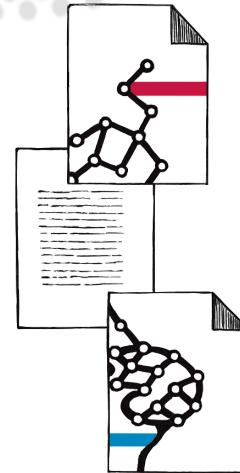




Data Release Documentation



Version	Author(s)
1.1	Taherian, F., Das, D., Kahraman, K., Sripad, P., Bonacchi, N., Brown, T., Oh, H., Melloni, L.
Date	Editor(s)
03.04.2024	Taherian, F., Das, D., Kahraman, K., Sripad, P., Bonacchi, N., Brown, T., Melloni, L.

TABLE OF CONTENTS

Updates.....	5
Future Releases.....	5
Experiment 1.....	5
Previous Releases.....	5
MEEG-DR-doc_2024-03-18_v1.0.....	5
Documentation Changes.....	5
Documentation-Changes_2024-04-17_v1.0.....	5
Introduction.....	6
Overview of COGITATE.....	6
Goals.....	7
Experiments.....	7
Experiment 1: Conscious Perception.....	8
Experiment 2: Video Game Engagement.....	10
COGITATE Dataset.....	10
Demography of Subjects.....	13
Quality Check.....	14
Exclusion Criteria.....	14
Description of COGITATE Data.....	15
Data Acquisition.....	18
Stimuli.....	18
Procedure.....	18
iEEG Related Modifications of the Design.....	20
M-EEG Data Acquisition.....	20
Hardware.....	20
Location of Electrodes and ECG/EOG Measurements.....	20
Head Position Indicator (HPI) Coils.....	21
Anatomical MRI Data Acquisition.....	22
Behavioral Data Acquisition.....	22
Eye Tracking Data Acquisition.....	22
Behavioral Data Code Scheme.....	22
Eye Tracker and MEG Code Scheme.....	23
Empty Room Recording.....	24
Resting-State (rM-EEG).....	25
Task (tM-EEG).....	25
Full Structure of Session.....	25
Inclusion Criteria.....	25
Quality Check and Exclusion Criteria.....	26
iEEG Data Acquisition.....	26

Hardware.....	26
Anatomical MRI Data Acquisition.....	27
CT Data Acquisition.....	27
Behavioral Data Acquisition.....	28
Eye Tracking Data Acquisition.....	28
Behavioral Data Code Scheme.....	29
Eye Tracker Data Code Scheme.....	29
iEEG Code Scheme.....	29
Photodiode Trigger Scheme.....	29
Log File Alignment.....	30
Surface Reconstruction and Electrode Localization.....	36
Finger Localizer Task.....	36
Task (tiEEG).....	37
Full Structure of Session.....	37
Inclusion Criteria.....	37
Quality Check.....	37
Exclusion Criteria.....	38
Deviations and Missing Data.....	38
Curation Procedures.....	39
Data Curation Procedure.....	39
Deviations from Data Curation Procedure.....	39
iEEG.....	39
Metadata Curation Procedure.....	40
Access to COG TATE Data.....	40
1. XNAT.....	41
Naming Convention on XNAT.....	48
2. Bundles.....	48
Naming Convention for Bundles.....	50
Experiment 1: Directory Structure of Data Bundles.....	51
Raw Data.....	51
Raw M-EEG Data Directory Structure.....	52
Raw iEEG Data Directory Structure.....	52
BIDS Format.....	52
BIDS M-EEG Data Directory Structure.....	53
BIDS iEEG Data Directory Structure.....	55
Support and Report Bugs.....	56
Links and Reference Materials.....	59
Appendices.....	60
Appendix 1. Screening Questionnaire.....	60
Appendix 2. Case Report Form.....	60
Appendix 3. Exit Questionnaire.....	60



Appendix 4. MEG Standard Operating Procedure.....	60
Appendix 5. GDPR Requirements.....	60
Appendix 6. Terms of Use.....	60
Appendix 7. Data Curation Standard Operating Procedure.....	60
Appendix 8. Metadata Curation Standard Operating Procedure.....	60
Appendix 9. iEEG Standard Operating Procedure.....	60
Acknowledgements.....	60
Glossary.....	64
References.....	65

Updates

This is the **second version (V1.1)** of the COGITATE data release document. New updates or any changes to the previous versions will be announced here, and on [COGITATE_wiki](#) as well.

In **V1.1**, the raw/unprocessed and [BIDS](#) (Brain Imaging Data Structure) format of iEEG (intracranial electroencephalography) data, collected by the Cogitate Consortium for [Experiment 1](#), are released.

Future Releases

Here are the items that will be released soon:

Experiment 1

- BIDS format of the M-EEG data (batch 2)
- Unprocessed/raw format of all M-EEG data (batch 1 and batch 2)
- Unprocessed/raw and BIDS format of fMRI data

Previous Releases

Here is the list of previous COGITATE data release documents:

Version	Data Release Document	Description
v1.0	MEEG-DR-doc_2024-03-18_v1.0	MEEG (batch 1) data release document

Documentation Changes

In each new version of the COGITATE data release document, relevant content for another modality/experiment is appended to the previous version, and there may be modifications made to the content of the previous versions. The list of these changes is compiled in the following documents:

Version	Data Release Documentation Changes	Description
v1.0	Documentation-Changes_2024-04-17_v1.0	Changes from MEEG (batch 1) data release document (v1.0) to iEEG data release document (v1.1)



Attention: M-EEG, MEEG, M/EEG, MEG/EEG or MEG might be used interchangeably throughout this document or the name of data folders, but all of them pertain to a singular data. This also applies to iEEG and ECoG (Electrocorticography).

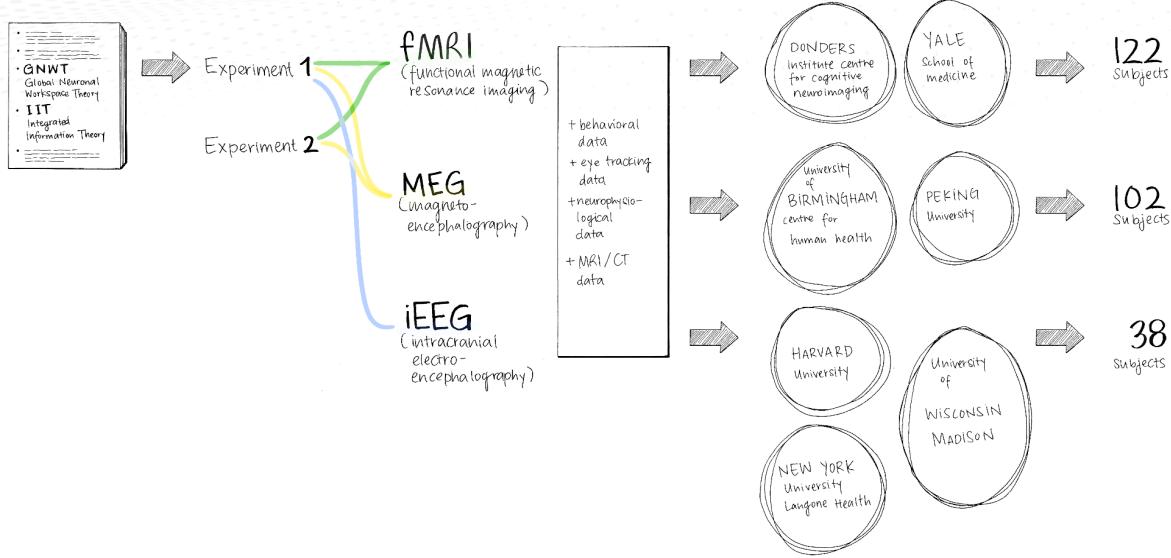
Introduction

This document provides guidance and detailed information on the datasets released by COG TATE, how to access them, the directory structure, and a description on various types of data acquired for each modality.

Overview of COG TATE

What are the mechanisms that give rise to consciousness? This question has been the focus of extensive research, leading to the development of several prominent theories, including Global Neuronal Workspace Theory (GNWT) and Integrated Information Theory (IIT). Critically, however, the focus so far has been on testing each theory independently, gathering evidence for/against them separately, leaving open a crucial question: which theory has higher explanatory power when tested against each other directly?

COG TATE is a pioneering Open Science adversarial collaboration to bridge this gap and evaluate GNWT and IIT through two studies, named [Experiment 1](#) (EXP1) and [Experiment 2](#) (EXP2). In these experiments, multimodal empirical tests are conducted on human volunteers, combining magneto-electroencephalography (M-EEG), functional magnetic resonance imaging (fMRI) and intracranial electrocorticography (iEEG) along with behavioral and eye tracking measurements. The reason for this approach is to maximize the sensitivity and specificity to the tests of each hypothesis, while accounting for trade-offs between temporal and spatial specificity inherent to the currently available methods in human neuroscience.



Goals

The aim of the COG TATE project is to accelerate research on consciousness and establish a groundbreaking model for scientific practices in cognitive neuroscience at large, by demonstrating the impact of team-based adversary research and open data to address some of the major riddles in the field, much like established practices in other fields of inquiry such as physics and genomics.

Furthermore, the resulting products of this research include a large and unique multimodal database, high-end analysis tools, and a new paradigm for probing consciousness in naturalistic settings. All experimental procedures, multimodal datasets, and analysis tools developed in this project will be made openly available to the public. These products will propel further discoveries in the field of consciousness, and in cognitive neuroscience in general, which will exceed and outlast the direct outputs of the proposed studies.

Experiments

The COG TATE consortium performed two experiments:

In Experiment 1 (EXP1), a set of clearly visible task relevant and irrelevant stimuli was shown to the subjects with different durations. The goal was to test the effects of maintenance of a percept in consciousness and task relevance and contradictory predictions regarding the involvement of prefrontal and posterior, category selective cortical areas in consciousness. Specifically, the main questions were: *How is the persistence of a stimulus in consciousness reflected in cortical hemodynamic and electrophysiological activity, i.e., are the neural responses phasic or sustained throughout a conscious experience? Do activity patterns in prefrontal areas relate to visual consciousness per se or to its consequences, i.e., task-related processes?*



In [Experiment 2](#) (EXP2), a novel paradigm was developed to test the key predictions of GNWT and IIT while overcoming a major obstacle in the field: *creating more naturalistic conditions of invisibility that do not degrade the physical input*. To achieve this goal, an engaging video game was used with the help of which salient stimuli were presented for relatively long durations in the background. Sometimes the stimuli was not consciously seen due to attentional engagement by the game. This approach allowed us to uniquely study neural activity elicited by seen or unseen stimuli under naturalistic conditions so that the stimuli can either be task relevant or task irrelevant.

Experiment 1: Conscious Perception

Objective

The primary aim of this experiment was to investigate neural activity in response to stimuli that are consciously perceived. It was designed to manipulate two key factors:

1. **Relevance of the Stimulus to the Task:** This factor was categorized into three levels—Task-relevant target, Task-relevant non-target, and Task-irrelevant stimulus.
2. **Stimulus Duration:** The stimuli were presented for durations of 500 ms, 1000 ms, and 1500 ms

This design framework allowed us to test several key hypotheses, including:

- Disentangling consciousness-related activations from task-related activations.
- Identifying brain regions that convey information about the content of consciousness.
- Examining the persistence of the content of consciousness over time.

Design

This experiment followed a 3x3x4x2 factorial design, with the following items:

Relevance of Stimulus to the Task (3)	<ul style="list-style-type: none">• Task-relevant target• Task-relevant non-target• Task-irrelevant stimulus
Stimulus Duration (3)	<ul style="list-style-type: none">• 500 ms• 1000 ms• 1500 ms
Stimulus Category (4)	<ul style="list-style-type: none">• Faces• Objects• Letters

	<ul style="list-style-type: none"> • False-fonts (meaningless symbols) <p>*20 identities for each category</p>
Stimulus Orientation (2)	<ul style="list-style-type: none"> • Side view (25% right and 25% left) • Front view (50%)

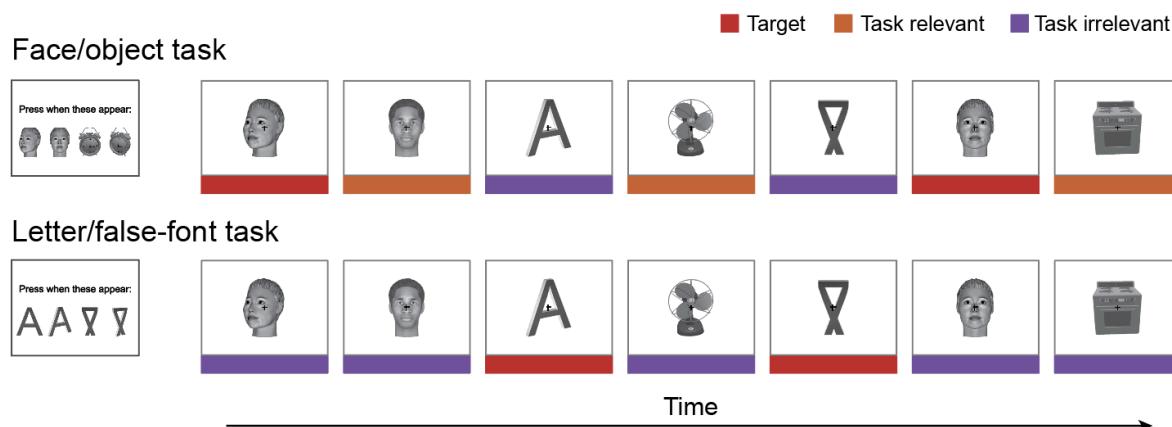
Sample Size

The sample sizes were determined based on common practices in the literature, resulting in a total of 122 subjects for fMRI, 102 for M-EEG, and 38 for iEEG. All subjects met specific criteria, including age and health conditions, to ensure data quality.

Task Description

A sequence of images including faces, objects, letters or meaningless symbols ('false fonts') with front or side (left or right) view were presented to the subjects. At the beginning of each sequence, the target images were presented and subjects were asked to memorize and remember them during the sequence. Subjects were instructed to press any buttons with their index finger when they saw targets (in either front or side views) as quickly and accurately as possible.

The duration of each sequence was approximately 2 minutes. The next sequence started when the subjects pressed the space key. Here is an example of the tasks:



For a comprehensive summary of more details about the experiments, please refer to the following supplementary resources:



[**COG TATE Main Scientific Paper 1 \(MSP-1\)\)**](#)

[**COG TATE Preregistration, v4**](#)



[**EXP1 Demo Video**](#)

Experiment 2: Video Game Engagement

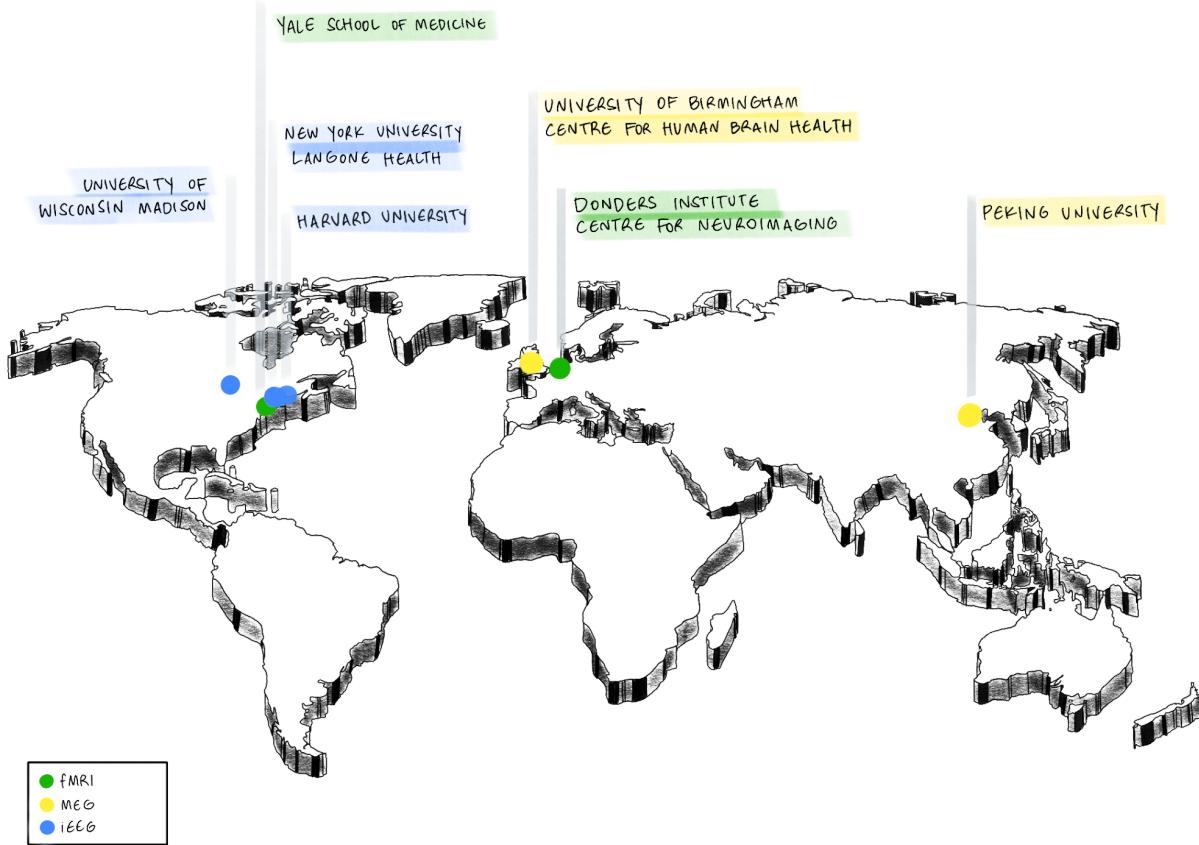
Not included in this document. It will be released soon!

Task Code and Stimuli repositories

The **task code** and **stimuli** used for EXP1 and for all modalities are available at [COG TATE code experiments](#).

COG TATE Dataset

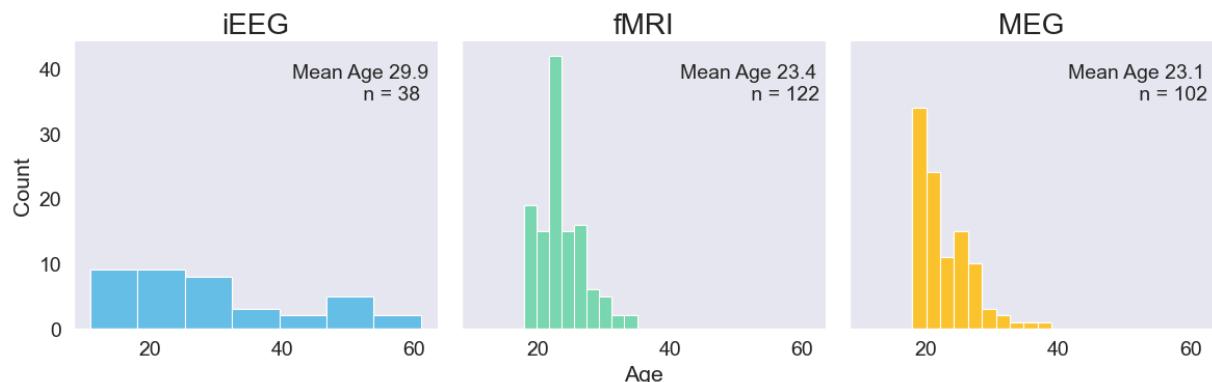
The COG TATE dataset is a comprehensive collection of multimodal neuroimaging data, encompassing a total of 262 subjects. COG TATE employs three distinct neuroimaging techniques: fMRI, M-EEG, and iEEG/ECoG.



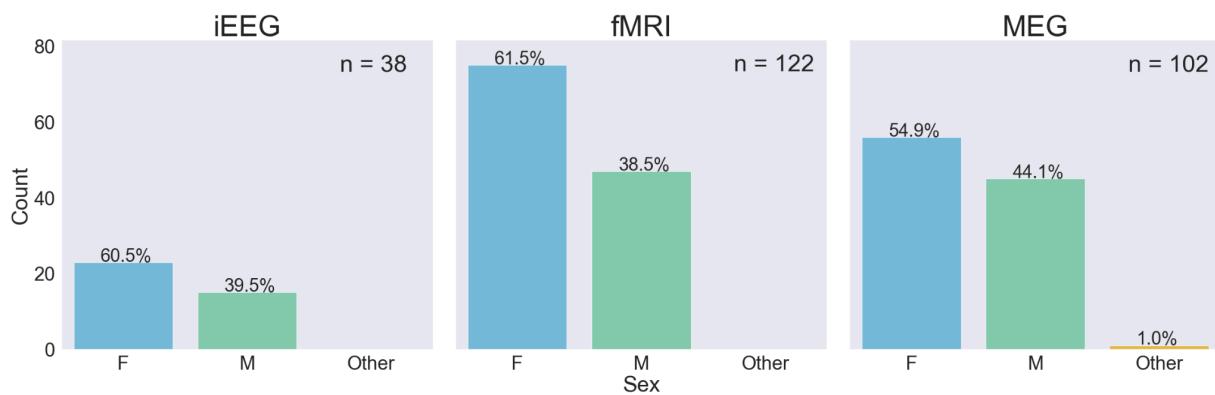
The MEG modality comprised 102 healthy subjects, also above the age of 18, with no known psychiatric or neurological issues. These participants were sourced from the Centre for Human Brain Health at the University of Birmingham (Birmingham, United Kingdom) and the Center for MRI Research of Peking University (Beijing, China).

Similarly, the fMRI modality included 122 healthy volunteers, all of whom were above the age of 18 and predominantly right-handed. These participants had no known history of psychiatric or neurological disorders and were recruited from the Yale Magnetic Resonance Research Center (New Haven, CT, United States) and the Donders Centre for Cognitive Neuroimaging (Nijmegen, Netherlands).

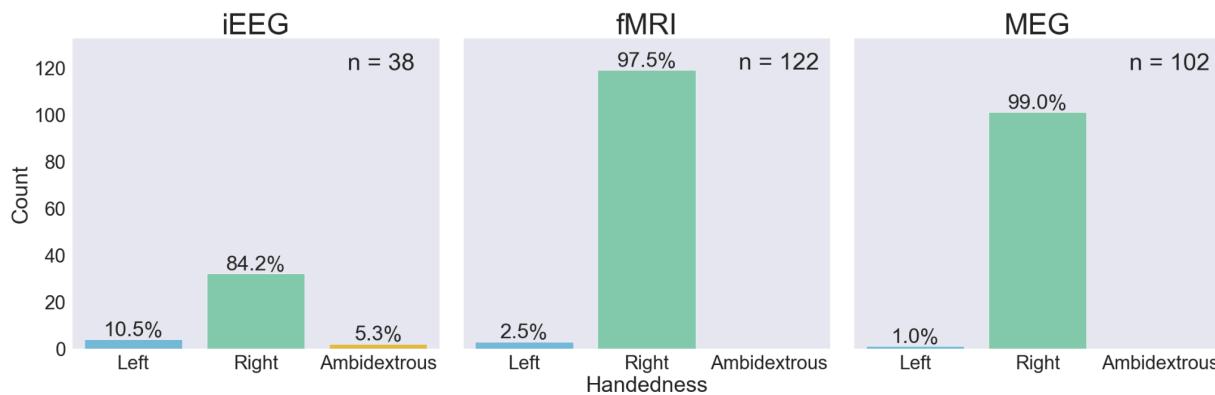
In contrast, the iEEG modality involved a more specialized cohort of 38 patients diagnosed with pharmaco-resistant focal epilepsy. These participants ranged in age from 10 to 65 years, had an IQ above 70, and met specific health criteria. They were recruited from multiple medical centers specializing in epilepsy treatment, including the Comprehensive Epilepsy Center at New York University (New York, NY, United States), Brigham and Women's Hospital, Boston Children's Hospital (Boston, MA, United States), and the University of Wisconsin School of Medicine and Public Health (Madison, WI, United States).



Age histograms across modalities



Sex proportions across modalities



Handedness proportions across modalities



Demography of Subjects

You can find the profile of participants for all modalities at [subjects_demography](#). Here is a brief explanation about the information collected from the subjects.

Demographic Information for M-EEG

The below items are included in the subjects' demography for M-EEG modality:

Participant_ID (participant identifier), sex (biological sex of participant), age (age of participant at the time of testing), handedness (right, left or ambidextrous), included in [MSP](#) (whether the data was used for the experiments or not), phase* (determining in which stage the data is, phase 2/phasel (optimization) or phase 3/phaselII (replication)), QC** status (passed/not), if Not (in QC status) QC rejection reason, weight (weight of participant at the time of study in pounds), height (height of participant at the time of study in inches), primary and secondary language, race (ethnicity of the participant), education, compensation (whether the subject was paid or not), colorblind (determining whether the subject can distinguish the colors and perceiving them correctly or not), visual correction (none or using any glasses or contact lenses), eye dominance (which eye is dominant), eye chart results (the outcome of a visual acuity test performed using the eye chart) and dioptre (visual acuity of the participant in Diopters).

***Phase:** COG TATE project has three phases. In **phase 1**, all data were acquired by theory neutral teams. To ensure replicability of the results, the entire dataset was split into two halves, each with an equal mixture of data from each of the labs for each modality. In **phase 2**, after evaluating data quality, the first half of the data were used for developing analysis tools (optimization of methods). The purpose of **phase 2** was to define the best analysis practices and to agree upon, in consultation with expert advisors. In **phase 3**, the replication phase, the second half of the data were analyzed using the concurred procedure, agreed upon protocols, thereby allowing an in-house replication of the results obtained in phase 2.

****QC (quality control):** A number of items were checked for all the data of each modality which are elaborated in the section of [Quality Check](#) and [Exclusion Criteria](#).

Demographic Information for fMRI

All of the items are similar to the M-EEG modality.

Demographic Information for iEEG

In addition to the properties mentioned for M-EEG modality, the below parameters were also provided for this modality:

Electrode scheme (the scheme used for implanting the electrodes, Stereo, Subdural grid & Strips), number of implanted electrodes, implant hemisphere (brain hemisphere where the

electrodes implanted, right, left, both or bilateral), IQ (score and name of the test used for assessment (FSIQ, WISC, VCI, POI, WMI, PSI, AMI, VMI), WADA (intracarotid sodium amobarbital, a test that determines which side of the subject's brain controls language and memory functions), seizure type (classification of seizure type), age of onset (age at which the first symptoms of seizure appeared), auditory normal hearing (indicator of whether the participant had normal hearing capabilities, yes or no), epilepsy seizure classification (categorization of epilepsy as per standard seizure classification), epilepsy seizure aura (description of any sensory or perceptual symptoms before a seizure occurred), epilepsy seizure semiology (signs and symptoms exhibited during epileptic seizures), epilepsy seizure frequency (frequency of seizures experienced by participant), epilepsy post ictal semiology (symptoms and signs after an epileptic seizure), epilepsy trigger (identified factors or circumstances that increased the likelihood of experiencing a seizure), epilepsy duration uncontrolled (the duration that seizures had not been successfully managed or medically controlled), epilepsy seizure onset zone (brain region identified as the initial site of seizure activity), epilepsy resection (details of any surgical resection performed for seizure control), epilepsy language lateralization (determination of the dominant hemisphere for language function), epilepsy past surgical history (record of any previous surgeries related to the treatment of epilepsy), epilepsy past medical history (medical history relevant to epilepsy diagnosis and treatment), epilepsy family history (presence of seizure or epilepsy disorders in family members), other neurological disorders (any other diagnosed neurological disorders besides epilepsy), epilepsy MRI findings (summary of MRI findings relevant to epilepsy diagnosis), epilepsy pathology findings (pathological findings from tissue analysis post-surgery or biopsy).

Quality Check

Data from all modalities were checked at three levels. The first level checks tested whether the datasets contained all expected files keeping their naming conventions, and that all personal information had been removed. The second level checks tested subjects' performance with respect to behavior. For [Experiment 1](#), subjects were excluded if their hit rate (Hit) was lower than 80% or false alarm (FA) was higher than 20% for M-EEG and fMRI. The third level checks assessed the quality of the neural data.

Exclusion Criteria

The generic exclusion criteria used across [Experiment 1](#) and [Experiment 2](#) included: (a) insufficient number of trials in each of the experimental conditions (<30 for M-EEG or <20 for fMRI), due to excessive muscular artifacts, movement, noisy recording, or subjects deciding to stop the experiments. If a given analysis showed that a good enough signal could be obtained with fewer trials, these numbers were amended; and (b) low performance in the attention tasks. In [Experiment 1](#), this translates into: <80% Hits, >20% FAs for fMRI and M-EEG subjects. In addition, data was excluded from analysis if it did not pass any of the predefined data quality checks.



Description of COG TATE Data

Although our data collection had a specific purpose, the data we gathered holds potential value for a range of diverse inquiries. Consequently, the COG TATE consortium has chosen to openly share all raw data collected (including the data that did not pass the quality criteria), to facilitate its utilization for various research endeavors and promote data reusability.

We have made available two primary formats for the data acquired during the experimental phase of the COG TATE project, specifically [Experiment 1](#):

1. Unprocessed/Raw Data
2. BIDS Format

1. Unprocessed/Raw Data

The unprocessed data format closely resembles the original acquired data, having undergone minimal processing to ensure compliance with [GDPR](#) (General Data Protection Regulation)/[HIPAA](#) (Health Insurance Portability & Accountability Act) anonymity standards.

2. BIDS Format

BIDS format, widely adopted in cognitive neuroscience, enhances data reusability. To facilitate others in leveraging our data, we have released it in [BIDS](#) format.

File Type Glossary

Here are the various file formats used for each modality of the COG TATE dataset along with a short description of them.

	Unprocessed /Raw	BIDS Format	Description
Eye Tracking & Behavioral Data	ASC/CSV	ASC/CSV	<p>The two eye trackers used within COG TATE are:</p> <ol style="list-style-type: none">1. EyeLink eye tracker2. Tobii eye tracker <p>1) EyeLink eye tracker: Most of the sites used this eye tracker which produces data in the EDF format, EyeLink Data Format. This data was immediately converted to ASCII text files using the converter provided by EyeLink. This is the ASC files that we used in our data analysis.</p>

			<p>2) Tobii eye tracker: The other eye tracker was the Tobii eye tracker used by New York University Langone for ECOG data. This eye tracker produces data in the form of CSV files.</p> <p>The files generated by eye tracking systems, containing information about eye movement and gaze behavior which typically store a time-stamped sequence of gaze data points and include information such as:</p> <ol style="list-style-type: none"> 1. Timestamps: The exact time at which each gaze data point was recorded. 2. Gaze Coordinates: The x and y coordinates on the screen where the person's gaze is directed. 3. Pupil Diameter: The size of the person's pupil, which can provide insights into changes in visual processing or cognitive load. 4. [EyeLink eye tracker] Fixations: Periods of stable gaze where the person is looking at a specific point without significant movement. 5. [EyeLink eye tracker] Saccades: Rapid eye movements between fixations, indicating shifts in attention. 6. [EyeLink eye tracker] Blinks: Instances when the person's eyes are closed, which can be important for data cleaning and analysis. 7. Triggers: Description of experimental events paired with their corresponding timestamps indicating when each event occurred. <p>Behavioral data is available in CSV format and it provides below information:</p> <ol style="list-style-type: none"> 1. Blocks 2. Events 3. Trials 4. Stimulus and jitter duration 5. Subject's responses
M-EEG	FIF	FIF	<p>File Format for the Input and Output of MEG and EEG data</p> <p>FIF files contain various types of information related to neuroimaging data, including:</p> <ol style="list-style-type: none"> 1. Raw sensor data: MEG and EEG measurements recorded from sensors placed on the scalp or near the head.

			<ol style="list-style-type: none"> 2. Event information: Time-stamped triggers or markers indicating the timing of events, such as stimulus presentations or subject responses. 3. Sensor locations and orientations: Information about the physical positions and orientations of sensors used in the measurements. 4. Head geometry: Information about the shape and structure of the subject's head, which is crucial for accurate source localization. 5. Covariance matrices: Statistical information about the relationships between sensor measurements at different time points or frequencies. 6. Anatomical MRI data: High-resolution structural images of the subject's brain, used for source localization and spatial alignment.
ECoG (iEEG)	EDF	BrainVision	<p>European Data Format files used for storing and exchanging time-series biological and physiological data</p> <p>EDF files are designed to accommodate data from multiple channels, allowing researchers to store and manage data collected simultaneously from different sensors or electrodes. The format supports both raw signal data and associated metadata, including information about sampling rates, units of measurement, patient demographics, and recording conditions.</p> <p>The BrainVision format, often shortened to BV, is a widely employed file format in neuroscience research, specifically designed for organizing neurophysiological data. It comprises three files, including the header file (.vhdr), marker file (.vmrk), and raw EEG data file (*.eeg), alongside possible auxiliary files. These components store essential details like recording parameters, event markers, and the raw EEG data.</p>
MR/CT	DICOM/NIFTI	DICOM/NIFTI	<p>DICOM (.dcm file extension) is a standard format utilized for storing CT (Computed Tomography) scans and MRI (Magnetic Resonance Imaging) data. These files encompass not only the image data but also essential metadata, including imaging parameters.</p> <p>NIFTI (.nii.gz file extension) serves as another format employed for a subset of subjects where our standard procedure encountered challenges. With the exception of the MR and CT scans for 12 subjects within the</p>

			iEEG data, all other datasets of similar nature are stored in DICOM format. Further details regarding these 12 problematic datasets are available in this section . NIFTI files encapsulate image data alongside metadata concerning spatial orientation, voxel dimensions, and additional imaging parameters.
--	--	--	--

Data Acquisition

The Cogitate dataset encompasses three distinct neuroimaging modalities, along with synchronized eye-tracking and behavioral data linked to each of these modalities. Here we detail the acquisition protocol for each modality in the corresponding data release: M-EEG, iEEG

Stimuli

Stimuli belonged to four categories that naturally fell into two groups that were clearly distinct from each other: pictures (20 faces and 20 objects) and symbols (20 letters and 20 false-fonts). Face stimuli were created using the FaceGen Modeler 3.1 program and object stimuli were taken from the Object Databank (Tarr, 1996). Faces and objects were grey-scaled (RGB: 125, 125, 125), and manipulated to have similar size and equal luminance using the SHINE toolbox (Willenbockel et al., 2010). Equal proportions of male and female faces were presented. They all had hair and belonged to different ethnicities (e.g., Caucasian, Asian, African, American) to facilitate face individuation. The orientation of the stimuli was manipulated, such that half of the stimuli from each category had a side view (30° and -30°) and the other half a front view (0°). All letter stimuli and false fonts were generated with MAXON CINEMA 4D Studio (RC-R20) 20.059 on macOS 10.14, appearing in gray (RGB: 125, 125, 125). Three views were rendered for each font set (real font, false/pseudo font) at 0°, 30° and -30° horizontal viewing angle with the following settings: Extrusion depth 9.79% of character height, camera distance 5.65 times character height and 18° above the center of the letter (High Angle), with a simulated focal length of 135 mm (35 mm equiv.). All stimuli were presented on a rectangular aperture at an average visual angle of 6° by 6°.

Procedure

Term	Definition
Stimuli	Visual items presented to the subjects during the experiment, with half being task-relevant and the other half task-irrelevant.
Trial	A single instance of presenting a stimulus to the subject and recording their response. Each trial consisted of presenting a single stimulus for a specific duration.

Block	A segment within each run containing a set of trials with a specific experimental setup or condition.
Break	Intervals between blocks that allowed subjects to rest and prepare for the next series of trials.

Stimuli were presented sequentially, all supra-threshold, with half being task-relevant and the other half task-irrelevant. Only one stimulus was shown on the screen at any given time. To define task relevance, subjects were instructed to detect (press a button; non-speeded response) two targets from different categories, regardless of their orientation. This online reporting enabled an explicit assessment of subjects' performance, engaging report-related areas for later analysis. Each block began with notification of the two target stimuli, either pictorial (faces and objects) or symbolic (letters and false fonts), creating a clear distinction between relevant and irrelevant stimuli. At the start of each block, specific target stimuli were revealed with instructions such as "detect face A and object B" or "detect letter C and false-font D." Targets did not repeat across blocks. Each run included two blocks of the Face/Object task and two blocks of the Letter/False-font task, with the order counterbalanced across runs. Subjects were instructed to maintain central fixation throughout each trial. Gaze was monitored online through an eye tracker, with repeated calibrations ensuring good quality data.

Each block comprised stimuli from all four categories, with each stimulus displayed for 500, 1000, or 1500 ms, followed by a blank interval, ensuring a consistent trial duration of 2000 ms. To avoid periodic presentation of the stimuli, random jitter was added to the end of each trial (mean inter-trial interval of 400 ms, jittered 200-2000 ms, with truncated exponential distribution). Within each block, three trial types were presented: i) Task Relevant Targets, consisting of the specific stimuli participants were tasked with detecting; ii) Task Relevant Non-Targets, encompassing stimuli from relevant categories that were not designated targets; and iii) Task Irrelevant Stimuli, comprising stimuli from the remaining categories.

Trial division was pseudorandomized with respect to two factors: the assignment of non-target stimuli to a miniblock, and the number of trials with specific durations within each task-relevance level. The assignment of non-targets was restricted by two constraints: i) Each stimulus was presented an equal number of times as a non-target stimulus (and as an irrelevant one) throughout the experiment. ii) Non-targets were, by definition, different from the targets in that specific miniblock. To meet these constraints, two pseudorandomized versions were made where the targets and non-target stimuli of each miniblock were determined. Three additional versions were prepared for each target/non-target division to balance the duration division between categories and orientation. In each version, slight imbalances (maximum difference of four trials) existed between the different durations for each condition (targets, non-targets, irrelevant). This imbalance was a necessary consequence of keeping the balance in duration and category across orientation and across categories as the numbers were not evenly divisible by three. The imbalances canceled out over the entire experiment so there was a perfect duration balance across the experiment: an equal amount of stimuli were presented for 500, 1000 and 1500 ms durations. The imbalances between durations within each relevant condition



were balanced across the three versions. In total, six different versions were assigned to the subjects in consecutive order, repeating every six subjects. To get more details, please see the “Trial counts” section in [COG TATE Preregistration, v4](#).

iEEG Related Modifications of the Design

Only half of the stimuli were used as targets for the iEEG experiments. The selection of target faces kept the balance of gender and ethnicity. Letters were chosen in equal amounts from the first and second parts of the alphabet, and the corresponding false-fonts were used. All stimuli were presented as relevant non-targets and irrelevant stimuli, matching the designs of the aforementioned procedure.

M-EEG Data Acquisition

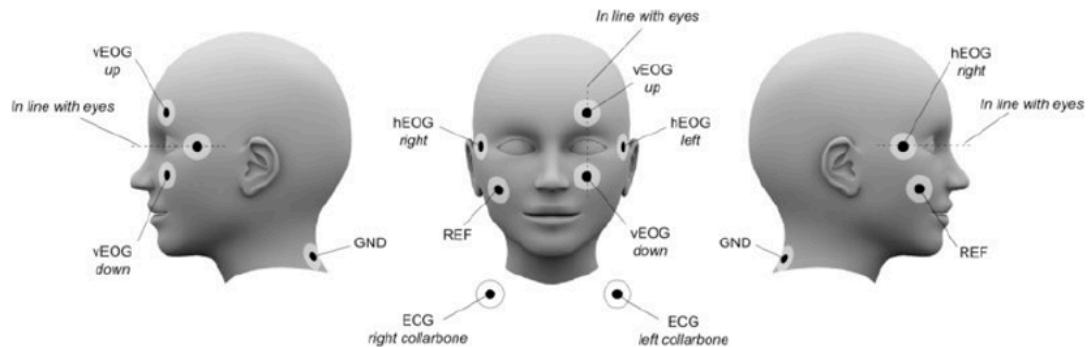
M-EEG recordings were acquired at the Centre for Human Brain Health (CHBH) of University of Birmingham in the United Kingdom, and at the Center for MRI Research of Peking University (PKU) in China.

Hardware

Both centers had a 306-channel, whole-head TRIUX MEG system from MEGIN (York Instruments; formerly Elekta). The MEG system comprised 204 planar gradiometers and 102 magnetometers in a helmet-shaped array. Simultaneous EEG was recorded using an integrated EEG system and a 64-channel electrode cap. The MEG system was equipped with a zero boil-off Helium recycling system and the noise-resilient ARMOR sensors and placed in a shielded room (2 layers of mu-metal and 1 layer of aluminum). To reduce environmental noise, the integrated active shielding system was used at PKU. In order to cover the brain more homogeneously, the MEG gantry was positioned at 68 degrees.

Location of Electrodes and ECG/EOG Measurements

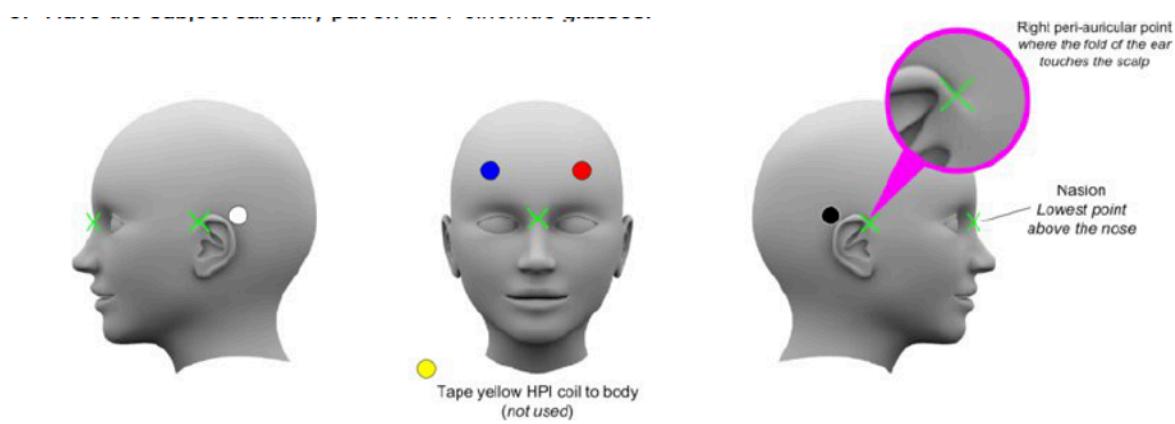
The location of the fiducials, the positions of the 64 EEG electrodes and the participant’s head shape were recorded using a 3-D digitizer system (Polhemus Isotrak). A set of bipolar electrodes were placed on the subject’s chest (upper left and upper right chest position) to record the cardiac signal (ECG). Two sets of bipolar electrodes were placed around the eyes (two located at the outer canthi of the right and left eyes and two above and below the center of the right eye) to record eye movements and blinks (EOG). Ground and reference electrodes were placed on the back of the neck and on the right cheek, respectively. The impedance of all of the electrodes was checked to be below 10 kOhm.



Standard locations of EOG and ECG electrodes

Head Position Indicator (HPI) Coils

The participant's head position inside the MEG system was measured at the beginning and at the end of each run using four head position indicator (HPI) coils placed on the EEG cap. Specifically, the HPI coils were placed next to the left and right mastoids and on the left and right forehead. Their location relative to anatomical landmarks was digitized with a Polhemus Isotrak System. During the measurement, high frequency (>200 Hz) signals were produced by those coils and the localization of these signals was used to estimate the head position in the sensor space. To avoid the potential artifacts produced by the non-linear interaction between the signals generated by these coils, head position measurement was performed only during resting periods (as opposed to continuously).



Standard locations of HPI coils. Coil Numbers: 1. Blue, 2. White, 3. Red, 4. Black, 5. Yellow

Anatomical MRI Data Acquisition

For each subject, a high resolution T1-weighted MRI volume (3T Siemens MRI Prisma scanner) was acquired. At CHBH, a 32-channel coil with a resolution of 1 x 1 x 1 mm, TR/TE = 2000/2.03 ms; TI = 880 ms; 8° flip angle, 208 sagittal slices and field of view (FOV): 256 x 256 matrix was acquired for source localization with individual realistic head modeling. At PKU, a 64-channel coil with a resolution of 0.5 x 0.5 x 1 mm, TR/TE = 2530/2.98 ms; TI = 1100 ms; 7° flip angle, 192 sagittal slices; FOV: 448 x 512 matrix was used. To avoid possible interference of body magnetization on the MEG recording, all MRI scans were acquired at least one week before the MEG session, or at any time afterwards.

Subject ID	MRI Imaging Technique	Sequence Type	TE (ms)	TR (ms)	Flip Angle	Resolution (mm)	Slice Thickness (mm)	Dimensions	Manufacturer	Model Name
All CA??? subjects	T1-weighted	MPRAGE	2.03	2000	8	1 x 1	1	256 x 256 x 208	SIEMENS	Prisma
All CB??? subjects	T1-weighted	MPRAGE	2.98	2530	7	0.5 x 0.5	1	448 x 512 x 192	SIEMENS	Prisma

Behavioral Data Acquisition

The task was executed using Matlab (CHBH: R2019b, PKU: R2018b) with Psychtoolbox v.3 (Pelli, 1997) on a custom PC at CHBH and a Dell XPS desktop PC at PKU. Visual stimuli were presented on a screen placed in front of the subjects with a PROPiXX DLP LED projector (VPiXX Technologies Inc.) at a resolution of 1920 x 1080 pixels and a refresh rate of 120 Hz. The distance between the subject's eyes and the screen was different at each site (CHBH: 119 cm, PKU: 85 cm) to achieve the same FOV of 36.6 x 21.2 degrees. Participants responded with both hands using two 5-button response boxes (CHBH: NAtA, PKU: SINORAD).

Eye Tracking Data Acquisition

Eye movements were monitored and recorded from both eyes (binocular eye-tracking) using the MEG-compatible EyeLink 1000 Plus eye-tracker (SR Research Ltd., Ottawa, Canada). Nine-point calibration was performed at the beginning of the experiment, and recalibrated if necessary at the beginning of each block/word. Pupil size and gaze location (x, y) in screen pixel coordinates were collected at a sampling rate of 1000 Hz.

The channel name that contains the eye tracker data in the FIF file is as follows: MISC1 (X), MISC2 (Y), and MISC3 (pupil).

Behavioral Data Code Scheme

Stimuli are coded as a 4-digit number.

- 1st digit = stimulus type (1 = face; 2 = object; 3 = letter; 4 = false font)
- 2nd digit = stimulus orientation (1 = center; 2 = left; 3 = right)



- 3rd & 4th digits = stimulus id (1...20; for faces 1...10 is male, 11...20 is female)

e.g., "1219" = 1 is face, 2 is left orientation and 19 is a female stimulus #19

Eye Tracker and MEG Code Scheme

Successive trigger scheme

The triggers were sent successively. The first trigger represented the stimulus type, followed by orientation, stimulus duration, and task relevance, all interspaced by 50 ms. Additionally, a trigger was sent upon key press.

1st Trigger (on Stimulus Onset): Stimulus Type

- 1 to 20: faces
 - 1 to 10 males,
 - 11 to 20 females
- 21 to 40: objects
- 41 to 60: letters
- 61 to 80: falses

2nd Trigger (2 Frames after Stimulus Onset): Stimulus Orientation

- 101: Center
- 102: Left
- 103: Right

3rd Trigger (4 Frames after Stimulus Onset): Stimulus Duration

- 151: 500 msec
- 152: 1000 msec
- 153: 1500 msec

4th Trigger (6 Frames after Stimulus Onset): Stimulus Task Relevance

- 201: Task relevant target
- 202: Task relevant non target
- 203: Task irrelevant

5th Trigger (8 Frames after Stimulus Onset): Trial ID Triggers

- 111-148: Trial number

Response Trigger

- 255: Following button press.

Stimulus Presentation End

- 96: Offset of stimulus presentation (onset of blank)
- 97: Offset of blank (onset of jitter period)
 - Note that both these are fixations, they are just divided into blank and jitter.

General Triggers to Mark Experiment Progression

86: Onset of experiment

81: Onset of recording

83: Offset of recording

Miniblock ID Triggers

161-200: Miniblock ID trigger

Zeroes

0: Zeros were sent between the successive triggers to reset the LPT, see below. These were also sent to the eye tracker but did not mean anything and they can safely be ignored.

How The LPT Triggers Were Sent

The LPT port of the computer was used for sending the triggers and it was done by using the sendTrig function. This function sets the port in a specific state (whatever trigger we want to send) and logs the trigger afterwards, noting if it is sent and what time the command for sending it is executed. For each trigger that is being sent, the port is being reset after a frame to 0.

In the beginning of the experiment, a few triggers were sent to mark experiment onset and onset of recording. Then, a miniblock was initiated. The participant was presented with the target screen and required to press the spacebar to proceed. When the participant pressed the space button, the miniblock ID was sent. Only once the miniblock trigger was sent the fixation appeared. This means that there was a small delay between key press and fixation onset. Following the first fixation, a jitter started, which was also logged. Then, the first stimulus was displayed. Upon the presentation of the stimulus, the successive triggers were initiated. The first trigger occurred directly after the onset of the stimulus, indicating the stimulus ID (1-80). Then, after 2 frames, the orientation trigger (101-103) was sent, followed by the duration trigger (151 to 153) at 4 frames, the task demand trigger (201-203) at 6 frames, and finally, the trial ID trigger (111 to 148) at 8 frames.

Empty Room Recording

Prior to each experiment, MEG signals from the empty room were recorded for 3-minutes.



Resting-State (rM-EEG)

The resting-state data for each participant was also recorded for 5-minutes and the subjects were asked to keep their eyes open and fixated on a point presented at the center of the screen. M-EEG signals were sampled at a rate of 1 kHz and band-pass filtered between 0.01 and 330 Hz prior to sampling.

Task (tM-EEG)

Following the empty room and rM-EEG recordings, subjects were asked to complete the task defined in the [Procedure](#) section. tM-EEG consisted of 10 runs, with 4 blocks each. During each block, a ratio of 34-38 trials was presented, with 32 non-targets (8 of each category), 2-6 targets (number chosen randomly), and each trial lasting 2.4 s approximately. Rest breaks between runs and blocks were included. Random jitter was added at the end of each trial (mean inter-trial interval of 0.4 s jittered 0.2-2.0 s, truncated exponential distribution) to avoid periodic presentation of the stimuli.

Task	Runs	Blocks	Trials	Total trials
Experiment 1	10	4	34-38 per block	1440

Full Structure of Session

Complete standard procedure of an M-EEG session is available in [MEG Standard Operating Procedure](#).

Inclusion Criteria

The items below were assessed for the subjects before the data was acquired:

- Age range: 18 to 35 (since over the age of 35 subjects might have a hard time maintaining central focus)
- Handedness: right
- Hearing problems: no
- Hearing aid: no
- Vision problems: no, or corrected-to-normal with soft lenses
- No MRI in the last week
- MRI compatible: no metal, medical implants, etc. No claustrophobia. Note: dental implants are allowed (particularly for non-magnetic materials) unless it generates big impacts on MEG signals, and this will be checked prior to MEG recording.
- No known history of psychiatric or neurological disorders, e.g.,
 - Not have been formally diagnosed with attention deficit (hyperactivity) disorder (AD(H)D).
 - Not have been formally diagnosed with autism spectrum disorder (ASD)

- Not suffer from epilepsy

Quality Check and Exclusion Criteria

For M-EEG, the first stage of the third-level checks focused on system-related and external noise generators. It was tested using the signal spectra in the empty room recording, the resting state session, and the experiment itself for all sensors. Any sensor and/or specific frequency revealing extensive noise using visual inspection, was flagged to document potential problems. Next, all experimental data blocks were visually inspected for abnormalities in spectra (peaks not explainable by physiology), and in ICA (Independent Component Analysis) components, and checked for extremely noisy (based on the score of differences between the original and Maxwell-filtered data > 7) and flat sensors. The latter step was performed in a collaboration between the data monitoring team and members of the centers where data was acquired to check whether any potential changes in preprocessing for particular subjects were needed. Finally, we tested if all experimental cells (i.e. task-relevant non-targets and task-irrelevant stimuli for each one of the four categories) have enough trials.

iEEG Data Acquisition

iEEG recordings were obtained from patients with pharmacologically resistant epilepsy undergoing invasive electrophysiological monitoring at the Comprehensive Epilepsy Center at New York University (NYU) Langone Health Center, Brigham and Women's Hospital, Children's Hospital Boston (Harvard Medical School), and University of Wisconsin School of Medicine and Public Health (WU).

Hardware

Brain activity was recorded with a combination of intracranially subdural platinum-iridium electrodes embedded in SILASTIC sheets (2.3 mm diameter contacts, Ad-Tech Medical Instrument and PMT Corporation) and/or depth stereo-electroencephalographic platinum-iridium electrodes (PMT Corporation; 0.8-mm diameter, 2.0-mm length cylinders; separated from adjacent contacts by 1.5 to 2.43 mm), or Behnke-Fried depth stereo-electroencephalographic platinum-iridium electrodes (Ad-Tech Medical, BF08R-SP21X-0C2, 1.28 mm in diameter, 1.57 mm in length, 3 to 5.5 mm spacing). The decision to implant, electrode targeting, and the duration of invasive monitoring was solely determined on clinical grounds and without reference to this or any other study. Electrodes were arranged as grid arrays (either 8 \times 8 with 10 mm center-to-center spacing, 8 \times 16 contacts with 3 mm spacing, or hybrid macro/micro 8 \times 8 contacts with 10 mm spacing and 64 integrated microcontacts with 5 mm spacing), linear strips (1 \times 8/12 contacts), depth electrodes (1 \times 8/12 contacts), or a combination thereof. Subdural electrodes covered extensive portions of lateral and medial frontal, parietal, occipital, and temporal cortex of the left and/or right hemisphere. Recordings from grid, strip and depth electrode arrays were done using a Natus Quantum amplifier (Pleasanton, CA) or a Neuralynx Atlas amplifier (Bozeman, MT). A total of 4057 electrodes (892 grids, 346 strips, 2819 depths) were implanted across 32 patients with drug-resistant focal epilepsy undergoing clinically motivated invasive monitoring. 3512 electrodes (780 grids, 307 strips, 2425 depths) that were

unaffected by epileptic activity, artifacts, or electrical noise were used in subsequent analyses. To determine the electrode localization for each patient, a postoperative CT (computed tomography) scan and a pre-operative T1 MRI were acquired and co-registered. Recordings were obtained continuously during the patients' stay in the hospital. All data was stored with stimulus and timing markers permitting offline synchronization.

Anatomical MRI Data Acquisition

Before the participants underwent surgery and electrode implantation, T1-weighted MR data were acquired from them. At NYU, imaging was performed using the Siemens Biograph mMR scanner. At Harvard, the imaging sequence utilized was MPRAGE (magnetization-prepared rapid gradient-echo), with a Siemens Skyra 3T scanner. At WU, imaging was conducted using different models of scanners from GE MEDICAL SYSTEMS. The rationale behind acquiring MR scans was the spatial resolution it offers for brain tissue visualization.

Subject ID	MRI Imaging Technique	Sequence Type	TE (ms)	TR (ms)	Flip Angle	Resolution (mm)	Slice Thickness (mm)	Dimensions	Manufacturer	Model Name
All CE??? subjects	T1-weighted	MPRAGE	2.5	2000	9	0.9 x 0.9	0.9	208 x 256 x 256	SIEMENS	Skyra 3T
All CF??? subjects	T1-weighted	MPRAGE	2.98	2300	9	1 x 1	1	208 x 256 x 256	SIEMENS	Biograph mMR
CG101	T1-weighted	CUBE VASC	3.624	8.764	13	0.48 x 0.48	1.2	512 x 512 x 300	GE MEDICAL SYSTEMS	SIGNA Artist
CG102	T1-weighted	Bravo Stealth	3.608	9.456	13	0.48 x 0.48	1.2	512 x 512 x 308	GE MEDICAL SYSTEMS	SIGNA HDxt
CG103	T1-weighted	Bravo Stealth	3.248	8.484	12	0.48 x 0.48	1.4	512 x 512 x 232	GE MEDICAL SYSTEMS	SIGNA Architect
CG104	T1-weighted	Bravo Stealth	3.116	7.8	12	0.48 x 0.48	1.19	512 x 512 x 282	GE MEDICAL SYSTEMS	SIGNA Architect
CG105	T1-weighted	Bravo Stealth	3.112	7.76	12	0.48 x 0.48	1.19	512 x 512 x 286	GE MEDICAL SYSTEMS	SIGNA Architect
CG106	T1-weighted	Bravo Stealth	3.752	9.52	13	0.48 x 0.48	1.2	512 x 512 x 272	GE MEDICAL SYSTEMS	Optima MR450w
CG107	T1-weighted	Bravo	3.112	7.572	12	0.43 x 0.43	1.98	512 x 512 x 204	GE MEDICAL SYSTEMS	Discovery MR750

CT Data Acquisition

Following surgery, post-operative CT scans were obtained from the subjects to assist in localizing the electrodes on specific brain tissue. At NYU, scans were performed using a Siemens SOMATOM Force scanner. At Harvard, imaging was conducted using the Medtronic O-arm MVS O2, manufactured by Medtronic. At WU, scans were acquired using various scanners manufactured by GE MEDICAL SYSTEMS.

Subject ID	Resolution (mm)	Slice Thickness (mm)	Dimensions	Manufacturer	Model Name
All CE???	0.4 x 0.4	0.8	512 x 512 x 192	Medtronic	O-arm MVS O2
All CF???	0.45 x 0.45	0.75	512 x 512 x 192	Medtronic	SOMATOM Force
CG101	0.5 x 0.5	1.25	512 x 512 x 160	GE MEDICAL SYSTEMS	Optima CT660
CG102	0.5 x 0.5	0.625	512 x 512 x 599	GE MEDICAL SYSTEMS	Optima CT660
CG103	0.5 x 0.5	0.625	512 x 512 x 157	GE MEDICAL SYSTEMS	Discovery CT750 HD
CG104	0.45 x 0.45	1.25	512 x 512 x 154	GE MEDICAL SYSTEMS	Revolution CT
CG105	0.43 x 0.43	2	512 x 512 x 202	GE MEDICAL SYSTEMS	Aquilion ONE
CG106	0.976 x 0.976	1	512 x 512 x 326	GE MEDICAL SYSTEMS	Aquilion ONE
CG107	0.47 x 0.47	3	512 x 512 x 153	GE MEDICAL SYSTEMS	Discovery CT750 HD

Behavioral Data Acquisition

The task was implemented using Matlab (Harvard: R2020b; NYU: R2020a, WU: 2021a), Psychtoolbox v.3 (Pelli, 1997), and run on a Dell Precision 5540 laptop, with a 15.6" Ultrasharp screen (screen size 345 x 195 mm²; resolution 1920 x 1080) at NYU and Harvard and on a Dell D29M PC with an Acer V196WL 19" LED LCD monitor (screen size 406.4 x 254 mm²; resolution 1440 x 990) at WU. The distance between the subject's eyes and the screen was 80 cm. But the actual distance was measured for each subject before the start of recording to ensure that the size of the stimulus was 6 x 6 of visual angle. Participants responded using an 8-button response box (Millikey LH-8; response hand(s) varied based on the setting in the patient's room).

Eye Tracking Data Acquisition

At Harvard and Wisconsin, EyeLink 1000 Plus Camera was used to collect eye-tracking data, and a thirteen-point calibration was performed several times during the experiment. The calibration was performed at the beginning of the experiment, and recalibrated in-between blocks, if necessary to meet precision requirements. At NYU, eye-tracking data was collected throughout the duration of the experiment using a Tobii-4C eye-tracker, and a nine-point calibration was performed several times during the experiment. Pupil size (a.u.) as well as x and y gaze coordinates on the screen (in screen pixels) were recorded at a sampling rate of 500 Hz at Harvard and Wisconsin and at a sampling rate of 90 Hz at NYU. The Eyelink system recorded monocular data, while the Tobii system recorded binocular data. For the former cases, only one

eye was recorded as determined by ocular dominance. The experiment was not influenced by the Eye-tracking recording.

Behavioral Data Code Scheme

The behavioral code scheme is similar to the M-EEG modality which is explained in [this section](#).

Eye Tracker Data Code Scheme

The eye tracker code scheme for the iEEG modality follows a similar structure to that described for M-EEG data. You can find detailed explanations [here](#).

iEEG Code Scheme

Photodiode Trigger Scheme

For ECOG patients, the type of port utilized by the M-EEG team (LPT) was incompatible with our recording system. Consequently, a photodiode was employed. A photodiode is an electronic device that records changes in luminance and converts them into voltage.



An example of a photodiode

In the experimental code, it was ensured that when a new event occurred on the screen (such as stimulus onset or stimulus offset), a white flash appeared in the bottom right corner. The photodiode device was positioned atop the flashing square and connected to the amplifier recording the iEEG channel signals. This additional channel facilitated the identification of event onsets in our task. This type of recording only allows binary signals (the photodiode is either on or off). However, specific events were encoded with varying numbers of subsequent pulses.

Stimulus Presentation Onset

The flashing square was flashed only once at the onset of each new stimulus.

Stimulus Presentation Offset

The flashing square was flashed only once at the offset of each stimulus.

Start of the Inter-Trial Interval

The flashing square was flashed only once at the beginning of the inter-trial interval. The inter-trial interval was initiated 2 seconds after stimulus onset and persisted for a random duration (following a truncated exponential distribution between 0.2 and 2 seconds, with a mean of 0.4 seconds).

Block Start

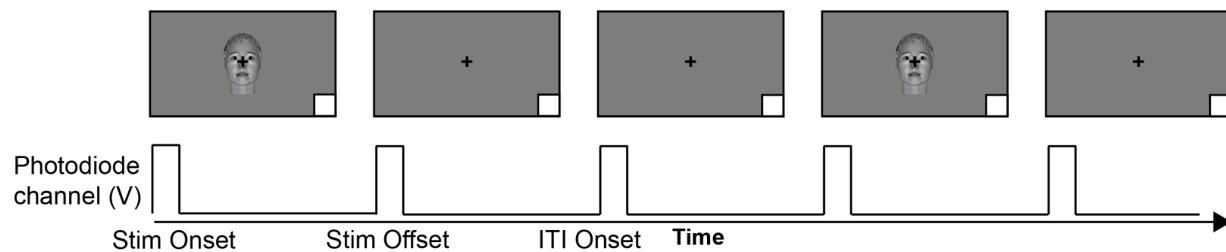
The start of an experimental block was marked by sending 4 consecutive pulses.

Block End

The end of an experimental block was marked by sending 2 consecutive pulses.

Experiment Start and End

The beginning and end of the experiment were marked by sending 3 consecutive pulses.



Schematic representation of the photodiode channel

Log File Alignment

The photodiode channel solely indicates when a specific event occurred in the iEEG signals, lacking any information about the nature of the event (unlike an LPT trigger in MEG). To identify specific events in the signal, the timing information from the log file was combined with that from the photodiode.

The log file contains a description of each presented event along with a corresponding time stamp from the experimental computer. The photodiode channel recorded time stamps for each event, indicating when it occurred according to the acquisition computer clock. The goal was to align the log file and the photodiode to associate each event in the photodiode signal with the corresponding event description in the log file. This step was crucial since misalignment could lead to incorrect event descriptions in the iEEG signal, compromising the entire analysis.

The procedure relies on the fact that both the log file and the photodiode had timestamps. These timestamps were recorded on different clocks. Unfortunately, computer clocks tend to

drift apart from one another, and these deviations accumulate significantly over extended periods of time. For instance, they could be several seconds apart after just one hour. Therefore, the timestamps of the photodiode and the log file could not be used interchangeably. However, over short periods of time, these drifts were negligible. What this meant was that the interval between two successive timestamps in the log file should be quite consistent with the intervals between two successive events in the photodiode. This provided us with the most thorough check possible: if the events in the log file and in the photodiode were aligned, then there should be only tiny differences between the differences between successive events in both. Here is a step-by-step description of the alignment procedure.

Extract the Photodiode Timestamps

The timestamps from the photodiode triggers were extracted as the first step. As illustrated in the figure "[Schematic representation of the photodiode channel](#)", a square pulse was generated for each event during the recording. The onset of each of these pulses was sought. To achieve this, a threshold was initially established, below which the photodiode was considered to be in the off state and above which it was considered to be on (based on visual inspection of the data, which was facilitated by the clean nature of photodiode signals). Subsequently, the signal was binarized using this threshold ($\text{signal_bin} = \text{signal} > \text{threshold}$), resulting in a signal consisting only of ones and zeros. Next, the discrete difference of the binary signal was computed ($y(i + 1) = y(i + 1) - y(i)$). This operation produced a “1” when the photodiode transitioned from off to on (onset) and a “-1” when it transitioned from on to off (offset). Since only the onset was of interest, the timestamps of the ones were extracted, representing the timestamps of the photodiode.

Verify Event Count Alignment

The first step in aligning the photodiode events and the log files was to check if the number of events in each matched. If they did not match, then there was a problem.

Aligning the Two Signals

To ensure alignment of both signals, the discrete difference between the photodiode and log file timestamps was computed, providing the interval between successive events for each signal. The resulting arrays were then plotted atop each other. Misalignment between the two sources of timing information could be easily detected, as they did not overlap. Perfect overlap between the two was necessary to consider the signals aligned. Additionally, the difference between the two signals was computed to ensure minimal deviation.

Integrating Information

Once the two signals were properly aligned, the log file events could be used as descriptors of the events marked at the timestamps from the photodiode.

The alignment procedure allowed the information from the log file to be utilized in generating well-described events in the iEEG data. The events were encoded as “/” separated strings and

stored in an events.tsv table. An extensive description of each event type and their levels are as follows:

Task Name: Dur

Task Description: Description of the experimental factors and their level with the correct syntax to access them from the MNE epochs object. Note that factor and level names are case-sensitive. We describe the syntax for each condition separately. However, note that you can fetch a combination of factors from the epochs using a forward slash. For example, if you wish to fetch the face target trials, you can combine both conditions like so: epochs['face/Relevant target'] will fetch all face target trials. In addition, the epochs objects are equipped with metadata, where the name of the column is the name of the factor, and the level follows the same nomenclature as below. This can be used for more sophisticated trial filtering and retrieval.

Experimental Design:

- Event Type:
 - Marks the different events occurring within a trial.
 - Factor Type: Categorical
 - Factor 1:
 - Name: stimulus onset
 - Description: Marks the onset of the visual stimuli. With epochs['stimulus onset'], extract all epochs time-locked to the visual stimulus onset.
 - Factor 2:
 - Name: stimulus offset
 - Description: Marks the offset of the visual stimuli. With epochs['stimulus offset'], extract all epochs time-locked to the offset of the visual stimulus.
 - Factor 3:
 - Name: jitter onset
 - Description: Marks the beginning of the inter-trial jitter. All trials lasted 2 sec., with an added jitter of 400 ms on average. With epochs['jitter onset'], extract all epochs time-locked to the beginning of the jitter period (2 sec. after stimulus onset).
- Block:
 - Marks the experimental blocks.
 - Factor Type: Discrete
 - Factor 1:
 - Name: block_*
 - Description: Experimental blocks. Our experiment consisted of 5 blocks, in between which participants were allowed to take a break. With epochs['block_1'], extract all epochs of the first experimental block.
- Miniblock:
 - Marks the experimental miniblocks.
 - Factor Type: Discrete
 - Factor 1:
 - Name: miniblock_*

- Description: Experimental miniblocks. Each experimental block consisted of 4 miniblocks. At the beginning of each miniblock, the two target stimuli were presented to the participant, which the participant had to remember to be able to detect in the stream of stimuli. With epochs['miniblock_1'], extract all epochs of the first experimental miniblock.
- Category:
 - Category of the visual stimuli.
 - Factor Type: Categorical
 - Factor 1:
 - Name: face
 - Description: Identifies face trials. With epochs['face'], extract all epochs in which a face was presented.
 - Factor 2:
 - Name: object
 - Description: Identifies object trials. With epochs['object'], extract all epochs in which an object was presented.
 - Factor 3:
 - Name: letter
 - Description: Identifies letter trials. With epochs['letter'], extract all epochs in which a letter was presented.
 - Factor 4:
 - Name: false
 - Description: Identifies false font trials (i.e., symbols). With epochs['false'], extract all epochs in which a false font was presented.
- Identity:
 - Identity of the visual stimuli.
 - Factor Type: Categorical
 - Factor 1:
 - Name: face_*
 - Description: Identifies the identity of face trials. With epochs['face_*'], extract all epochs in which that specific face was presented. From 1-9, leading 0.
 - Factor 2:
 - Name: object_*
 - Description: Identifies the identity of object trials. With epochs['object_*'], extract all epochs in which that specific object was presented. From 1-9, leading 0.
 - Factor 3:
 - Name: letter_*
 - Description: Identifies the identity of letter trials. With epochs['letter_*'], extract all epochs in which that specific letter was presented. From 1-9, leading 0.
 - Factor 4:
 - Name: false_*

- Description: Identifies the identity of false font trials (i.e., symbols). With epochs['false_*'], extract all epochs in which that specific false font was presented. From 1-9, leading 0.
- Orientation:
 - Orientation of the displayed stimuli.
 - Factor Type: Categorical
 - Factor 1:
 - Name: Center
 - Description: Identifies stimuli presented in the center orientation. With epochs['Center'], extract all epochs in which a stimulus was presented in the center orientation.
 - Factor 2:
 - Name: Left
 - Description: Identifies stimuli presented in the Left orientation. With epochs['Left'], extract all epochs in which a stimulus was presented in the Left orientation.
 - Factor 3:
 - Name: Right
 - Description: Identifies stimuli presented in the Right orientation. With epochs['Right'], extract all epochs in which a stimulus was presented in the Right orientation.
- Duration:
 - Duration a visual stimulus was presented for.
 - Factor Type: Categorical
 - Factor 1:
 - Name: 500ms
 - Description: Identifies stimuli presented for 500 ms. With epochs['500ms'], extract all epochs in which the stimulus was displayed for 500ms.
 - Factor 2:
 - Name: 1000ms
 - Description: Identifies stimuli presented for 1000 ms. With epochs['1000ms'], extract all epochs in which the stimulus was displayed for 1000ms.
 - Factor 3:
 - Name: 1500ms
 - Description: Identifies stimuli presented for 1500 ms. With epochs['1500ms'], extract all epochs in which the stimulus was displayed for 1500 ms.
- Task Relevance:
 - Task relevance of a given trial.
 - Factor Type: Categorical
 - Factor 1:
 - Name: Relevant target

- Description: Identifies target stimuli. Target stimuli are presented at the beginning of each miniblock, and participants must detect them among the sequence of presented stimuli by pressing a button. With epochs['Relevant target'], extract all target trials.
- Factor 2:
 - Name: Relevant non-target
 - Description: Identifies task-relevant non-target stimuli. We considered task-relevant stimuli that were of the same category as the target but of a different identity. With epochs['Relevant non-target'], extract all task-relevant non-target trials.
- Factor 3:
 - Name: Irrelevant
 - Description: Identifies task-irrelevant non-target stimuli. We considered task-irrelevant stimuli that were of a different category than the target. With epochs['Irrelevant'], extract all task-irrelevant non-target trials.
- Response:
 - Rated response of the participants.
 - Factor Type: Categorical
 - Factor 1:
 - Name: Hit
 - Description: Participants correctly identified a target by pressing a button. With epochs['Hit'], extract all target trials for which the participants pressed a key.
 - Factor 2:
 - Name: CorrRej
 - Description: Participants correctly rejected a non-target stimulus and did not press any button. With epochs['CorrRej'], extract all non-target trials for which the participants did not press a key.
 - Factor 3:
 - Name: Miss
 - Description: Participants failed to press a button when a target stimulus was presented. With epochs['Miss'], extract all target trials in which participants failed to press a button.
 - Factor 4:
 - Name: FA
 - Description: Participants mistakenly pressed a button when a non-target stimulus was presented. With epochs['FA'], extract all non-target trials in which participants pressed a button.
 - Factor 5:
 - Name: n.a.
 - Description: For the events stimulus offset and jitter onset, the response is set to n.a. as the response relates to the visual stimulus, not to the other events. This should not be used to access the data.

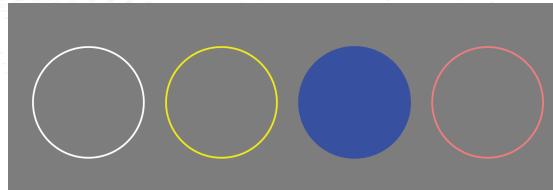
Surface Reconstruction and Electrode Localization

Subject-specific pial surfaces were automatically reconstructed based on a pre-implant T1 weighted MR image using the [Freesurfer](#) image analysis suite ('recon-all', Dale et al., 1999). Post-implant CT images were co-registered with the pre-implant MR images using FLIRT (Jenkinson and Smith, 2001), as implemented in [FSL](#) (Smith et al., 2004). For NYU patients, we used a semi-automatic approach to generating electrode labels. For manual cases, co-registered MR and CT slices were examined using FSLView (Smith et al., 2004). For grids, we localized three corner electrodes and the remaining electrodes coordinates were then automatically interpolated along the shared plane using the known inter-electrode distances. Strip and depth electrodes were localized manually when they did not follow straight trajectories. When depth electrodes were in a straight line, the first and last electrodes were localized manually, and electrodes in between were automatically interpolated and labeled based on known inter-electrode distances and serial labeling convention. For WU patients, electrodes were localized manually using the [SubNuclear](#) toolbox. Electrode locations were further refined within the space of the pre-operative MRI using three-dimensional non-linear thin-plate spline warping (Rohr et al., 2001), which corrected for post-operative shift and distortion. The warping was constrained with manually selected points through the brain, which was visually aligned with landmarks in pre-implantation MRI and post-implantation CT. For Harvard subjects, individual contacts from depth electrodes were labeled manually from the CT image using the [BioImageSuite](#)'s Electrode Editor tool (legacy version 3.5; Joshi, et al., 2011). The coordinates in CT image-space were converted to coordinates within the patient's segmented MRI brain-space using the [iELVis](#) toolbox (yangWangElecPjct; Yang, Wang, et al., 2012; Groppe et al., 2017). For all sites, the electrode spatial coordinates were transformed from the individual patient space into the standard space of the Montreal Neurological Institute (MNI-152) template for plotting purposes. At NYU, this transformation was performed using the DARTEL algorithm (Ashburner, 2007) implemented in SPM8 (Wellcome Department of Imaging Neuroscience, London, United Kingdom). At Harvard, this transformation was performed using the [iELVis](#) toolkit. At WU the transformation was performed with the [SubNuclear](#) toolbox using img2imgcoord utility.

Each electrode was labeled using the Desikan and Destrieux atlases with the MNE function "mne.get_montage_volume_labels".

Finger Localizer Task

In the Finger Localizer task, participants were presented with four circles, one of which was filled with a specific color, serving as a cue for participants to press the corresponding colored button on the response box. The filled state of the circle persisted for the duration of the response time, followed by an additional delay of 200 milliseconds. The Inter-Trial Intervals (ITIs) were uniformly distributed, with a mean of 0.55 seconds and a range from 0.400 to 0.700 seconds. The experimental protocol comprised 80 trials, distributed equally among the four colors, with 20 trials per color, and the sequence of trials was randomized. This task aimed to identify brain regions responsible for motor control, particularly those governing finger movements, and to pinpoint electrodes selectively activated by specific motor responses, such as button presses.



An illustration depicting a trial in which the participant is required to press the blue button

Please note: Although participants completed this task concurrently with [Experiment 1](#), we did not utilize the data in the analysis, as it was primarily acquired for use in [Experiment 2](#). Consequently, the data pertaining to the Finger Localizer task is not included in this version of our data release.

Task (tiEEG)

Participants proceeded to Experiment 1 either after or before completing the [Finger Localizer task](#). tiEEG consisted of 5 runs containing 4 blocks each, and 34-38 trials per block, 32 non-targets (8 of each category) and 2-6 targets, with each trial lasting 2.4 s approximately, for a total of 720 trials. Rest breaks between runs and blocks were included. Random jitter was added at the end of each trial (mean inter-trial interval of 0.4 s jittered 0.2-2.0 s, truncated exponential distribution) to avoid periodic presentation of the stimuli. Additional information about the task can be found [here](#).

Task	Runs	Blocks	Trials	Total trials
Experiment 1	5	4	34-38 per block	720

Full Structure of Session

Complete standard procedure of an iEEG session is available in [iEEG Standard Operating Procedure](#).

Inclusion Criteria

For the iEEG studies, subjects were 10-65 years old, able to provide informed consent, had IQ > 70, fluent in English, with self-reported normal hearing, normal or corrected-to-normal vision, and cognitive and language abilities within or above the normal range in formal neuropsychological testing performed before surgery. They must not have had an electrographic seizure within 3-hours prior to testing.

Quality Check

A comprehensive quality assessment was conducted on the iEEG data. The data underwent manual annotation by epileptologists, excluding channels within the epileptic onset zone, as well as those exhibiting artifacts or showing complete flatness due to electrode contact issues.



Channel rejection was independently performed by both the data monitoring and iEEG teams, with results compared to ensure consistency. Additionally, electrode reconstruction was verified to align with subjects' CT scans. Finally, we inspected for significant disturbances in the spectra.

Exclusion Criteria

Subjects who were unable to complete a sufficient number of trials due to excessive muscular artifacts, movement, noisy recordings, or a decision by the subject to terminate the experiment were excluded. In addition, data was also excluded if it did not pass any of the pre-defined data quality checks.

Deviations and Missing Data

Attention 1: The iEEG data for the seven subjects listed below do not have the correct sampling rate in the original EDF files. The correct sampling rate for these subjects is 2048 Hz. The raw data has not been updated with the correct sampling rate to avoid alignment issues when constructing the events. However, the same measurements in the BIDS release, provided in the [BrainVision](#) format, have the correct sampling rate as expected.

Subject_ID	CE106	CE107	CE108	CE112	CE115	CE118	CE120
------------	-------	-------	-------	-------	-------	-------	-------

Attention 2: Subject "CG102" is missing the first "Run" and contains only DurR2, DurR3, DurR4, DurR5.

Attention 3: The following subjects are missing "Exit Questionnaires":

Subject_ID	CF103	CF105	CF106	CF110	CF112
	CF113	CF116	CF117	CF119	CF120
	CF121	CF122	CF124	CF125	CF126

Please note: MR and CT data were collected for the subjects at Brigham and Women's Hospital and Children's Hospital Boston. However, due to the data protection policies, they are not included in the COG TATE Data Release.

Curation Procedures

Data Curation Procedure

A detailed explanation about the multiple steps that were taken to prepare the data to be released in public will be available in [Appendix 7](#).

Deviations from Data Curation Procedure

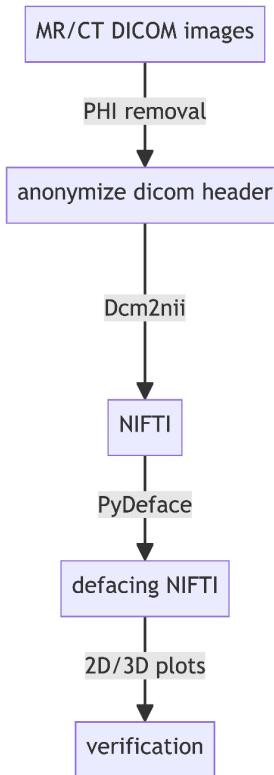
iEEG

Our approach to defacing MR/CT DICOM images involved utilizing the built-in face masking tool from XNAT. However, for 12 iEEG subjects, we encountered difficulties in executing this step. This was primarily due to variations in the DICOM images, which could include cropped images, aliasing artifacts, broken headers, missing slices, non-equidistant slices within a series, and other issues. Below is the list of subjects where the execution of the XNAT facemasking tool failed:

Subject_ID	CF103	CF104	CF112	CF113	CF116	CF117
	CF120	CF121	CF122	CF124	CF125	CF126

To address this issue, we implemented a slightly different workflow that allowed us to successfully deface MR/CT images of these 12 subjects. However, this new approach differed in its ability to regenerate the original DICOM images post-defacement (the original output from the XNAT facemasking tool). Instead, it generated defaced NIFTI images as the primary output. For our current version of data release, we have decided to share only the defaced NIFTI images for these subjects. Details about this workflow are provided below:

1. Anonymization: MR/CT DICOM images underwent anonymization to remove the subject's Protected Health Information (PHI).
2. NIFTI Conversion: Anonymized DICOM images were then converted to the NIFTI image format using the dcm2niix package (version: 1.0.20220505) (Li et al., 2016).
3. Defacing of NIFTI: Defacing of the NIFTI images was performed using the PyDeface package (version: 2.0.2) (Gulban et al., 2022).
4. Verification: This step involved checking the quality of the defaced NIFTI images using 2D/3D image plots to compare before and after the defacing stage.



Alternative workflow for defacing 12 challenging MR/CT DICOM Images

Miscellaneous: In the MR data for subject CF103, one DICOM slice was inadvertently dropped during the conversion process from DICOM to NIFTI format. However, the resulting NIFTI file remains functional and usable.

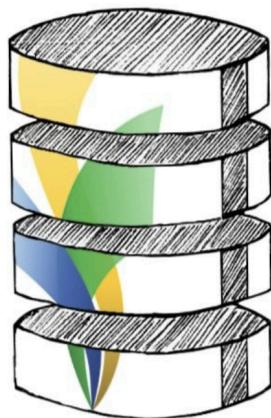
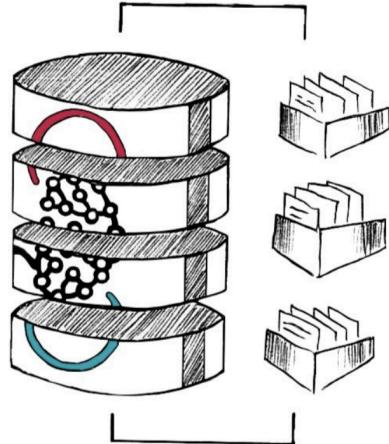
Metadata Curation Procedure

Comprehensive guidance on extracting and managing metadata relevant to the COGITATE project is available in [Appendix 8](#).

Access to COGITATE Data

There are two ways of accessing the COGITATE data:

1. "Live" Database Release: [XNAT](#) (eXtensible Neuroimaging Archive Toolkit)
2. Archival Format: Bundles

[Go to XNAT](#)[Access Data Bundles](#)

1. XNAT

This database offers a web interface for navigating the data and an API (Application Programming Interface) for programmatically retrieving specific databases based on user interests. Comprehensive instructions on how to register, access, and query our database are provided below.

Step 1: Registration

If you are a new user and have not registered yet, you should visit [Cogitate_XNAT_registration](#). Once the registration is done, a verification step, the same as the “Creating an Account”, is needed.

If you have already registered, you can skip this step and login at [Cogitate_XNAT](#).



Welcome to **XNAT**, the data release platform of the **COG TATE** consortium.

On this website you will be able to access and download all the data gathered as part of this collaboration. We provide neural activity using **iEEG**, **fMRI** and **MEG/EEG**, while subjects are performing two tasks aimed at testing two theories of consciousness: **GNWT** and **IIT**.

Stay tuned for more information on how to register and access the data provided here.

1. [Cogitate data release documentation](#)
2. [Cogitate website](#)
3. [Using XNAT](#)

By registering to access data using this service, you agree to the below terms and conditions.

1. [Terms of Use](#)
2. [General Data Protection Regulation](#)

For any questions please contact [Cogitate Support](#).

USER	<input type="text"/>
PASSWORD	<input type="password"/>
Register Forgot login or password?	
<input type="button" value="Login"/>	

Step 2: Navigating at XNAT

After completing the registration step, you can log in with your User and Password. You can see the list of available datasets under the “Projects” tab.



MPI-CURATE currently contains \$proj_count projects, \$sub_count subjects, and \$sd_count imaging sessions.

Projects	Subjects	undefined	undefined	undefined	ECOG	MEG/EEG	EyeTracker																																																
ID: <input type="text"/>	Title: <input type="text"/>	Description: <input type="text"/>																																																					
Keywords: <input type="text"/>	Investigator: <input type="button" value="(SELECT)"/>																																																						
<input type="button" value="Submit"/>																																																							
<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="background-color: #0070C0; color: white;">Projects</th> <th colspan="3" style="background-color: #0070C0; color: white;">Recent Data Activity</th> </tr> </thead> <tbody> <tr> <td>Curated Cogitate MEEG Project EXP1 Project ID: curated_MEEG_EXP1 You are a member for this project.</td> <td>curated_M...</td> <td>MR</td> <td>CA103_MR_0</td> </tr> <tr> <td>COGSAMPLE - EXP1 Project ID: cogsample_exp1 Cogitate sample dataset with only EXP1 data. You are a member for this project.</td> <td>curated_M...</td> <td>MR</td> <td>CB071_MR_0</td> </tr> <tr> <td>Cogsample ALL Project ID: cogsample_all Contains all the subjects and experiments from the sample project. You are a member for this project.</td> <td>curated_E...</td> <td>ECOG</td> <td>CG101_ECOG_1</td> </tr> <tr> <td></td> <td>curated_M...</td> <td>MEEG</td> <td>CB999_MEEG_1</td> </tr> <tr> <td></td> <td>curated_M...</td> <td>MEEG</td> <td>CB015_MEEG_1</td> </tr> <tr> <td></td> <td>curated_E...</td> <td>ECOG</td> <td>CF109_ECOG_1</td> </tr> <tr> <td></td> <td>curated_M...</td> <td>MEEG</td> <td>CB013_MEEG_1</td> </tr> <tr> <td></td> <td>curated_E...</td> <td>ECOG</td> <td>CF107_ECOG_1</td> </tr> <tr> <td></td> <td>curated_M...</td> <td>MEEG</td> <td>CB011_MEEG_1</td> </tr> <tr> <td></td> <td>curated_M...</td> <td>MEEG</td> <td>CB010_MEEG_1</td> </tr> <tr> <td></td> <td>curated_M...</td> <td>MEEG</td> <td>CB008_MEEG_1</td> </tr> </tbody> </table>								Projects	Recent Data Activity			Curated Cogitate MEEG Project EXP1 Project ID: curated_MEEG_EXP1 You are a member for this project.	curated_M...	MR	CA103_MR_0	COGSAMPLE - EXP1 Project ID: cogsample_exp1 Cogitate sample dataset with only EXP1 data. You are a member for this project.	curated_M...	MR	CB071_MR_0	Cogsample ALL Project ID: cogsample_all Contains all the subjects and experiments from the sample project. You are a member for this project.	curated_E...	ECOG	CG101_ECOG_1		curated_M...	MEEG	CB999_MEEG_1		curated_M...	MEEG	CB015_MEEG_1		curated_E...	ECOG	CF109_ECOG_1		curated_M...	MEEG	CB013_MEEG_1		curated_E...	ECOG	CF107_ECOG_1		curated_M...	MEEG	CB011_MEEG_1		curated_M...	MEEG	CB010_MEEG_1		curated_M...	MEEG	CB008_MEEG_1
Projects	Recent Data Activity																																																						
Curated Cogitate MEEG Project EXP1 Project ID: curated_MEEG_EXP1 You are a member for this project.	curated_M...	MR	CA103_MR_0																																																				
COGSAMPLE - EXP1 Project ID: cogsample_exp1 Cogitate sample dataset with only EXP1 data. You are a member for this project.	curated_M...	MR	CB071_MR_0																																																				
Cogsample ALL Project ID: cogsample_all Contains all the subjects and experiments from the sample project. You are a member for this project.	curated_E...	ECOG	CG101_ECOG_1																																																				
	curated_M...	MEEG	CB999_MEEG_1																																																				
	curated_M...	MEEG	CB015_MEEG_1																																																				
	curated_E...	ECOG	CF109_ECOG_1																																																				
	curated_M...	MEEG	CB013_MEEG_1																																																				
	curated_E...	ECOG	CF107_ECOG_1																																																				
	curated_M...	MEEG	CB011_MEEG_1																																																				
	curated_M...	MEEG	CB010_MEEG_1																																																				
	curated_M...	MEEG	CB008_MEEG_1																																																				

Once you click the project's name, you will see the list of subjects in the farthest left column.

Subjects

Add Tab  << first < prev 1 next > last >> 200  1 of 1 Pgs (102 Rows) Reload Options ▾

Subject	M/F	Hand	YOB	MR Sessions	MEEGs	EyeTrackers
CA101	U				1	
CA102	U				1	1
CA103	U			1	1	1
CA104	U				1	1
CA105	U				1	
CA106	U			1	1	1
CA107	U			1	1	1
CA108	U				1	
CA109	U			1	1	1
CA110	U				1	1
CA111	U				1	1
CA112	U			1	1	1

In each subject's folder, the demographic information of that subject and the various sets of data acquired for Experiment 1 are provided. As an example, for a subject with the ID of CA103, the MR session, Eye tracker and MEEG datasets are listed as the below figure.



Custom Field Sets

Additional Demographics

Form UUID:7c721510-adcd-40df-9168-abf203f71005

Additional Demographics

Gender ⓘ	Female	Hand Dominance ⓘ	Right
Year of Birth	-	Handedness level ⓘ	-
Race ⓘ	-	Are you colorblind? ⓘ	No
Eye Dominance ⓘ	Right	Year when subject revoked within consortium	-
Eyedominance Description ⓘ	-	Year when subject revoked open access	-
Primary Language ⓘ	German	Additional comments ⓘ	-

Experiments

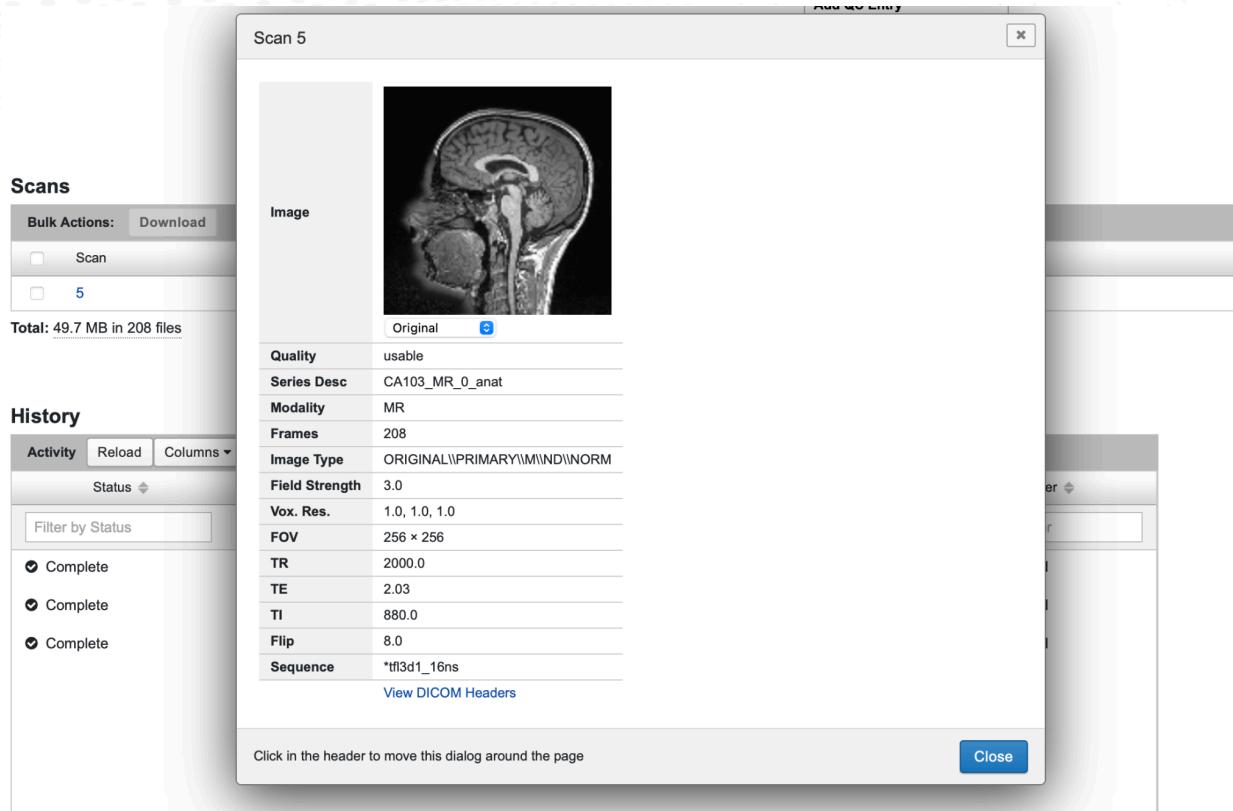
Date	Experiment	Project	Label
1900-Jan-01	MR Session	Curated MEEG EXP1	CA103_MR_0
	EyeTracker	Curated MEEG EXP1	CA103_ET_1
	MEEG	Curated MEEG EXP1	CA103_MEEG_1

In the MR session folder, you can view and access the MR scan of the subject along with the related imaging parameters.

Scans

Bulk Actions: Download					
	Scan	Type	Series Desc	Usability	Files
	<input type="checkbox"/> 5	CA103_MR_0_anat	CA103_MR_0_anat	usable	49.7 MB in 208 files

Total: 49.7 MB in 208 files



In the Eye tracker folder, the eye tracking data of different runs and some details related to them, including the recorded eye, sampling frequency, distance to screen and screen size are available.



Upload details about eyetracking

Scan DurR1

Has the ET data been prepared	
Manufacturer	Quality usable
Number of scans	Manufacturer EYELINK
Sampling Frequency	Recorded Eye Both
Distance to Screen	Sampling Frequency 1000.0
ScreenSize	Distance to Screen -1.0 (cms)
Recorded Eye	Screen Size -1 cms
ET Upload Comments	Click in the header to move this dialog around the page

Close

Scans

Bulk Actions: Download					
<input type="checkbox"/> Scan	Type	Series Desc	Usability	Files	Note
<input type="checkbox"/> DurR1	DurR1	EyeTracker	usable	52.4 MB in 2 files	
<input type="checkbox"/> DurR2	DurR2	EyeTracker	usable	51.7 MB in 2 files	
<input type="checkbox"/> DurR3	DurR3	EyeTracker	usable	50.7 MB in 2 files	

Under the folder of MEEG, there are some tabs on the top where you can find information regarding the Case Report Form, Exit Questionnaire, experiment checklist form, data details within the BIDS framework, and at the bottom, you can download different runs of MEG data.



Behavior Exit Questionnaire EXP1

Cogitate checklist EXP1

BIDS EEG Coordinate System

BIDS Specification EEG

BIDS MEEG Coordinate System

BIDS Specification MEEG

Upload Form MEEG

Form UUID:7c185942-0d1b-4a29-8b12-f2b9d4f7707e

Upload form for MEEG files.

Has the MEEG data been prepared for upload? -**Number of scans**

-

Sampling Frequency

-

Powerline Frequency

-

Highpass Filter Cutoff Frequency

-

Digitized Landmarks

-

Digitized Head Points

-

Lowpass Filter Cutoff Frequency

-

Dewar Position

-

Recording Duration

-

Number of channels

-

MEG Channel Count

-

Highpass Software Filter

-

Lowpass Software Filter

-

MEG Coordinate System

-

MEEG Upload Comments

-

Scans

Bulk Actions:		Download			
<input type="checkbox"/>	Scan	Type	Series Desc	Usability	Files
<input type="checkbox"/>	DurR1	DurR1	FIF	usable	1.1 GB in 1 files
<input type="checkbox"/>	DurR2	DurR2	FIF	usable	1.1 GB in 1 files
<input type="checkbox"/>	DurR3	DurR3	FIF	usable	1.0 GB in 1 files
<input type="checkbox"/>	DurR4	DurR4	FIF	usable	1.0 GB in 1 files
<input type="checkbox"/>	DurR5	DurR5	FIF	usable	1.0 GB in 1 files
<input type="checkbox"/>	RestinEO	RestinEO	FIF	usable	428.8 MB in 1 files
<input type="checkbox"/>	Rnoise	Rnoise	FIF	usable	251.2 MB in 1 files

Total: 6.0 GB in 7 files

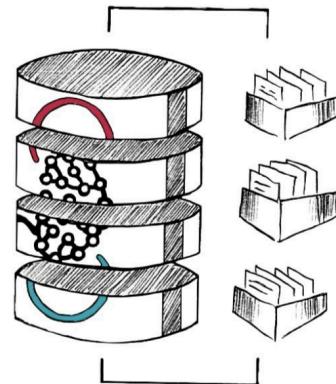
Naming Convention on XNAT

The data on XNAT is organized into subjects and sessions under a given project. The subjects are identified using the format “CX???” and the sessions follow the format CX??_MODALITY_VISIT_PARADIGMRUN e.g. CA103_MEEG_1_DurR1 indicated MEEG measurement for subject ID CA103 during the first visit with Dur experimental paradigm run 1 (R1).

2. Bundles

This approach involves providing a collection of links to the prepared bundles of the data and accompanying metadata, which are available in zip format. These links grant users the ability to download specific modalities, example datasets, or the complete dataset.

Data bundles



Experiment 1 - Sample datasets

Data Bundle	Format	Size (GB)	Download time	Download
All fMRI, MEG, iEEG, Behav, ET	Raw	212	00:30:21	Download
fMRI + Behav + ET	Raw	110	00:30:21	Download
MEG + Behav + ET	Raw	71	00:10:09	Download
iEEG + Behav + ET	Raw	31	00:04:26	Download

Experiment 1 - *Full datasets*

Data Bundle	Format	Size (GB)	Download time	Download
All fMRI, MEG, iEEG, Behav, ET	Raw	212	00:30:21	Download
fMRI + Behav + ET	Raw	110	00:30:21	Download
MEG + Behav + ET	Raw	71	00:10:09	Download
iEEG + Behav + ET	Raw	31	00:04:26	Download
Behav + ET	Raw	15	00:02:08	Download

Here is a brief explanation about how to access the data bundles:



Step 1: Create a Data User Account

Access to the data bundles requires a quick and easy registration process.

- a. Provide user information, including name and email address.
- b. Read and accept the [Terms of Use](#) and [GDPR Requirements](#) (General Data Protection Regulation).
- c. Once you register, you will receive four (4) emails to the email account you registered with. In some cases, checking your junk mail may be necessary.
 - i. **Welcome email:** general information
 - ii. **Data User Account Verification email:** Within the verification email, you must click on the 'verify my account' option to finalize step 1 of creating a data user account in order to gain access to all current and future data releases.
 - iii. **Resource Material:** A handy email that contains all the important links that serve as reference materials to Cogitate data.
 - iv. **Mailing List Subscription:** In order to stay up-to-date and informed about news related to COGITATE data releases, you must activate your email subscription (this is in compliance with GDPR requirements).

Tip: The registration procedure needed for accessing the data bundles is a separate step than what is required to access XNAT.



Step 2: Login and logout of your Data User account

- a. To login to your account, go to the Login button on the top right of the page. Enter your email and password used when registering. You should now have access to the Cogitate Data User main page and Data Bundles.
- b. To log out of your account, go to the top navigation bar and hover over Data. In the dropdown menu, click on Data User Account. A panel will open on the right side of the screen - click on Account Settings in the bottom of that panel. Then the option to Sign out will appear under your username. Click on Sign out.

Tip: The Login button will remain as 'Login' even after signing in to your account. The only way of knowing whether you are logged in or out, is by clicking on Data User Account, under the Data heading or being able to download data (i.e. indicating you are in fact, logged in)



Step 3: How To Download the Data

- a. Login to your account
- b. Scroll down and click on the "Access Data Bundles"
- c. Click on the download button next to each dataset

Naming Convention for Bundles

Raw data bundles follow the below naming convention. The project root directory consists of subdirectories named after the subject's ID which is of the format "CX??". The subject directories consist of various sub directories as described below. Except for the metadata directory the sessions follow the pattern subject-ID_PARADIGM_MODALITY. If the modality data is paradigm agnostic, e.g. MR, CT then the paradigm is left blank.

We currently have two paradigms in the data EXP1 indicating the experiment described and FingerLoc for the finger localiser. The session directories further contains individual scans following the format CX??_MODALITY_1_DurR1.EDF.

The metadata subdirectory further consists of various assessments and questionnaires that provide valuable information.

Experiment 1: Directory Structure of Data Bundles

Raw Data

Raw data files are organized hierarchically: Experiment modality --> Subjects --> data folders. The metadata related to each level of the hierarchy is contained in a mandatory folder called 'metadata'.

Each data folder follows a naming convention {subject_context_modality}\[_modifier]}. The section of the names are separated by underscores. This naming convention aims at making it easy to identify the data files that relate to the same moment in time and that were acquired simultaneously.

- subject -> this refers to the subject ID
- context -> the task or context. This section is optional and can be empty, e.g. if a subject had a standalone MR scan the context is left blank, resulting in a double underscore like in the case of the CT scan or MR scan in the above example
- modality -> or type of data collected

The Cogitate consortium collected several types of data/metadata during the experiments:

- BEH: behavioral events
- ET: Eye tracking data
- MR: Magnetic resonance data (anatomical scans)
- EXQU: Exit Questionnaire
- CRF: Case Report Form

All metadata related to the subject can be found under the aptly named 'metadata' folder under the subject folder (this refers mainly to the EXQU and CRF files).

The remaining metadata for the experiment as well as the demographic information on the subjects can be found in the metadata folder above the subject.

This folder includes experiment wide metadata in json format and a csv table with the demographic data of all subjects:

- devices: A list of devices used to collect the data
- protocols: a link to the Standard Operating Procedure (SOP) document used for the data collection
- subjects_demographics: the full set of subjects and their metadata for the specific experiment modality
- tasks_**taskname**: a description of the behavioral task or context with which we named the data bundles.
- wirings: a pdf file showing how the **devices** are connected to each other depicting the experimental setup.

Raw M-EEG Data Directory Structure

```

COG_MEEG_EXP1_RELEASE/
    └── metadata/
        ├── devices_MEEG.json
        ├── protocols_MEEG.json
        ├── subjects_demographics_MEEG.json
        ├── tasks_EXP1.json
        ├── tasks_RestinEO.json
        ├── tasks_Rnoise.json
        └── wirings_MEEG.pdf
    └── CB036
        ├── metadata/
        │   ├── CB036_EXP1_CRF.json
        │   └── CB036_EXP1_EXQU.json
        ├── CB036_EXP1_BEH/
        ├── CB036_EXP1_LPTTriggers/
        ├── CB036_EXP1_MEEG/
        ├── CB036_EXP1_ET/
        ├── CB036_RestinEO_MEEG/
        ├── CB036_RestinEO_ET/
        ├── CB036_Rnoise_MEEG/
        └── CB036_MR/
# Experiment modality level metadata folder
# List of devices used to collect the data
# A link to the Standard Operating Procedures (SOP)
# Demographic information of MEEG subjects
# Description of the 1st Cogitate task
# Description of the Resting state task
# Description of the Rnoise task
# Wiring diagram of devices_MEEG.json connections
# Subject folder
# Subject level metadata folder
# Subject Case Report Form (CRF)
# Subject Exit Questionnaire responses
# Behavioral Events data collected during EXP1
# Trigger data for synchronization
# MEEG data collected during EXP1 (fif)
# Eye Tracking data collected during EXP1 (asc)
# MEEG data collected during RestinEO task (fif)
# Eye Tracking data collected during RestinEO task
# MEEG data collected during Rnoise task (fif)
# MR anatomical scan data (fif)

```

Raw iEEG Data Directory Structure

```

COG_ECOG_EXP1_RELEASE/
    └── metadata/
        ├── devices_ECOG.json
        └── protocols_ECOG.json
    (SOP) document used for the data collection
    └── subjects_demographics_ECOG.json
respective demographic information
    ├── tasks_EXP1.json
    ├── tasks_FingerLoc.json
    └── wirings_ECOG.pdf
devices_ECOG.json are connected to each other
    └── CE103
        ├── metadata/
        │   ├── CE103_EXP1_CRF.json
        │   └── CE103_EXP1_EXQU.json
        ├── CE103_EXP1_BEH/
        ├── CE103_EXP1_ECOG/
        ├── CE103_EXP1_ET/
        ├── CE103_FingerLoc_ECOG/
        └── CE103_FingerLoc_BEH/
Finger Localizer task
    ├── CE103_CT/
    ├── CE103_MR/
    └── CE103_ElecCoords/
coregistration end electrode reconstruction pipeline
# Experiment modality top level folder
# Experiment modality level metadata folder
# List of devices used to collect the data
# A link to the Standard Operating Procedure
# Full set of experiment modality subjects with their
# Description of the 1st Cogitate task
# Description of the Finger Localizer task
# Wiring pdf file showing how the devices described in
# Subject folder
# Subject level metadata folder
# Subject Case Report Form (CRF)
# Subject Exit Questionnaire responses
# Behavioral Events data collected during EXP1
# ECOG data files collected during EXP1
# Eye Tracking data collected during EXP1
# Ecog data collected during the Finger Localizer task
# Behavioral event data collected during the
# CT scan data (no task)
# MR anatomical data
# Contains coordinate output files of MR/CT

```

BIDS Format

The BIDS (Brain Imaging Data Structure) file structure for M-EEG (Magnetoencephalography) and iEEG (intracranial EEG) data adheres to a standardized format for organizing neuroimaging data. Each file follows a structured naming convention indicating subject, session, task, and



data type. Here is a breakdown of the key elements within each modality's data directory structure:

- dataset_description.json: Provides general information about the dataset.
 - participants.json and participants.tsv: Contain demographic information about subjects.
 - README.md: Offers an overview of the data and BIDS format.
 - Subject-specific data: Organized under sub-[SubjectID]/.
 - Session-specific data: Organized under ses-[SessionID]/.
 - Anatomical and functional data: Stored in appropriate folders (anat/ for anatomical, meg/ for MEG, and ieeeg/ for iEEG).
 - Metadata: Related to subjects and experiments is stored in metadata/ directories.

This structured approach ensures clarity and consistency in data organization, facilitating ease of access and analysis for researchers and collaborators.

BIDS M-EEG Data Directory Structure

```

COG_MEEG_EXP1_BIDS_RELEASE/
|-- dataset_description.json
    # General information about BIDS version,
    # and the link of COGITATE website
|-- derivatives
    # Including metadata and coreg
    (coregistration)
        |-- additional_metadata
            |-- dataset_description.json
                # Containing all of the metadata
                # General information about BIDS version,
                # and the link of COGITATE website
        type of dataset
            |-- METADATA
                # Including metadata including the list of
                # devices, link to COGITATE GitHub repository, types of tasks, stimuli and responses and wiring diagram of
                # MEG data
            |-- analysis.json
                # Analysis steps, the order of them and
                # the link of analysis code repository
            |-- devices_MEEG.json
                # List of devices used for MEG data
        acquisition
            |-- protocols.json
                # Link of COGITATE wiki and MEG SOP
            |-- tasks_EXP1_MEEG.json
                # Description of behavioral task, stimuli
                # and responses
            |-- tasks_RestinEO_MEEG.json
                # Description of resting-state task and
                # acquisition
        type of the response
            |-- tasks_Rnoise_MEEG.json
                # Description of empty room task
            `-- wiring_MEEG.pdf
                # Wiring diagram of MEG
            |-- README.md
                # Containing an explanation about
                # acquisition
        additional_metadata directory
            |-- sub-CA103
                # Subject folder
            `-- METADATA
                # Containing Case Report Form, Exit
                # Questionnaire and subject's demography
                    |-- CA103_CRF.json
                        # Case Report Form
                    |-- CA103_demographics.json
                        # Subject's demography
                    `-- CA103_EXQU.json
                        # Exit Questionnaire
                `-- coreg
                    # The results of the coregistration
                    # BIDS version, Data Type, and description
                    # of the files of this directory
                    |-- README.md
                    |-- sub-CA103
                        # Subject folder
                        |-- ses-1
                            # Session 1/visit 1
                            |-- meg
                                # MEG folder
                                `-- sub-CA103_ses-1_trans.fif
                                    # The output of coregistering MEG sensors

```

```

and head to the anatomical data
|-- participants.json
demography
|-- participants.tsv
|-- README.md
|-- sub-CA103
`-- ses-1
    |-- anat
    |   |-- sub-CA103_ses-1_T1w.json
    |   `-- sub-CA103_ses-1_T1w.nii.gz
    |-- meg
    |   |-- sub-CA103_ses-1_acq-calibration_meg.dat
    |   |-- sub-CA103_ses-1_acq-crosstalk_meg.fif
    |   |-- sub-CA103_ses-1_coordsystem.json
coil and coordinate system, units, description and anatomical
names, types, units, sampling rate, status, and frequency cutoffs of the filter applied to the recorded
data during run 1
|   |-- sub-CA103_ses-1_task-dur_run-01_channels.tsv # Contains information on the channel
names, types, units, sampling rate, status, and frequency cutoffs of the filter applied to the recorded
data during run 1
|   |   |-- sub-CA103_ses-1_task-dur_run-01_events.json # Description of sample, value and trial
type
|   |   |-- sub-CA103_ses-1_task-dur_run-01_events.tsv # Contains information about the
events/stimuli presented during Experiment 1, run 1, event's onset time and duration, type of event,
event code (trigger code) and sample
|   |   |-- sub-CA103_ses-1_task-dur_run-01_meg.fif # Contains the raw/unprocessed MEG data
during the task of Experiment 1/session 1, run 1
|   |   |-- sub-CA103_ses-1_task-dur_run-01_meg.json # Contains power line and sampling
frequencies, duration of recording, MEG, EOG and ECG and trigger channel counts during run 1
|   |   |-- sub-CA103_ses-1_task-dur_run-02_channels.tsv
|   |   |-- sub-CA103_ses-1_task-dur_run-02_events.json
|   |   |-- sub-CA103_ses-1_task-dur_run-02_events.tsv
|   |   |-- sub-CA103_ses-1_task-dur_run-02_meg.fif
|   |   |-- sub-CA103_ses-1_task-dur_run-02_meg.json
|   |   |-- sub-CA103_ses-1_task-dur_run-03_channels.tsv
|   |   |-- sub-CA103_ses-1_task-dur_run-03_events.json
|   |   |-- sub-CA103_ses-1_task-dur_run-03_events.tsv
|   |   |-- sub-CA103_ses-1_task-dur_run-03_meg.fif
|   |   |-- sub-CA103_ses-1_task-dur_run-03_meg.json
|   |   |-- sub-CA103_ses-1_task-dur_run-04_channels.tsv
|   |   |-- sub-CA103_ses-1_task-dur_run-04_events.json
|   |   |-- sub-CA103_ses-1_task-dur_run-04_events.tsv
|   |   |-- sub-CA103_ses-1_task-dur_run-04_meg.fif
|   |   |-- sub-CA103_ses-1_task-dur_run-04_meg.json
|   |   |-- sub-CA103_ses-1_task-dur_run-05_channels.tsv
|   |   |-- sub-CA103_ses-1_task-dur_run-05_events.json
|   |   |-- sub-CA103_ses-1_task-dur_run-05_events.tsv
|   |   |-- sub-CA103_ses-1_task-dur_run-05_meg.fif
|   |   |-- sub-CA103_ses-1_task-dur_run-05_meg.json
|   |   |-- sub-CA103_ses-1_task-noise_channels.tsv # Contains information on the channel
names, types, units, sampling rate, status, and frequency cutoffs of the filter applied to the recorded
data during noise recording
|   |   |-- sub-CA103_ses-1_task-noise_meg.fif # Contains the raw/unprocessed MEG data
during noise recording of Experiment 1/session 1
|   |   |-- sub-CA103_ses-1_task-noise_meg.json # Contains power line and sampling
frequencies, duration of recording, MEG, EOG and ECG and trigger channel counts during noise recording
|   |   |-- sub-CA103_ses-1_task-rest_channels.tsv # Contains information on the channel
names, types, units, sampling rate, status, and frequency cutoffs of the filter applied to the recorded
data during resting-state recording
|   |   |-- sub-CA103_ses-1_task-rest_meg.fif # Contains the raw/unprocessed MEG data
during resting-state recording of Experiment 1/session 1

```

```

|     |   `-- sub-CA103_ses-1_task-rest_meg.json      # Contains power line and sampling
frequencies, duration of recording, MEG, EOG and ECG and trigger channel counts during resting-state
recording
|     `-- sub-CA103_ses-1_scans.tsv                  # List of MEG data files

```

BIDS iEEG Data Directory Structure

```

COG_ECOG_EXP1_BIDS_RELEASE/
|-- dataset_description.json                                # General information about BIDS
version, type of dataset, Authors, Acknowledgments, Funding, Ethics Approvals, and the link of COGITATE
website
|-- derivatives                                            # Directory containing derived
data
|   |-- additional_metadata                                # Containing all of the metadata
and CT scans
|   |   |-- dataset_description.json                      # General information about BIDS
version, type of dataset and list of metadata
|   |   |-- METADATA                                      # Metadata folder
|   |   |   |-- analysis.json                            # Analysis steps, the order of
them and the link of analysis code repository
|   |   |   |-- devices_ECOG.json                         # List of devices used for iEEG
data acquisition
|   |   |   |-- manifest.json                           # List of all MR and CT datasets
|   |   |   |-- protocols.json                          # Link of COGITATE wiki
|   |   |   |-- tasks_EXP1_ECOG.json                   # Description of behavioral
task, stimuli and responses
|   |   |   |-- tasks_FingerLoc_ECOG.json            # Description of Finger
Localizer task, stimuli and responses
|   |   |   `-- wiring_ECOG.pdf                        # Wiring diagram of iEEG
|   |   |-- README.md                                    # Containing an explanation
about additional_metadata directory
|   |   |-- sub-CF102                                     # Subject folder
|   |   |   '-- METADATA                                # Metadata folder
|   |   |       |-- CF102_CRF.json                     # Case Report Form
|   |   |       |-- CF102_demographics.json           # Subject's demography
|   |   |       `-- CF102_EXQU.json                    # Exit Questionnaire
|   |   '-- ct                                         # CT scans folder
|   |       |-- dataset_description.json              # General information about BIDS
version and format of CT scans
|   |       |-- README.md                            # Containing an explanation
about ct folder
|   |       |-- sub-CF102                            # Subject folder
|   |       |   '-- ses-1                           # Session 1/visit 1
|   |       |       '-- anat                         # Folder of CT data
|   |       |           '-- sub-CF102_ses-1_ct.nii.gz # CT data
|   |-- participants_epilepsy.json                   # Contains descriptions of
additional metadata including electrode implantation, seizure classification, IQ scores, auditory
capabilities, seizure frequency, surgical history, MRI findings, and family medical history
|-- participants_epilepsy.tsv                      # Contains additional metadata
including electrode implantation, seizure classification, IQ scores, auditory capabilities, seizure
frequency, surgical history, MRI findings, and family medical history
|-- participants.json                               # Contains descriptions of
metadata including biological sex, age, handedness, weight, height, primary and secondary languages,
ethnicity, education level, colorblindness status, visual correction method, eye dominance, eye chart test
results, and strength of visual correction in diopters
|-- participants.tsv                             # Contains metadata including
biological sex, age, handedness, weight, height, primary and secondary languages, ethnicity, education
level, colorblindness status, visual correction method, eye dominance, eye chart test results, and
strength of visual correction in diopters.

```

```

|-- README.md                                     # Overview of iEEG data, the
BIDS format and description of the contents of the dataset
|-- sub-CF102
|   '-- ses-1
|       |-- anat
|           '-- sub-CF102_ses-1_T1w.nii.gz
|       '-- ieeg
|           '-- sub-CF102_ses-1_atlas-desikan_labels.json
|_ses-1_atlas-desikan_labels.tsv file
|       '-- sub-CF102_ses-1_atlas-desikan_labels.tsv      # Subject folder
obtained from the Desikan atlas using mne.get_montage_volume_labels, along with the electrode coordinates in the subject's native T1 space
|       '-- sub-CF102_ses-1_atlas-destrieux_labels.json    # Session 1/visit 1
|_ses-1_atlas-destrieux_labels.tsv file
|       '-- sub-CF102_ses-1_atlas-destrieux_labels.tsv      # Folder of MR anatomical data
obtained from the Destrieux atlas using mne.get_montage_volume_labels, along with the electrode coordinates in the subject's native T1 space
|       '-- sub-CF102_ses-1_laplace_mapping_ieeg.json        # MR anatomical data
the electrodes
|       '-- sub-CF102_ses-1_space-fsaverage_coordsystem.json # Folder of iEEG data
of the channels
|       '-- sub-CF102_ses-1_space-fsaverage_electrodes.tsv    # Contains descriptions about
coordinates in fsaverage space
|       '-- sub-CF102_ses-1_task-Dur_channels.tsv            # Contains electrode labels
channel types (which channel is seeg, ecog...)
|       '-- sub-CF102_ses-1_task-Dur_events.json             # Contains referencing scheme of
|       '-- sub-CF102_ses-1_task-Dur_events.tsv               # Contains the x y z coordinates
timestamps, duration and descriptions aligned in the edf timeline. This file was generated by parsing the photodiode trigger and combining the photodiode time stamps with the log files information
|       '-- sub-CF102_ses-1_task-Dur_ieeg.eeg                 # Contains electrodes'
|       '-- sub-CF102_ses-1_task-Dur_ieeg.json                # Contains information about the
recording parameters and further meta-information
|       '-- sub-CF102_ses-1_task-Dur_ieeg.vhdr                # Contains the events
description of the event collected during the data recording
|       '-- sub-CF102_ses-1_scans.tsv                         # Raw EEG data file
# Header file - Contains
# Marker file - Contains
# Time of acquisition

```

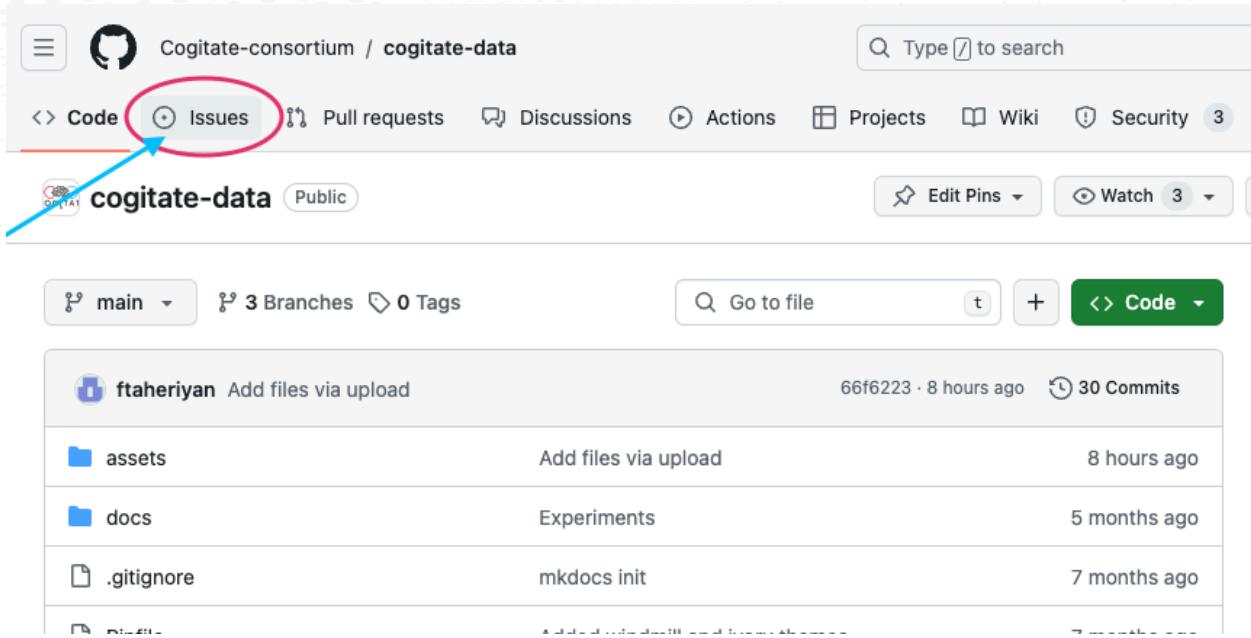
Support and Report Bugs

There are various ways the COG TATE team can support you as a Data User, depending on the type of help you are seeking.

For general questions, email us at: cogitate-support@ae.mpg.de

For reporting issues or bugs:

1. Visit github.com/cogitate-consortium/cogitate-data/issues and ensure you are logged into your GitHub account
 - a. Don't have a GitHub account? Create one [here](#).
2. Click on **Issues**



Cogitate-consortium / cogitate-data

Type ⌘ to search

< Code Issues Pull requests Discussions Actions Projects Wiki Security 3

cogitate-data Public Edit Pins Watch 3

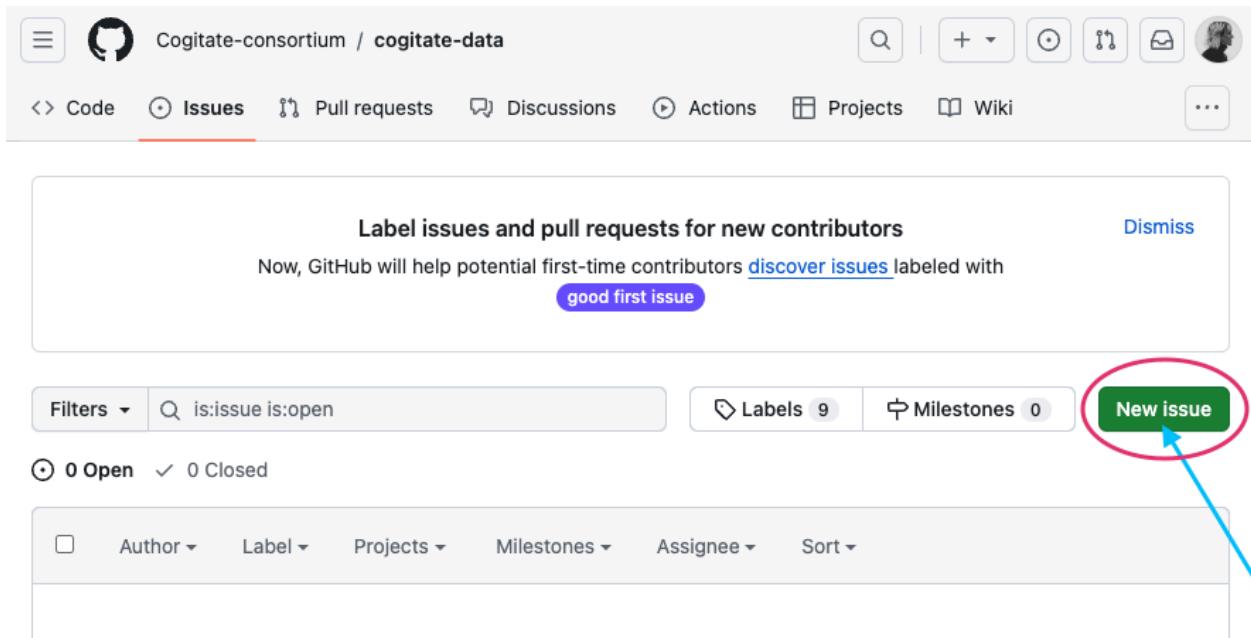
main 3 Branches 0 Tags Go to file + Code

ftaheriyani Add files via upload 66f6223 · 8 hours ago 30 Commits

- assets Add files via upload 8 hours ago
- docs Experiments 5 months ago
- .gitignore mkdocs init 7 months ago

3. Create New Issue

- provide maximal details possible



Cogitate-consortium / cogitate-data

Issues

Label issues and pull requests for new contributors Dismiss

Now, GitHub will help potential first-time contributors [discover issues](#) labeled with [good first issue](#)

Filters is:issue is:open Labels 9 Milestones 0 New issue

0 Open 0 Closed

Author Label Projects Milestones Assignee Sort

4. Click Submit new issue.

Cogitate-consortium / cogitate-data

Code Issues Pull requests Discussions Actions Projects Wiki Security



Add a title

Title

Assignees

No one—[assign yourself](#)

Add a description

[Write](#)

[Preview](#)

[H](#)

[B](#)

[I](#)

[≡](#)

[↔](#)

[🔗](#)

[...](#)

Add your description here...

provide maximal detail to help ensure your issue is properly reported and resolved in a timely manner - thank you :)

Markdown is supported

Paste, drop, or click to add files

[Submit new issue](#)

Labels

None yet

Projects

None yet

Milestone

No milestone

Development

Shows branches and pull requests link to this issue.

Helpful resources

[GitHub Community Guidelines](#)

5. Your issue will be logged with our Data Release team and dealt with in a timely manner.

Links and Reference Materials

COGITATE Website	https://www.arc-cogitate.com/
COGITATE Main Scientific Paper	https://doi.org/10.1371/journal.pone.0268577
COGITATE - Preregistration v4 - December 2022	https://osf.io/gm3vd
GitHub Repository	https://github.com/Cogitate-consortium
COGITATE Wiki	https://cogitate-consortium.github.io/cogitate-data/
Subjects Demography	https://s.gwdg.de/DGPX0l
YouTube Demos	COGITATE Experiment 1 HOW TO create a COGITATE Data User account HOW TO login and logout of your COGITATE Data User account HOW TO Download a COGITATE Data Bundle
XNAT Support	https://wiki.xnat.org/documentation/

Modality	Place of Acquisition	Number of Subjects	Reference Materials
M-EEG	University of Birmingham, Center for Human Brain Health (CHBH)	102	M-EEG Wiring Diagram
	Peking University (PKU)		M-EEG Experiment 1 Code M-EEG SOP
fMRI	Donders Center for Cognitive Neuroimaging (DCCN)	122	
	Yale Magnetic Resonance Research Center (MRRC)		
ECoG (iEEG)	Harvard University at Boston Children's Hospital	38	iEEG Wiring Diagram
	New York University Langone (NYU)		iEEG Experiment 1 Code
	University of Wisconsin		iEEG SOP



Appendices

[Appendix 1. Screening Questionnaire](#)

[Appendix 2. Case Report Form](#)

[Appendix 3. Exit Questionnaire](#)

[Appendix 4. MEG Standard Operating Procedure](#)

[Appendix 5. GDPR Requirements](#)

[Appendix 6. Terms of Use](#)

Appendix 7. Data Curation Standard Operating Procedure

[Appendix 8. Metadata Curation Standard Operating Procedure](#)

[Appendix 9. iEEG Standard Operating Procedure](#)

Acknowledgements

This project and data are made possible through the support of a grant from [Templeton World Charity Foundation](#). The opinions expressed throughout our documentation and data are those of the authors and do not necessarily reflect the views of Templeton World Charity Foundation.



Members of the Cogitate Consortium contributed an incredible amount of time, effort and expertise to design the experiments, collect the data and subsequently curate the data and code for public release. We would like to acknowledge the significant contributions of the following team members:

Experimental design:

Yoav Roll, Alex Lepauvre, Aya Khalaf, and Katarina Bentz, Felix Bernoulli

Konstantinos Vasileiadis, Aris Semertzidis, Nikos Gregos

Data Collection:

fMRI

- Donders (Floris de Lange, David Richter, Sarah Kusch, Monique Smulders)
- Yale (Hal Blumenfeld, Aya Khalaf, Abdel Sharaf)

M-EEG

- Birmingham (Oscar Ferrante, Ole Jensen, Dorottya Hetenyi, Tara Ghafari)
- PKU (Ling Liu, Huan Luo, Shujun Yang)

iEEG

- NYU (Sasha Devore, Simon Henin, Stephanie Montenegro, Jay Jeschke, Alia Seedat)
- Harvard (Gabriel Krieman, Katarina Bendtz, David Mazumder, Marcelo Armendariz)
- WU (Urszula Gorska)

Data Quality Checks:

- Rony Hirschhorn (Tel Aviv University)
- Urszula Gorska (WU)
- Csaba Kozma (WU)
- Liad Mudrik (Tel Aviv University)
- Michael Pitts (Reed College)

Data Curation:

COG TATE's Data Release Team: Niccolò Bonacchi, Tanya Brown, Kyle Kahraman, Lucia Melloni, Praveen Sripad, Fatemeh Taheriyan, Diptyajit Das

Flywheel Cloud Services (XNAT): Mohana Ramaratnam, James Dickson, Angela Farrar, Dan Marcus, Blake Griggs



Appendix 1. Screening Form

M-EEG Screening Form

This questionnaire should be filled by the participant before the experiment. It asks for some information such as name, weight (kg), email, data of birth and phone number of the subject following by a couple of questions including:

- Have you participated in a MEG study before?
- Do you suffer from any medical condition that may be relevant (e.g. epilepsy, diabetes, asthma)?
- Do you suffer from claustrophobia?
- Have you been formally diagnosed with attention deficit (hyperactivity) disorder (AD(H)D)?
- Have you been formally diagnosed with autism spectrum disorder (ASD)?
- Other information (e.g. spectacle prescription)

Appendix 2. Case Report Form

This form was for reporting any issues that might have happened during the experiment. After the end of the experiment and saving the data, the operator filled out this form.

M-EEG Case Report Form

For M-EEG modality, the below items were asked for different sections of the data acquisition (Data for Empty Room, Resting-state, run 1, ..., run 5) in the Case Report Form:

- Eye tracking not working
- Eye tracking showing bad patterns
- Problems in task performance
- Strong head motion
- Strong body motion
- Trigger monitoring
- Bad MEG sensors
- Bad EEG sensors
- Notes: explaining about the issue in a more detail

iEEG Case Report Form

The iEEG case report form contained essential details such as the subject's ID, task name, and the distance from the stimulus interface to the subject in centimeters. It also included comprehensive information about the eye-tracking setup, specifying which eye was tracked, the

sampling rate, calibration method, pupil shape, and camera lens used. Additionally, trigger information, including the trigger channels utilized and other relevant parameters, was documented. Furthermore, some fields were provided for notes and miscellaneous information. Here are the specific fields listed in the iEEG case report form:

- Subject ID
- Task Name
- Stimulus Interface Computer Name
- Distance (cm)
- Practice Run
- Pre-Task Equipment Disinfected
- Post-Task Equipment Disinfected
- Other
- Notes
- Screen Brightness
- Eye Tracking Sticker
- Glasses or Lenses
- Trigger Information:
 - Trigger Channel
 - Trigger Reference Channel
 - Trigger Channel Gain
 - Audio Channels
- Eye Tracker Setup:
 - Eye Tracked
 - Sampling Rate
 - Calibration Type
 - Pupil
 - Camera Lens

Appendix 3. Exit Questionnaire

This Form should be filled by the participant after the experiment and asks the below questions:

- How difficult was it to stay focused for the entire duration of the experiment?
- Did you notice that the stimuli duration varied?
- When stimuli were presented for a short period of time, were you able to focus on them for as long as they were presented?
- When stimuli were presented for a medium period of time, were you able to focus on them for as long as they were presented?
- When stimuli were presented for a long period of time, were you able to focus on them for as long as they were presented?
- Was it difficult to keep the targets in memory for the entirety of a block?
- For each stimulus category, how hard was it to recognize them among the other stimuli? [Faces]

- For each stimulus category, how hard was it to recognize them among the other stimuli? [Object]
- For each stimulus category, how hard was it to recognize them among the other stimuli? [Letters]
- For each stimulus category, how hard was it to recognize them among the other stimuli? [Symbols]
- Were specific stimuli orientations harder to recognize than others?
- If yes, which ones?
- What did you think of the block length?
- Would the task have been easier if the stimulus duration didn't vary so much?
- Do you have any additional comments to share?

Glossary

Term	Definition
Wiring Diagram	A diagram that shows how different devices or sensors were connected
Standard Operating Procedure (SOP)	A document that outlines the steps or procedures to be followed for each modality
GDPR (General Data Protection Regulation)	A comprehensive data protection and privacy regulation in the European Union (EU) and the European Economic Area (EEA) - It offers instructions on leveraging the data of the users who register to download COGITATE data
HIPAA (Health Insurance Portability & Accountability Act)	Standards for the protection of sensitive patient health information - HIPAA is used to identify and remove personal identifiers in the curation process
BIDS (Brain Imaging Data Structure)	A standard format for organizing and describing neuroimaging data - This is one of the formats of the released COGITATE datasets
XNAT (Extensible Neuroimaging Archive Toolkit)	An open-source software platform designed for managing, sharing, and analyzing neuroimaging and related data in research settings - This is the platform used for COGITATE project to store the data and facilitate accessing it for the users

References

1. <https://bids-specification.readthedocs.io/en/stable/>
2. Melloni, L., Mudrik, L., Pitts, M., Bendtz, K., Ferrante, O., Gorska, U., ... & Tononi, G. (2023). An adversarial collaboration protocol for testing contrasting predictions of global neuronal workspace and integrated information theory. *PLoS One*, 18(2), e0268577. <https://doi.org/10.1371/journal.pone.0268577>
3. [COG TATE - Preregistration v4 - December 2022](#)
4. [GDPR](#) (General Data Protection Regulation)
5. [HIPAA](#) (Health Insurance Portability & Accountability Act)
6. Tarr, M. J. (1996). The Object Databank. Carnegie Mellon University, [Tarr Lab page](#), [Databank direct link](#).
7. Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: the SHINE toolbox. *Behavior Research Methods*, 42(3), 671-684. <https://doi.org/10.3758/BRM.42.3.671>
8. Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial vision*, 10(4), 437-442. <https://doi.org/10.1163/156856897x00366>
9. Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*, 9(2), 179-194. <https://doi.org/10.1006/nimg.1998.0395>
10. Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical image analysis*, 5(2), 143-156. [https://doi.org/10.1016/S1361-8415\(01\)00036-6](https://doi.org/10.1016/S1361-8415(01)00036-6)
11. Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., ... & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, 23, S208-S219. <https://doi.org/10.1016/j.neuroimage.2004.07.051>
12. Rohr, K., Stiehl, H. S., Sprengel, R., Buzug, T. M., Weese, J., & Kuhn, M. H. (2001). Landmark-based elastic registration using approximating thin-plate splines. *IEEE Transactions on medical imaging*, 20(6), 526-534. [DOI: 10.1109/42.929618](#)
13. Joshi, A., Scheinost, D., Okuda, H., Belhachemi, D., Murphy, I., Staib, L. H., & Papademetris, X. (2011). Unified framework for development, deployment and robust testing of neuroimaging algorithms. *Neuroinformatics*, 9, 69-84. <https://doi.org/10.1007/s12021-010-9092-8>
14. Yang, A. I., Wang, X., Doyle, W. K., Halgren, E., Carlson, C., Belcher, T. L., ... & Thesen, T. (2012). Localization of dense intracranial electrode arrays using magnetic resonance imaging. *Neuroimage*, 63(1), 157-165. <https://doi.org/10.1016/j.neuroimage.2012.06.039>
15. Groppe, D. M., Bickel, S., Dykstra, A. R., Wang, X., Mégevand, P., Mercier, M. R., ... & Honey, C. J. (2017). iELVis: An open source MATLAB toolbox for localizing and visualizing human intracranial electrode data. *Journal of neuroscience methods*, 281, 40-48. <https://doi.org/10.1016/j.jneumeth.2017.01.022>
16. Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage*, 38(1), 95-113. <https://doi.org/10.1016/j.neuroimage.2007.07.007>



17. Li, X., Morgan, P. S., Ashburner, J., Smith, J., & Rorden, C. (2016). The first step for neuroimaging data analysis: DICOM to NIfTI conversion. Journal of neuroscience methods, 264, 47-56. <https://doi.org/10.1016/j.jneumeth.2016.03.001>
18. Omer Faruk Gulban, Dylan Nielson, John Lee, Russ Poldrack, Chris Gorgolewski, Vanessasaurus, & Chris Markiewicz. (2022). poldracklab/pydeface: PyDeface v2.0.2 (v2.0.2). Zenodo. <https://doi.org/10.5281/zenodo.6856482>.
19. <https://wiki.xnat.org/documentation/>
20. [Templeton World Charity Foundation](#)