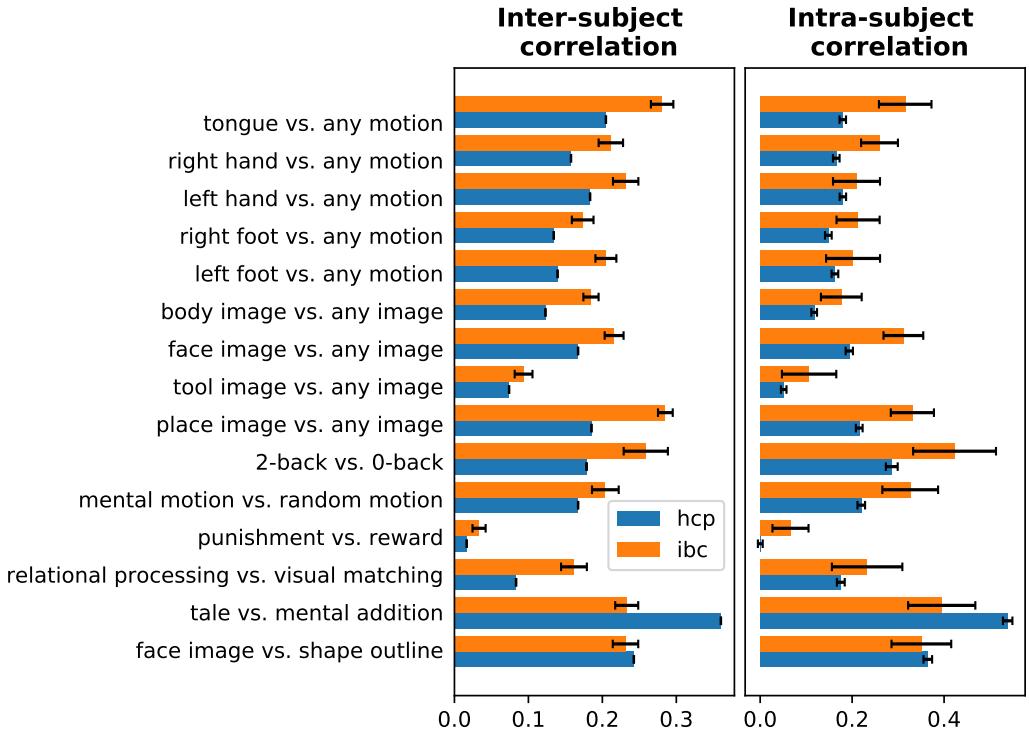

FIGURE C.1 Comparison of intersubject and intrasubject variability across contrasts between IBC and HCP datasets. The bar chart represents the means of the distributions referring to the correlations of *z*-maps per contrast. *(Left panel)* Intersubject correlations were estimated from pairs of contrasts for all subjects pairs, whereas *(Right panel)* intrasubject correlations were estimated from all possible pairs of contrasts from "PA" and "AP" runs, within sessions. Bars in orange correspond to the correlations of contrasts obtained from the IBC dataset; bars in blue correspond to the correlations of contrasts obtained from the HCP dataset. Error bars represent the 95% CI.

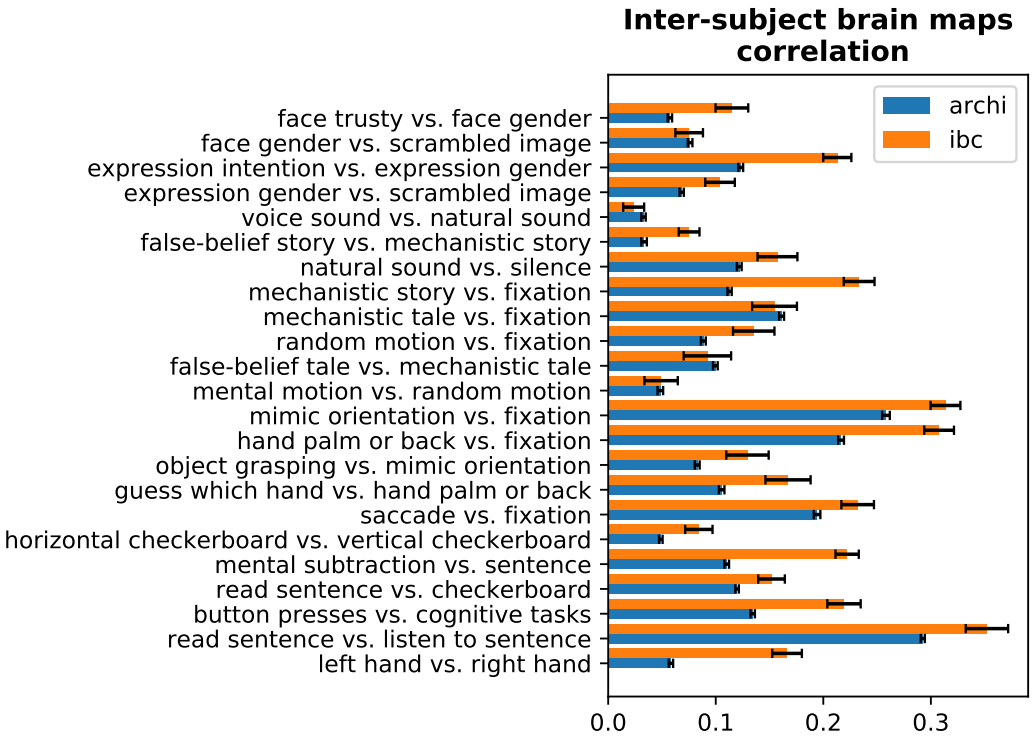

FIGURE C.2 Comparison of intersubject variability across contrasts between IBC and ARCHI datasets. The bar chart represents the means of the distributions referring to the correlations of *z*-maps per contrast, for all subjects pairs. Bars in orange correspond to the correlations of contrasts obtained from the IBC dataset; bars in blue correspond to the correlations of contrasts obtained from the ARCHI dataset. Error bars represent the 95% CI.

D | SUPPORTING MATERIAL ON THE REGION-EXTRACTION ANALYSIS OF THE LANGUAGE NETWORK

This Appendix section contains supporting material for Section 3.5.

D.1 | Subject-Specific Regions-of-Interest

Figure D.1 displays a panel of 13 glass brains; they feature the individual projection, after dual regression (see Section 2.7.8 for more details about this method), of six regions-of-interest from the left hemisphere for every participant of the IBC-dataset first-release cohort.

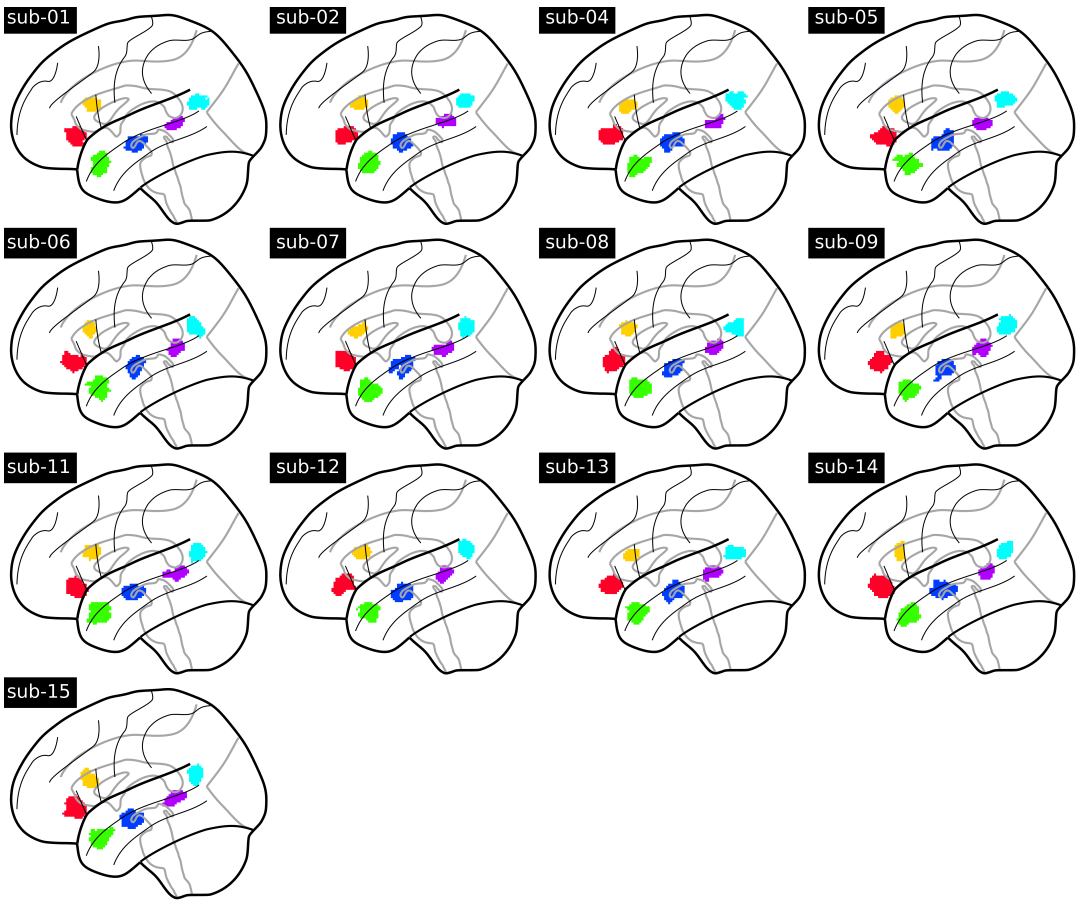

FIGURE D.1 Subject-specific regions-of-interest on evidence in a glass brain. This panel provides glass brain maps, on the left sagittal plane, of six regions-of-interest of the language network, as described in Pallier et al. (2011), projected onto individual templates obtained after dual regression, for every participant featuring the IBC dataset first-release. Each region-of-interest is marked in a different color on the glass brain: *(red)* left Inferior Frontal Gyrus (IFG) *pars orbitalis*, *(yellow)* left IFG *pars triangularis*, *(cyan)* left Temporoparietal Junction, *(green)* left Temporal Pole, *(dark blue)* left anterior Superior Temporal Sulcus (STS), and *(purple)* left posterior STS.

D.2 | Voxel Classification of Regions-of-Interest from *Pallier*'s Study using all IBC Contrasts

Figure D.2 shows prediction scores obtained between pairs of regions against the corresponding chance levels, using all main contrasts in the first release of the IBC-dataset. Scores were slightly higher than those obtained using a selection of contrasts related to semantic processing. Such results thus reveal that variability in the functional profiles of these regions-of-interest are mostly but not totally explained by cognitive processes pertaining to language.

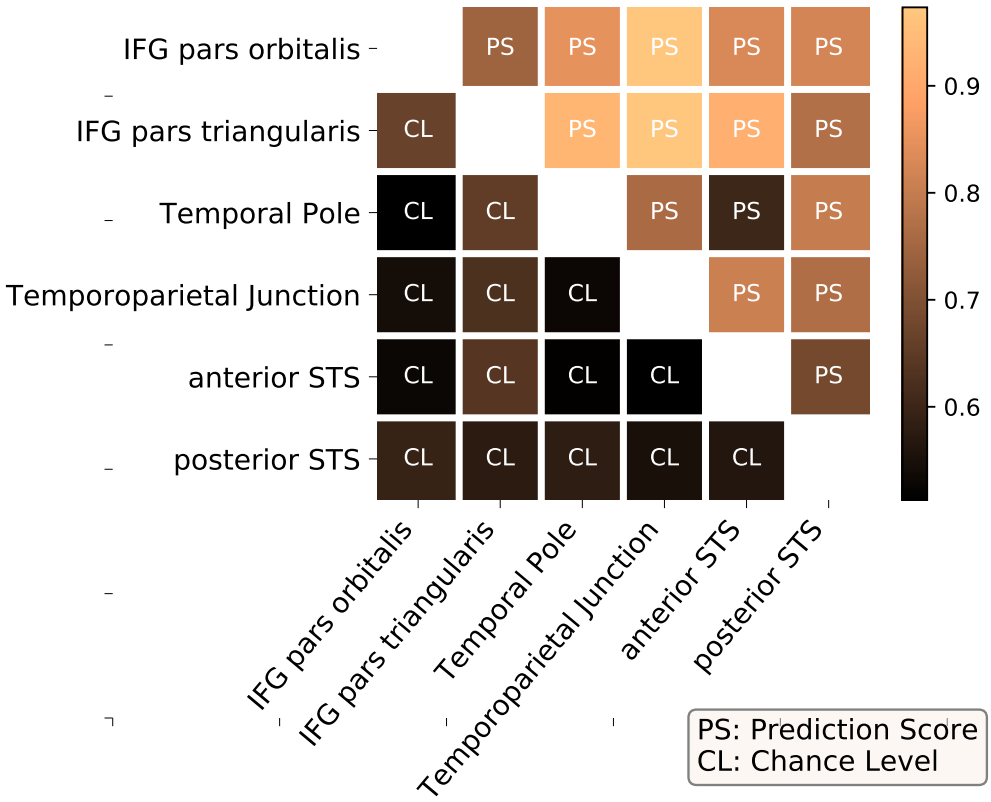

FIGURE D.2 Prediction scores of voxel classification as belonging to one out of two regions-of-interest against chance level. The scores were estimated based on the functional activation of voxels from different pairs of regions-of-interest pertaining to the language network. Combinations of two brain regions from a total amount of six regions—Inferior Frontal Gyrus (IFG) *pars orbitalis*, IFG *pars triangularis*, Temporal Pole, Temporoparietal Junction, anterior Superior Temporal Sulcus (STS) and posterior STS—were employed in a cross-validation experiment to calculate each prediction score. The functional profiles of the voxels were extracted from all main contrasts of all tasks of the IBC-dataset first release.

D.3 | Region-of-Interest extraction from the HCP dataset

To extract regions-of-interest of the language network from the HCP dataset, we firstly selected the contrast *story* versus *math* from the HCP Language task (for more information about this task, consult Section 2.4.2) as the one that better isolates specific-related mechanisms referring to language, like narrative comprehension and sentence recognition. The volume-based unthresholded *z*-maps for the contrasts *story* and *math* for $n = 786$ subjects were obtained from the NeuroVault collection #4337 (https://identifiers.org/neurovault.collection:4337). The average of their joint effects across subjects and its statistical significance were estimated by fitting an RFX linear model in the unthresholded *z*-maps and computing afterwards a Paired Sample t-Test between both contrasts; this accounts for the subject effects between pairs of contrasts, since each pair belongs to a different subject. A *z*-transform of the resulting group-level *t*-map was performed afterwards for standardization of the results at every voxel. Following a careful visual inspection of the activation map using *FSLeyes* version: 0.33.1 (https://zenodo.org/record/3858136#.XvvbfJaxU5k) together with the probabilistic *Harvard-Oxford cortical and subcortical structural atlases* (Makris et al., 2006; Frazier et al., 2005; Desikan et al., 2006; Goldstein et al., 2007), this one was thresholded at z=16, as at this level we were able to discriminate and identify functional regions. Binary masks were then created out of the regions' delineations obtained from this thresholded *z*-map using Nilearn.

To draw the unique cognitive profile of the selected regions based on the IBC contrasts and, subsequently, assess the statistical significance of their differences, we employed the same procedures as described in Section 2.7.8. Figure D.3 thus displays the functional profiles of the four regions-of-interest identified from the contrast *story* versus *math* of the HCP Language task. They are: *(a)* the Ventromedial Prefrontal Cortex (Ventromedial PFC), *(b)* the left Inferior Frontal Gyrus (IFG) *pars orbitalis* plus *pars triangularis*, *(c)* the left anterior Superior Temporal Sulcus (STS) plus left Temporal Pole, and *(d)* the left Temporoparietal Junction. Results are consistent to those obtained in Figure 6 for the overlapping regions. Interestingly, stronger effects were obtained for the contrasts involved in combinatorial semantics—i.e. *read sentence* versus *read words* and *read sentence* versus *read jabberwocky* contrasts—in this new set of regions and they are particularly notorious when one compares the functional profiles obtained for the Temporoparietal Junction in both cases. These results reveal the expected specificity of the delineated regions-of-interest from the *story* versus *math* toward higher-order semantic modules, as effects-of-interest concerning narrative comprehension and semantic integration are present in this contrast. Nevertheless, the HCP regions-of-interest lack specificity with respect to lower-order semantic processes, since (for instance) it was not possible to identify the left posterior Superior Temporal Sulcus as in *Pallier*'s set.

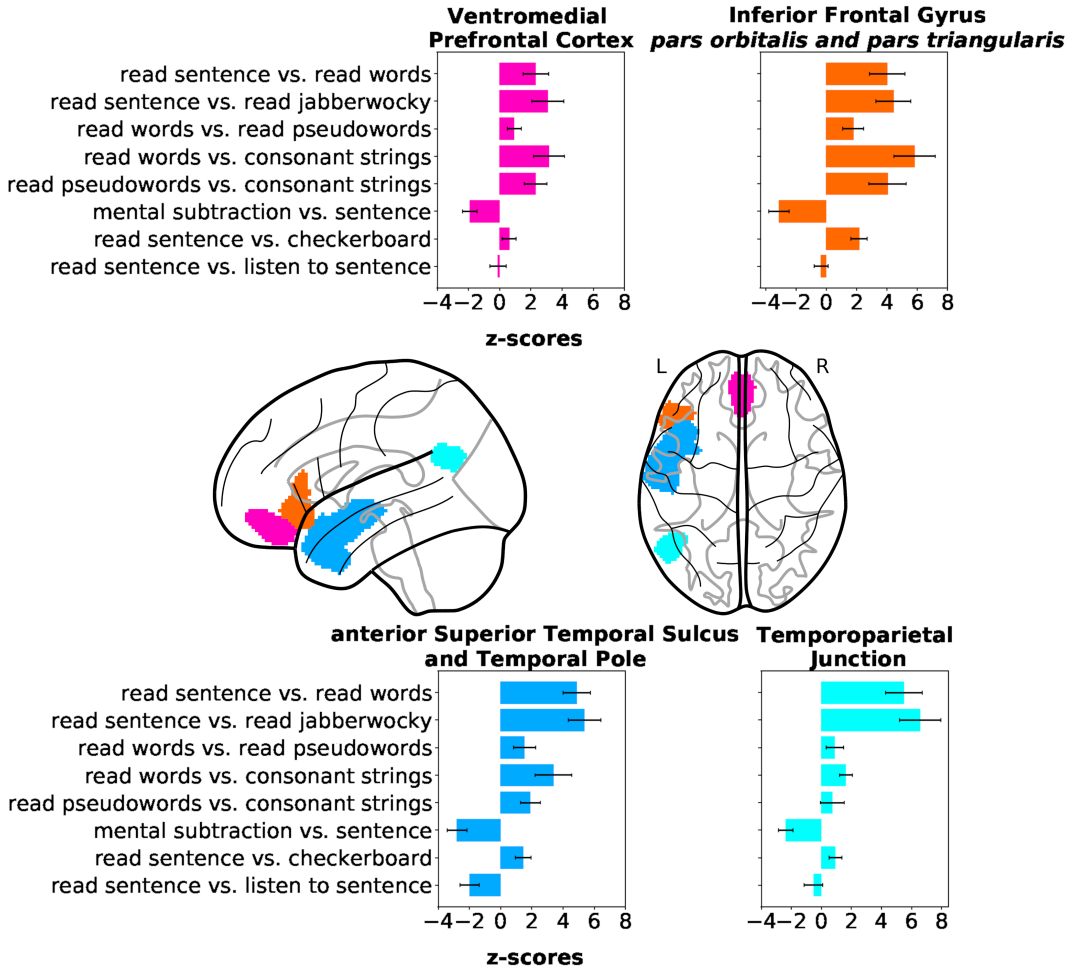

FIGURE D.3 Comparison of the cognitive profiles for a set of regions-of-interest delineated from the contrast *story* versus *math* of the HCP Language task. To verify that results are not strictly tied to particular choices concerning the selection of regions-of-interest, the same analysis—as described in Section 2.7.8—was conducted in a different set of regions-of-interest extracted from the group-level contrast *story* versus *math* of the HCP Language task, which was estimated from data referring to 786 participants of the HCP900 dataset. Not only both sets overlap to a considerable extent but also they share consistent results between them (see Figure 6 for a direct comparison with the present one). Each bar plot represents, for a set of contrast *z*-maps, the means of the distributions across subjects of the average of *z*-scores in a set of voxels inside a regions-of-interest. Error bars represent the 95% CI. Bar colors identify four regions-of-interest that are in evidence on the glass brain: *(magenta)* Ventromedial Prefrontal Cortex, *(orange)* left Inferior Frontal Gyrus, *(light blue)* left anterior Superior Temporal Sulcus plus left Temporal Pole, and *(cyan)* left Temporoparietal junction.

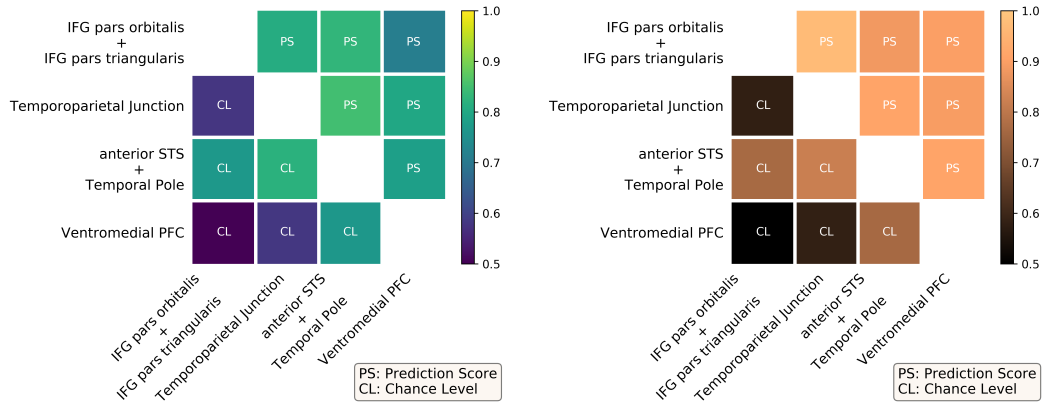

FIGURE D.4 Prediction scores of voxel classification as belonging to one out of two regions-of-interest against chance level for the regions-of-interest delineated from the contrast *story* versus *math* of the HCP Language task. The scores were estimated based on the functional activation of voxels from different pairs of regions-of-interest pertaining to the language network. Combinations of two brain regions from a total amount of four regions—Ventromedial Prefrontal Cortex (Ventromedial PFC), left Inferior Frontal Gyrus (IFG), left anterior Superior Temporal Sulcus (STS) plus left Temporal Pole and left Temporoparietal Junction—were employed in a cross-validation experiment to calculate each prediction score. The functional profiles of the voxels were extracted from: *(left)* the contrasts of the IBC dataset that relate to semantic processing in language (see Figure D.3); and *(right)* all main contrasts of all tasks of the IBC-dataset first release.

Notes

1 According to *Brain Imaging Data Structure* (BIDS) Specification, *task* is definined as "a set of structured activities performed by the participant that are usually accompanied by stimuli and responses, and can greatly vary in complexity. (...) In the context of brain scanning, a task is always tied to one data acquisition (*aka run*). Therefore, even if during one run the subject performed multiple conceptually different behaviors (with different sets of instructions) they will be considered one (combined) task." (BIDS v1.2.1.: https://bids-specification.readthedocs.io/en/stable/)

2 Main contrast stands for a contrast that isolates an effects-of-interest in a given task. Importantly, main contrasts within tasks are linearly independent and, consequently, this also stands true across tasks.

3 The cohort has one more individual than in Pinho et al. (2018).

4 The baseline of the predictive score is actually negative, since this prediction reflects noise from the training and the test set.

5 It combines both conditions *read sentence* and *listening to sentence* of the ARCHI Standard task.

6 https://www.doi.org/10.18112/openneuro.ds002685.v1.0.0

7 https://identifiers.org/neurovault.collection:6618