HHS Public Access

Author manuscript

J Neural Eng. Author manuscript; available in PMC 2019 September 28.

Published in final edited form as:

J Neural Eng. ; 16(5): 056015. doi:10.1088/1741-2552/ab2c58.

Decoding task engagement from distributed network electrophysiology in humans

Nicole R Provenza[1,2,12], **Angelique C Paulk**[3,6], **Noam Peled**[4], **Maria I Restrepo**[5], **Sydney S Cash**[6], **Darin D Dougherty**[7], **Emad N Eskandar**[3,10], **David A Borton**[1,8,9,13], **Alik S Widge**[7,11,13]

[1]Brown University School of Engineering, Providence, RI, United States of America

[2]Charles Stark Draper Laboratory, Cambridge, MA, United States of America

[3]Massachusetts General Hospital Neurosurgery Research, Boston, MA, United States of America

[4]MGH/HST Martinos Center for Biomedical Imaging, Charlestown, MA, United States of America

[5]Center for Computation and Visualization, Brown University, Providence, RI 02912, United States of America

[6]Massachusetts General Hospital Neurology, Boston, MA, United States of America

[7]Massachusetts General Hospital Psychiatry, Boston, MA, United States of America

[8]Carney Institute for Brain Science, Providence, RI, United States of America

[9]Department of Veterans Affairs, Providence Medical Center, Center for Neurorestoration and Neurotechnology, Providence, RI, United States of America

[10]Present affiliation: Chair, Department of Neurological Surgery, Montefiore Medical Center, New York, NY, United States of America

[11]Present affiliation: Department of Psychiatry, University of Minnesota, Minneapolis, MN, United States of America

Abstract

Objective.—Here, our objective was to develop a binary decoder to detect task engagement in humans during two distinct, conflict-based behavioral tasks. Effortful, goal-directed decision-making requires the coordinated action of multiple cognitive processes, including attention, working memory and action selection. That type of mental effort is often dysfunctional in mental disorders, e.g. when a patient attempts to overcome a depression or anxiety-driven habit but feels

[12]Author to whom any correspondence should be addressed. nicole_provenza@brown.edu.

[13]Co-Principal Investigators, equal contributors.

Author contributions

NRP, DDD, ENE, DAB, and ASW conceived the concept of detecting effortful mental states as a strategy for controlling adaptive DBS. ENE performed all neurosurgeries. ACP curated and cleaned neural recordings. NP performed electrode localization. NRP carried out all data analysis with input from DAB, ASW, ACP, and IR

Data and materials availability

Deidentified data is available upon request at https://transformdbs.partners.org.

unable. If the onset of engagement in this type of focused mental activity could be reliably detected, decisional function might be augmented, e.g. through neurostimulation. However, there are no known algorithms for detecting task engagement with rapid time resolution.

Approach.—We defined a new network measure, fixed canonical correlation (FCCA), specifically suited for neural decoding applications. We extracted FCCA features from local field potential recordings in human volunteers to give a temporally continuous estimate of mental effort, defined by engagement in experimental conflict tasks.

Main results.—Using a small number of features per participant, we accurately decoded and distinguished task engagement from other mental activities. Further, the decoder distinguished between engagement in two different conflict-based tasks within seconds of their onset.

Significance.—These results demonstrate that network-level brain activity can detect specific types of mental efforts. This could form the basis of a responsive intervention strategy for decision-making deficits.

Keywords

neural decoding; network connectivity; task engagement

1. Introduction

Deficits in effortful executive function, including cognition, attention, and conflict resolution, are a core feature of many mental illnesses [1–6]. Existing treatments for mental illness are less than 50% effective, in part because they rarely address these cognitive symptoms. Many patients are left with behavioral deficits that impact their ability to function normally [7]. Electrical deep brain stimulation (DBS) can directly modulate the circuits underlying abnormal behaviors [8] and has been proposed as a more effective approach for treating mental illnesses, including major depressive disorder (MDD) and obsessive-compulsive disorder (OCD) [9–11]. Randomized clinical trials of DBS for MDD and OCD have, however, had mixed results [8, 12–14]. Titrating DBS to achieve a desired psychologic effect, an approach commonly called ‘closed-loop’ or ‘adaptive’ DBS (aDBS), may be a more effective approach [8, 15, 16]. In a psychiatric aDBS paradigm, detection of a specific goal-directed effort could trigger stimulation to modulate neural activity in a way that augments executive function during that effort [15, 16]. Stimulating the brain during well-defined behavioral states, for example during experimental psychophysical tasks, has repeatedly been shown to enhance mental functions, including value judgments and associative processes [17–21].

A challenge is that there is no established signature of this cross-cutting concept of engagement in mentally demanding tasks, in part because it depends on multiple cognitive systems. To complete an effortful task, one must attend to relevant stimuli, hold pertinent information in working memory, incorporate past experiences with present context to make a decision, and act. A single decision requires multiple interdependent processes, including attention, working memory, decision making, and action. As such, these behaviors are not encoded in any single region, but instead involve the rapid coordination of multiple brain regions and are reflected in the activity of brain networks. Detecting network states

associated with task engagement and stimulating during those states could be a more efficient path towards aDBS [16].

A particularly useful state to target for intervention is the mental process surrounding decision-making during conflict. In controlled laboratory environments, patients with mood, anxiety, and developmental disorders show abnormal behavioral and neural responses to decisional conflict [22–25]. These abnormal responses may manifest as perseverative incorrect responses, slowed reaction times, or impulsive errors [26–28]. In some cases, patients appear behaviorally normal, but show markedly abnormal neural responses, suggesting that their decision-making circuits are unusually taxed by conflict [24, 29]. Brain stimulation applied during conflict can bias decisions and behaviors, suggesting that such stimulation can enhance the effortful control required for optimal decisions [17, 30]. Thus, if we can identify when a patient is engaged in mental processes related to resolving conflict (e.g. including attention, working memory, action selection, etc), DBS could be applied to prospectively bias a decision-making process before it begins. Detecting the presence of mental states related to task engagement is critical to the therapeutic strategy. Brain stimulation delivered during other mental activity might interfere with normal, healthy processing or might amplify an abnormality.

A further challenge is that goal-directed efforts, including the resolution of decisional conflict, take place over the course of seconds. Task engagement must thus be detected at a similarly fine temporal resolution. While motor intent can be decoded at those time scales from single units and local field potentials [31–33], there is no known method to decode specific mental efforts within seconds of initiation. Additionally, goal-directed states involve the coordinated activity of multiple brain areas. Detecting task engagement associated with decision-making might thus require broad, multi-site measurement of brain activity. Detection would need to be based on efficient computation of high-level neural features that describe network-level brain activity [29, 34]. In the present study, we demonstrate such a decoder of task engagement using a functional connectivity operator on cortical and sub-cortical local field potential recordings across up to 30 brain regions. The operator, a derivative of Canonical Correlation Analysis (CCA) termed Fixed CCA (FCCA), is a simple linear combination of individual channels that can be efficiently applied to successive windows of neural data. Considering goal-directed mental efforts during two conflict-based tasks, we show that our classification algorithm can accurately distinguish task engagement from free behavior on the timescale of seconds. We further show that mental state classification can be achieved with a small subset of critical neural features, and that the approach can be extended to distinguish between two distinct types of goal-directed effort.

2. Methods

2.1. Study design

We collected an exhaustive sample of all the willing participants during the fixed study period.

Research subjects were 17 participants, each with a history of long-standing pharmaco-resistant complex partial seizures, that underwent clinically-indicated invasive monitoring to

confirm seizure focus. Neural activity during task performance was recorded while participants were hospitalized awaiting seizures for clinical mapping. The decision to implant electrodes and the number, types, and location of implantations were all determined on clinical grounds by a team of caregivers independent of this study. Participants were informed that participation in the experiment would not alter their treatment in any way and that they could withdraw at any time without jeopardizing their clinical care. Each participant gave fully informed consent according to study sponsor guidelines, and all procedures were approved by the local institutional review board at Partners Healthcare (Massachusetts General Hospital) and the Army HRPO.

2.2. Behavioral tasks

Participants performed two conflict-based tasks: the multisource interference task (MSIT) and the emotional conflict resolution (ECR) task (figures 1(A)–(D)). Task stimuli were presented with Presentation or Psychophysics Toolbox [35–37]. MSIT [38, 39] and the ECR task [40] have been designed to reliably and robustly activate networks associated with cognitive and emotional conflict, respectively.

MSIT runs comprised one to five 64-trial blocks. During each trial, a fixation cross was presented for two seconds followed by an image of three numbers between zero and three. One of the numbers, the ‘target’, was different than the other two numbers, the ‘distractors’. Participants were given a keypad and instructed that keypad numbers one, two, and three represented the numbers one, two, and three on the screen. Participants were then instructed to identify the target by pressing the button corresponding to its value, ignoring its position. Each trial was considered as ‘congruent’ or ‘incongruent’ depending on difficulty level. If the distractor was flanked by invalid targets (zero), the trial was congruent. If the distractor was flanked by valid targets, the trial was incongruent. In incongruent trials, the target was also out of position relative to its corresponding keypad button [38, 39]. Congruence changes from trial to trial were evenly balanced in number and frequency within each block (i.e. there were never more than two trials of the same type in a row, and the number of congruent trials followed by congruent trials was the same as the number of congruent trials followed by incongruent trials, and vice versa).

ECR task runs comprised one to six 64-trial blocks. During each trial, a fixation cross was presented for 2 s followed by an image of an emotive face showing happy or fearful expressions. Images were drawn from a set of expressions with five identifiable male faces and six identifiable female faces [41]. The words ‘FEAR’ and ‘HAPPY’ were overlaid on the faces such that the word and expression either matched (congruent) or did not match (incongruent) for each trial. The images were presented in a pseudorandom order such that the identity, gender, congruence, and valence (fear or happy) did not occur more than twice in a row.

The ECR task exposes functional connectivity differences between psychiatric participants and healthy controls. Generalized anxiety disorder (GAD) patients are less able to regulate emotional conflict during the ECR task than are healthy controls [40, 42]. Likewise, the MSIT task exposes functional connectivity differences between major depressive disorder (MDD) participants and obsessive-compulsive disorder (OCD) participants relative to

healthy controls [43, 44]. These tasks thus not only measure decisional conflict but may have relevance for the diagnosis and classification of mental disorders. The tasks were selected as part of the development of a larger psychophysical diagnostic battery [15].

Datasets containing less than 20% of data from any class (MSIT, ECR, or non-task) were excluded. No outliers were excluded.

2.3. Electrode localization

Electrodes were localized and mapped to standard brain regions using a volumetric image co-registration procedure. Using Freesurfer scripts (http://surfer.nmr.mgh.harvard.edu), the preoperative T1-weighted MRI was aligned with a postoperative CT. Electrode coordinates were manually determined from the CT and placed into the native space [45].

Mapping to brain regions was performed using the electrode labeling algorithm (ELA) (https://github.com/pelednoam/ieil). The ELA estimates the probability that a particular brain region contributes to the dipoles that constitute the source of the recorded signal. Electrodes implanted in the brain receive signals from multiple sources, including white matter and gray matter. The ELA identifies the signal source by identifying the probability that a given electrode is in a labeled region of the brain (gray matter). A probability was assigned to each electrode based on the likelihood that it was in a gray matter region of interest, as parcellated using Freesurfer and the DKT (Desikan-Killiany-Tourville) brain atlas [46, 47].

2.4. Invasive electroencephalography recordings

Depth electrodes (Ad-tech Medical, Racine WI, USA, or PMT, Chanhassen, MN, USA) with diameters of 0.8–1.0 mm and 8–16 2.4 mm platinum/iridium-contacts on each lead were stereotactically placed for seizure localization. The participants received bilateral electrodes ranging from five to nine electrodes in the right hemisphere, and five to eight electrodes in the left hemisphere. Each participant's electrode montage was determined solely by clinical mapping considerations.

Intracranial local field potential (LFP) recordings were acquired using one or two neural signal processor (NSP) recording systems (Blackrock Microsystems Inc., Salt Lake City, UT) at a sampling rate of 2 kHz. Table S1 (stacks.iop.org/JNE/16/056015/mmedia) includes a summary of the invasive electroencephalography (iEEG) LFP recordings collected per participant. Depth recordings were referenced to one scalp EEG electrode during acquisition.

2.5. Data preprocessing and signal conditioning

Data analysis was performed using custom analysis code in Matlab (MathWorks) and Fieldtrip [48]. All data were down-sampled to 1000 Hz and demeaned relative to the entire recording. Line noise and its harmonics up to 200 Hz were removed by subtracting a bandpass filtered signal from the raw signal on each channel. Neighboring channels were bipolar referenced relative to one another to reduce the effects of volume conduction [49].

2.6. Channel selection

Channels that exhibited excessive line noise or no discernible signal were removed from the analysis. Based on clinical reports and on visual inspection, electrodes surrounding the epileptic focus and/or exhibiting abnormal activities were also excluded.

2.7. Feature estimation: fixed canonical correlation and coherence analysis

A derivative of traditional canonical correlation analysis (CCA), fixed canonical correlation analysis (FCCA) and fixed canonical coherence analysis (FCHA) was used to estimate functional connectivity between regions over time. Here, traditional CCA was modified using the equation described below to apply a fixed projection space across all observations to compare feature estimations over time.

First, channels were organized into corresponding regions of interest based on electrode localization results. For example, regions X and Y contain two groups of signals that vary with time, t. The number of signals included in regions X and Y was determined by the minimum number of signals mapped to region X and region Y by the ELA. In this way, the same number of signals was considered for each region. Time domain inputs result in canonical correlation estimates, and frequency domain inputs result in canonical coherence estimates. Frequency domain inputs are a continuous power signal averaged over frequency bands of interest (4–8 Hz, 8–15 Hz, 15–30 Hz, 30–55 Hz, and 65–200 Hz) after time-frequency decomposition using Morlet wavelet decomposition using the 'ft_specest_wavelet' function in Fieldtrip (width = 3, gwidth = 7) [48].

For two sets of data, $X[=]\ p \times n$ and $Y[=]\ p \times n$, covariance matrices were defined as Cov $(X)= \Sigma_{XX} = XX^T$ and Cov $(X,\ Y)= \Sigma_{XY} = XY^T$, where Cov $(Y,\ Y)= \Sigma_{YY} = XY$. For coefficient vectors a and b, G and H are defined as linear combinations of X and Y, where $G = a^T * X$ and $H = b^T * Y$. The goal is to find the optimum linear combination of the measurements within each set, such that the resulting series are maximally correlated. Correlation between X and Y can be written in terms of the variance and covariance:

$$\rho = \frac{a^T X^T Y b}{\sqrt{(a^T X^T X a)(b^T Y^T Y b)}}. \quad (1)$$

As a continuation of equation (1) described above, two sets of data, X and Y, can be reduced using singular value decomposition, where $X = U_X \Sigma_X V_X$ and $Y = U_Y \Sigma_Y V_Y$. The simplified solution of the maximization problem can be found by computing:
$Q_{XY} = \mathrm{Cov}(X) * \mathrm{Cov}(X, Y) * \mathrm{Cov}(Y, Y)$, which simplifies to:

$$Q_{XY} = U_X V_X^T V_Y U_Y^T. \quad (2)$$

Q_{XY} can be reduced using singular value decomposition, where a and b can be defined in terms of the covariance of X and Y, and the singular value decomposition of Q_{XY}:

$$Q_{XY} = U_Q \Sigma_Q V_Q . \quad (3)$$

The principal vectors of a and b can be used to define a fixed basis or projection space where:

$$a = XX^T U_Q \quad (4)$$

$$b = YY^T V_Q \quad (5)$$

Thereafter, Fixed Canonical Correlation and Coherence values can be calculated for any new sample of data via a simple multiplication step (equation (1)).

2.8. Classification

2.8.1. Class label assignment.—Data were windowed using a four second window sliding every two seconds. Each four second window of data was labeled as ‘task’ or ‘non-task’ depending on whether the participant was actively engaged in the MSIT and ECR behavioral tasks. Task engagement was defined by periods where the task was on the presentation monitor, (i.e. both image presentation and fixation periods are included in the ‘task’ label assignment). A three-second buffer was placed at the start and end of ‘task’ periods to allow for ramp-up and ramp-down periods (i.e. the three seconds before and after the task began and ended were labeled as ‘task’). The length of ‘task’ periods are determined by the number of task blocks the participant agreed to participate in. ‘Non-task’ labels were assigned to portions of the recording collected immediately before and after task performance, as well as during any breaks the participant opted to take between task blocks. This decision was based on our assumption that these free-behaving data interleaved with periods of blocked task engagement are well suited for functional connectivity estimations of the unique features of the task state [50]. The length of ‘non-task’ periods was arbitrarily determined by the number and length of breaks the participant opted to take during the experiment. During ‘non-task’ periods, the participant was behaving freely and was not engaged in the task. Common free behaviors included asking questions of the experimenter, conversing with family members, using a smartphone or tablet recreationally, and toileting. Task blocks only were used to compute the fixed projection spaced described in equations (4) and (5).

2.8.2. Classification model and prediction.—A support vector machine (SVM) classifier with a linear kernel was designed using the LIBSVM Matlab implementation [51]. FCCA and FCHA features and corresponding ‘task’ versus ‘non-task’ class labels were used as classifier input.

2.8.3. MSIT versus ECR versus non-task classification model and prediction.—A multiclass error-correcting output codes (ECOC) model was designed using the Matlab function, 'fitcecoc'. Three binary SVM models were trained using one versus one coding design, such that one class is positive, one class is negative, and the other is ignored. FCHA and corresponding 'ECR task', 'MSIT', and 'non-task' class labels were used as classifier input. The one versus one coding scheme applies a majority voting scheme to determine the predicted label of new observations.

2.8.4. Feature selection.—For MSIT versus non-task feature selection, training set features were correlated with class labels, and the top 25 correlated features across all frequency bands of interest were run through a greedy-like feature selection process. This was done under the assumption that the features most correlated with class labels would typically align with the optimal features selected through the greedy-like process. For example, consider a feature set of size n. During our greedy-like feature selection process, classification accuracy is re-calculated for the feature set (of size $n - 1$) after each feature is removed, and subsequently replaced. The feature whose removal retains maximum classification accuracy is then removed, and the process is repeated until the size of the feature set is one. Optimal features were defined as the feature set that yielded maximum classification accuracy, up to five features. Feature sets from all frequency bands were included in feature selection.

MSIT versus ECR versus non-task feature selection was conducted in the same manner as MSIT versus non-task feature selection, using the top 50 correlated features across all frequency bands of interest. Up to 25 optimal features were chosen to achieve optimal classification performance.

2.8.5. Time resolution of state transition detection.—Multi-task datasets (MSIT versus ECR versus non-task, $N = 4$) were used to evaluate the time resolution of detecting state transitions. The classifier was trained on four fifths of the dataset and tested on the entire dataset. Testing on the entire dataset was necessary to capture the data before, during, and after transitions from one state (e.g. MSIT) to another (e.g. ECR) in chronological order. Based on the predicted class labels, the percent of transitions that were detected within 10 s of state onset was quantified. Of these successful detections, the mean amount of time between state onset and detection was quantified offline.

2.8.6. Evaluation of decoder stability.—Four participants, P6, P9, P13, and P15 completed two sessions of the MSIT, and three participants, P6, P12, and P13, completed two sessions of the ECR task, where sessions were separated by at least four hours. Classification performance and feature selection were compared within participants over both sessions of each task. Only bipolar electrodes present in both recordings after preprocessing were included in the analysis. For example, if an electrode was included in the first recording session and eliminated during the second recording session (e.g. due to excessive line noise), the electrode would be excluded in both test datasets. The stable basis space required for feature selection was determined from only the training set.

To improve classification performance, stable, discriminative features were defined under two conditions: (1) task and non-task FCHA values must be significantly different during both the first and the second recording session ($\alpha < 0.05$), (2) task and non-task FCCA and FCHA values must be consistent (not significantly different within a data type) between the first and second recording session ($\alpha > 0.05$). For each participant, features were identified that met the stability conditions using only training set data.

2.9. Statistical analysis

2.9.1. Cross validation.—Classifier performance was assessed with a five-fold cross validation strategy. The time series of features was separated into five folds. Four folds were used for training and the remaining fold was used for testing. This process was repeated until every fold was used for testing.

For multi-session classification, four fifths of data from the first recording session and four fifths of the data from the second recording session was used for training. The remaining fifths from each session were used for testing. This process was repeated until every fold was used for testing.

2.9.2. Class balancing.—Task data exceeded non-task data for 18 of 24 recording sessions analyzed. To balance unequal ‘task’ and ‘non-task’ class sizes, a data-level approach agnostic of the classification algorithm was used. The class of smaller size was augmented by random oversampling with replacement within the training set and testing set to make up the difference between class sizes [52]. To preserve the chronological structure of the data during classification, the feature set was windowed before training and testing set labels were assigned. The supplemental feature set was then concatenated to the original feature set. Datasets with less than 20% non-task data were dropped from analysis (7 recording sessions).

2.9.3. Construction of confidence intervals and chance classification performance.—Five-fold cross validation and class balancing was bootstrap-resampled 1000 times with replacement for construction of confidence intervals. Accuracy, true positive rate (sensitivity), true negative rate (specificity), false positive rate, and false negative rate were calculated on each iteration of the five-fold cross validation strategy to gauge classifier performance. Task observations were designated as the positive class, while non-task observations were designated as the negative class. Classification performance due to chance was calculated by shuffling class labels before class sizes were balanced. Labels were randomly assigned in a way that preserved original class sizes before classes were balanced. The median accuracy across the 1000 repetitions of five-fold cross validation was reported for each participant for true and shuffled labels in figures 2(A)–(C).

2.9.4. Hypothesis testing.—The Wilcoxon Rank Sum test was used to compare median SVM accuracy across different feature types for each participant ($\alpha < 0.05$). Bonferroni correction was used to adjust p-values with a factor of 21 (seven choose two, i.e. the number of possible pairwise comparisons). The Wilcoxon Rank Sum test was also used to test for separation from chance by comparing median SVM accuracy for true and shuffled

labels for each feature type. The Bonferroni correction here was a factor of 7, i.e. the number of different analyses performed.

The binomial test was used to determine if optimal features in a particular frequency band of interest were more likely to occur than what would be expected by chance ($\alpha < 0.05$). Assuming that optimal features in each of the five frequency bands of interest are equally likely to occur due to chance, binomial probability is 0.2.

3. Results

3.1. Participants accurately performed psychophysical tasks during invasive electroencephalography recordings

Local field potentials (LFP) were recorded from the brain via intracranial electroencephalography (iEEG) (table S1) while 17 participants performed the Multi-Source Interference Task (MSIT) (figures 1(A) and (B)), emotion conflict resolution (ECR) task (figures 1(C) and (D)), or both. The participants performed the tasks correctly with high accuracy (figures 1(E) and (F)), and showed expected conflict effects on behavior (figures 1(G) and (H)) [38, 41]. Median accuracy for MSIT congruent and incongruent trials was 100% ± 2.47% and 97.1% ± 5.52% respectively, and median accuracy for ECR congruent and incongruent trials was 93.0% ± 13.3% and 86.4% ± 21.8% respectively. LFP data was labeled according to whether the participant was engaged in MSIT, ECR, or non-task free behavior. Each iEEG electrode contact was assigned to an anatomic region of interest label (ROI, figure 1(I)), and fixed canonical correlation (FCCA) and fixed canonical coherence (FCHA) network features were extracted from the LFP. The dataset was balanced to ensure an equal number of observations in each class. FCCA and FCHA features were then fed to a support vector machine (SVM) classifier to distinguish task engagement from free behavior (figure 1(J)).

3.2. Fixed canonical correlation and coherence features distinguish task engagement from free behavior

We quantified SVM performance by how accurately, sensitively, and specifically the decoder detected task and non-task states. FCCA and FCHA features accurately decoded task engagement in humans performing two conflict tasks, MSIT ($N = 14$) and ECR ($N = 12$) (figures 2(A) and (B); table S2 and S3). FCCA time-domain features classified behavior significantly above chance. Classification performance was similar for FCHA, the frequency-domain analogue (tables S4 and S5), where we subdivided the frequencies into bands identified to be behaviorally relevant (4–8 Hz, 8–15 Hz, 15–30, 30–55 Hz, 65–200 Hz). Median accuracy for MSIT versus non-task classification using FCCA and FCHA features across all frequency bands was 70.94% ± 6.41% and 74.10% ± 8.69%, respectively. Median accuracy for ECR versus non-task classification using time domain features and all frequency domain features was 71.70% ± 5.97% and 69.68% ± 5.58%, respectively. There was no significant difference between time-domain and frequency-domain classification results when using all frequency bands of interest. Further, decoding from power (as opposed to FCHA) features was at chance performance (MSIT versus non-task and ECR versus non-task classification of 50.34% ± 3.58%, and 50.05% ± 3.41%, respectively, figure

3). This verifies that the effortful processes we sought to decode require a network-level analysis.

Limiting features to individual frequency bands of interest did not improve and in some cases decreased classification performance (figures 2(A) and (B)). Pertinent information for decoding task engagement is not encoded in any single frequency band, and the encoding appears to be idiosyncratic to each participant. Performance was greater for correctly identifying non-task states than task states, i.e. specificity exceeded sensitivity (figures 2(C)–(F)). This may be explained by the relative frequency of the two states. The difference between sensitivity and specificity across participants had a positive, linear relationship with the number of observations in each class in the dataset before class balancing (figure 4). For datasets with close to an equal number of task and non-task observations before class balancing (e.g. Participant 14), the discrepancy was close to zero. Further, to correct for any possible distortion induced by our class resampling [52], we repeated the analysis shown in figure 2 using a balanced accuracy metric [53], which mitigates this limitation at the cost of lower statistical power. Bonferroni corrected *P*-values were overall closer to 1, but still above chance for MSIT versus Non-task classification using all FCCA and FCHA features, and ECR versus Non-task classification using FCCA features, and 4–8 Hz and 8–15 Hz FCHA features (tables S6 and S7).

3.3. A subset of optimal frequency domain features maintains decoder performance

Next, we were interested in isolating a small number of important features for classifying task and non-task activity without a significant loss in accuracy. We noted that some canonical features showed visible changes at task-state boundaries (figure 5(A)), and reasoned a small number of that these strongly modulated features might be sufficient for decoding. We identified features that strongly correlated with task-states through a principled approach involving iterative feature dropping (figure 5(B)). Classification with the optimal time-domain and frequency-domain features exceeded chance (MSIT versus non-task time: $P = 7.8 * 10^{-18}$; MSIT versus non-task frequency: $P = 4.0 * 10^{-13}$; ECR versus non-task time: $P = 2.4 * 10^{-19}$; ECR versus non-task frequency: $P = 6.4 * 10^{-21}$, Wilcoxon signed rank test) (figure 5(C)) and was equivalent to classification with all features. Frequency domain features yielded greater median performance than using time domain features, however this difference was not significant (MSIT time versus frequency: $P = 0.22$; ECR time versus frequency: $P = 0.56$, Wilcoxon signed rank test). Between one and five FCHA features per participant yielded optimal performance. Median accuracy for MSIT versus non-task classification and ECR versus non-task classification using up to five optimal FCHA features was 78.10% ± 7.39% and 78.97% ± 7.36%, respectively (figure 5(D)). Next, we sought to identify trends in optimal features across patients for distinguishing MSIT and ECR from non-task.

Certain features were highly discriminative across both tasks (figure 5(E)). Each FCHA feature incorporates three variables: two ROIs, and the frequency band of connectivity between them. Trends in optimal features were assessed by (1) the number of optimal connections between region pairs pertaining to each individual frequency band, and (2) the degree of connectivity to each ROI, agnostic of frequency band. Across patients, the 4–8 Hz

frequency band distinguished MSIT, while the 65–200 Hz frequency band distinguished ECR (figure 5(F)). After normalizing with respect to the average number of electrodes implanted in each ROI, we identified ROIs with a high degree of connectivity as measured by FCHA features (points beyond one standard deviation above the trendline, figure 5(G)). These ROIs by definition have unusually high influence over decoding. Connectivity to/from the ventrolateral prefrontal cortex (vlPFC) distinguished both tasks from the non-task state. The amygdala and temporal lobe showed a greater degree of connectivity during the MSIT than the ECR task, while the lateral orbitofrontal cortex (lOFC) and hippocampus showed a greater degree of connectivity during the ECR task than the MSIT. Differing optimal features for detecting MSIT and ECR suggest that it is possible to distinguish between the two.

3.4. Decoder distinguishes two distinct modalities of conflict-based tasks

Changes in canonical coherence estimates can also distinguish between MSIT and ECR despite the similarity of these tasks. MSIT versus ECR versus non-task decoder accuracy using all FCHA features was significantly greater than chance performance ($P = 0.028$, Wilcoxon signed rank test). Decoding with a subset of 25 optimal FCHA features was again non-inferior in terms of accuracy to all features ($P = 0.2$, Wilcoxon signed rank test), and was numerically slightly better (figure 6(A)). The dlPFC and caudate were implicated across participants for distinguishing between these task states, however our sample size was small due to the unique alternating task design ($N = 4$). Median non-task detection rate, MSIT detection rate, and ECR task detection rate using optimal canonical coherence features was 49.17% ± 6.80%, 75.18% ± 5.42%, and 87.60% ± 4.07%, respectively (figure 6(B) and table S10). In this way, the decoder could be applied across time to detect the type of mental activity in which a patient is engaged, moment to moment (figure 6(C)). The lag between state onset and detection for the transition to MSIT, ECR, and non-task states was quantified. The decoder successfully detected 82.2% ± 15.34% of the state transitions. Of the successful detections, the decoder detected state transitions within 1.30 s ± 2.15 s after state onset.

3.5. Subhead 5: decoder features and classification are stable between sessions

In addition to distinguishing between multiple tasks, another important consideration for a translatable decoder is stability over time. In a subset of participants (MSIT $n = 4$, ECR $n = 3$), the same task was performed during two recording sessions separated by at least four hours. We extended our automatic feature selection approach to identify highly stable features, defined as those that discriminated task from non-task but did not change substantially between recording sessions (figures 6(D) and (E)). These ‘optimally stable features’ showed classification above chance on a held-out dataset containing data from both recording sessions (figure 6(F)), (MSIT versus non-task: $P = 0.029$; ECR versus non-task: $P = 0.029$, Wilcoxon signed rank test). Classification using all features, by contrast, did not exceed chance performance (figure 6(F)), (MSIT versus non-task: $P = 0.51$; ECR versus non-task $P = 0.70$, Wilcoxon rank sum test).

4. Discussion

The ability to productively engage in and complete goal-directed behaviors is critical for mental function and is impaired for persons with mental illness [54]. Detecting when someone is engaged in a specific goal-directed behavior could be used as a control signal to trigger stimulation and augment executive function during that behavior [15, 55]. Here, a decoder was developed to discriminate free behavior from engagement in two distinct conflict-based tasks. A modified version of canonical correlation and coherence analysis, FCCA and FCHA, extracted neural features related to goal-directed behaviors during decisional conflict. These network-level neural features, particularly in the frequency-domain, successfully distinguished periods of task engagement from free behavior. Network level analysis was necessary: band-power features at the single electrode level yielded decoding accuracies that were effectively at chance (50%) performance. Decoder performance was maintained even when using five or fewer optimal features per participant. The algorithm distinguished two very similar mental tasks within a single recording session in all four participants with available data for task comparison. Further, stable features were identified across two temporally separated recording sessions that enabled accurate detection over multiple days.

We propose that detecting distinct mental states related to task engagement is a viable approach for enabling aDBS for psychiatric illness. Here, we define task engagement as a global state that encompasses many well-studied, discrete processes, including perceptual judgement, object recognition, working memory, goal representation, and action selection. We have provided a proof-of-concept demonstration using behavioral tasks related to decisional conflict. In future clinical applications, when a patient is engaged in mental efforts related to decisional conflict, our decoder could be used to trigger stimulation that would improve the chance of an adaptive, healthy response. Alternatively, our decoder could act as a higher-level controller, where detection of efforts to resolve decisional conflict would then trigger a second closed-loop algorithm that determined whether the ongoing behavior was more likely to be healthy or pathological. We have described prototypes of such ‘lower level’ closed-loop stimulation algorithms in other recent work [15, 56, 57].

The ability to make fine-grained discriminations is necessary for aDBS applications [16]. A critical feature of our work is that it can identify task engagement within seconds of onset, unlike past classifiers. Canonical correlation and coherence have been used to infer brain network connectivity during periods of speech and rest [58]. Cognitive states have been decoded from EEG with synchrony operators [59]. Neither of those prior efforts achieved time resolution—both worked only with blocked analysis of many trials. By adding a stabilized, fixed linear operator (the ‘F’ of FCCA and FCHA), we enable discrimination by repeatedly applying the same operator to successive neural data samples. During multi-state decoding, the decoder successfully detected 82.2% ± 15.34% of the state transitions within 10 s of state onset. Further, of the successful detections, the decoder detected state transitions within 1.30 s ± 2.15 s after state onset. Fine-grained temporal resolution would be essential to successful aDBS, since individual decisions are usually made over the course of a few seconds.

We found that network functional connectivity features at particular frequency bands of interest strongly distinguished task engagement from non-task states, while power frequency features computed at the single electrode level did not distinguish task from non-task behavior. Recent work has shown that network features are essential for distinguishing other complex behavioral states, specifically variation in mood over time [60]. During both tasks, discrimination was strongly driven by connectivity to and from the ventrolateral prefrontal cortex (vlPFC). vlPFC is implicated in cognitive control, task switching, and working memory. vlPFC shows fMRI activation during MSIT, along with the dACC, dlPFC, and superior parietal cortex [38, 61]. 4–8 Hz oscillations, which in our analysis discriminated MSIT from free behavior, are strengthened in vlPFC when patients receive DBS that enhances MSIT performance [55]. The amygdala and temporal lobe were highly discriminative during the MSIT, while the lateral orbitofrontal cortex (lOFC) was highly discriminative during the ECR task. The lOFC is implicated in making choices that involve response suppression, and is inhibited in response to presentation of angry faces. lOFC activation is not typically associated with the ECR task via fMRI [41, 62], but has recently been shown to be involved in emotion processing through acute neurostimulation studies [63]. These results suggest that by virtue of its coarse temporal resolution, non-invasive imaging may not fully capture neural dynamics important to task processing. However, like fMRI, FCCA and FCHA are not directed connectivity metrics. Future studies could explore directed connectivity metrics akin to recent results studying feedback learning during MSIT [34].

One might expect that optimal features would align with oscillatory activity that has been implicated in conflict resolution (e.g. mid-frontal theta band activity) [64–67]. FCHA features in the theta band were particularly prevalent among the identified subset of optimal features for MSIT detection. However, FCHA features in the high gamma band (65–200 Hz), and not the theta band, were prevalent in the optimal features for discriminating ECR. Thus, the neural signature of task engagement is broader than the neural signature for conflict resolution during high interference trials. Further, our analysis was not designed to identify neural signatures of conflict. A decoder focused on that problem would specifically try to discriminate high-from low-conflict trials (a ‘contrast’ approach). We created decoders that detect the activity common to high- and low-conflict trials, i.e. that detect the general state of focus on decisional conflicts. This is a different approach than the classic contrast analysis used in cognitive electrophysiology, and thus we expect it to yield different results.

Likewise, one might expect that optimal features for detecting task engagement would align with the brain regions involved in the default mode network. Task and rest are distinguished in fMRI studies by the activity of the default mode network, a resting-state network with hubs in posterior cingulate cortex, the medial prefrontal cortex (mPFC), parietal lobe, and medial temporal nuclei [68, 69]. Optimal features did not overlap with default mode network, as it is typically defined by periods of explicit quiet rest. Non-task periods here were freely chosen behaviors that consist of mostly non-effortful mental activities, e.g. watching television or light conversation. No electrodes were implanted in the vACC during our study, however, and connectivity to the posterior cingulate or mPFC did not distinguish task engagement.

Task engagement could be accurately decoded with a small number of optimal features. We show that the method we propose can be implemented with as few as five FCHA or FCCA features per participant (figure 5(C)) that span a total of four to nine ROIs (tables S8, S9 and S11). When we consider that a single linear lead/shank can easily sample 3 to 4 brain structures (figure 1(I)), the proposed method could easily be implemented with two–three total leads per brain hemisphere. The use of two implanted leads per hemisphere is now a common research practice and has been used to produce additional clinical benefit in some patients [33, 70–72]. Feature selection could reduce the decoder's computational requirements to a level suitable for real-time use, including in an implanted system. The reduced-feature decoders shown in figure 5 could process a single four-second window of data on a mid-grade commodity desktop PC in 50 ms. Existing closed-loop systems are limited, but next generation neural implants will likely have more computing power and capability and may be able to implement an algorithm that detects mental states based on network connectivity [73–75].

Also relevant to the design of future untethered systems, the optimal features differed across tasks and participants. The former was an expected result, as the MSIT and ECR tasks are similar, yet distinct. The ECR tasks involves substantial additional emotional processing and regulation components that the MSIT does not elicit [38, 39, 41]. The differences in optimal features among participants may be due to variability in the numbers of electrodes and their precise location within each brain region. Decoder performance was not, however, dependent on any one implant scheme. Successful decoder performance across participant specific electrode montages suggests that the networks involved in task-related brain states are broad and specific types of mental activity needed to perform conflict-related tasks is only present at the network level. In other words, a standard clinical implant montage with sufficiently broad coverage should support FCCA and FCHA decoding of a variety of goal-directed efforts. Moreover, while electrode placement was clinically variable, it is still quite possible to map electrodes to homologous regions across subjects, as we demonstrated here using our previously published electrode localization algorithm. We have successfully deployed this approach in other investigations, e.g. [76], as have other research groups using different approaches to assign electrodes to homologous regions [77].

Decoder stability is also important for translation; the ideal decoder would not require manual, daily re-calibration. Simply training on data from two temporally separated recording sessions was not sufficient to predict task engagement above chance levels. Identifying stable features through a principled statistical approach, however, recovered performance above chance. The multi-session classification was still slightly less accurate than single-session classification. This suggests that non-stationarities may exist in the data. These could arise from physical movement of electrodes, dropped signals, changes in recorded neural activity due to the evolving foreign body response, or variance in task performance on different days. Such non-stationarities may not present such a challenge in patients receiving a chronic implant, where the device-tissue interface can stabilize over weeks [78]. Further, training a classifier on more temporally separated data should improve stability. In the motor brain computer interface (BCI) field, iEEG recording sessions collected over tens to hundreds of days produce more robust and stable decoders [79–82].

The motor BCI field has already developed several approaches to decoder stabilization, adaptation, and long-term recalibration [83–85].

In addition to improving decoder stability, expanding the breadth of non-task related brain states used to train the classifier could strengthen classification performance over a wide range of activities and potentially improve stability. Here, our non-task periods involved free behavior that was not standardized across participants. This variability may have led to variability in results across participants. However, despite the non-uniform activities during non-task periods, the classifier was able to distinguish task engagement with high accuracy both within sessions and across days. Further, we expect that training the classifier during a wide array of non-task related brain states will strengthen the stability of the classifier over time. That is, when considering whether to activate an aDBS algorithm that may dramatically alter a patient's learning or self-regulation capabilities, it is critical to know when *not* to deploy that algorithm. Our approach leads to classifiers that emphasize specificity of detecting the task state, exactly what is needed for this application. That specificity is evident in the greater accuracy of classifying the non-task periods relative to the task periods, despite the fact that non-task is both more heterogeneous and the minority class in this dataset. One challenge in including a wide array of non-task data is that task-related networks may spontaneously reactivate during rest. Capturing a wide variety of activities during non-task state recordings will likely be important for training and stabilizing the classifier over time.

The decoding methods we have described may apply to many more contexts than the two specific tasks analyzed here (e.g. reward evaluation, evidence accumulation, learning, arithmetic, etc) as the decoder performs well for multiple conflict tasks and distinguishes between tasks. The generalizability of this approach will, however, need to be demonstrated in future studies using a broader cognitive task battery with emphasis on non-conflict tasks. Beyond aDBS, effortful cognitive state detection could be useful for monitoring of coma or states of altered consciousness, rehabilitating cognitive function after strokes, or performance monitoring in high stakes operational environments. While more development is required before clinical translation, stable, sensitive, and specific detection of goal-directed efforts could ultimately be used to drive DBS or other time-dependent interventions and improve the lives of psychiatric patients.

Acknowledgments

We thank participants and their families; and K Farnes, M Robertson, G Belok for assistance with data collection.

Funding

This research was funded by the Defense Advanced Research Projects Agency (DARPA) under Cooperative Agreement Number W911NF-14–2-0045 issued by ARO contracting office in support of DARPA's SUBNETS Program. The views, opinions, and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the US Government. DAB and NRP were supported by NIH NINDS BRAIN Initiative via contract 1UH3NS100549–01. NRP was supported by the Charles Stark Draper Laboratory Fellowship.

Competing interests

The decoding methods reported here overlap with a pending patent application filed by authors NRP, ENE, DDD, ASW, and DAB, ASW, DAB, and DDD receive device donations from Medtronic. ASW and DDD have received consulting income from Medtronic and have pending patent applications related to deep brain stimulation for mental illness. DDD further reports consulting income from Insys, speaking fees from Johnson & Johnson, and research support from Cyberonics and Roche.

References

[1]. Simpson JR, Snyder AZ, Gusnard DA and Raichle ME 2001 Emotion-induced changes in human medial prefrontal cortex: I. During cognitive task performance Proc. Natl Acad. Sci. USA 98 683–7 [PubMed: 11209065]

[2]. Whalen PJ, Bush G, McNally RJ, Wilhelm S, McInerney SC, Jenike MA and Rauch SL 1998 The emotional counting stroop paradigm: a functional magnetic resonance imaging probe of the anterior cingulate affective division Biol. Psychiatry 44 1219–28 [PubMed: 9861465]

[3]. Koyama T, Kato K, Tanaka YZ and Mikami A 2001 Anterior cingulate activity during pain-avoidance and reward tasks in monkeys Neurosci. Res 39 421–30 [PubMed: 11274741]

[4]. Niki H and Watanabe M 1979 Prefrontal and cingulate unit activity during timing behavior in the monkey Brain Res 171 213–24 [PubMed: 111772]

[5]. Nishijo H, Yamamoto Y, Ono T, Uwano T, Yamashita J and Yamashima T 1997 Single neuron responses in the monkey anterior cingulate cortex during visual discrimination Neurosci. Lett 227 79–82 [PubMed: 9180208]

[6]. Procyk E, Tanaka YL and Joseph JP 2000 Anterior cingulate activity during routine and non-routine sequential behaviors in macaques Nat. Neurosci 3 502–8 [PubMed: 10769392]

[7]. Insel TR 2009 Disruptive insights in psychiatry: transforming a clinical discipline J. Clin. Invest 119 700–5 [PubMed: 19339761]

[8]. Widge AS, Malone DA and Dougherty DD 2018 Closing the loop on deep brain stimulation for treatment-resistant depression Front. Neurosci 12 175 [PubMed: 29618967]

[9]. Mayberg HS 2009 Targeted electode-based modulation of neural circuits for depression J. Clin. Invest 119 717–25 [PubMed: 19339763]

[10]. Malone DA et al. 2009 Deep brain stimulation of the ventral capsule/ventral striatum for treatment-resistant depression Biol. Psychiatry 65 267–75 [PubMed: 18842257]

[11]. Greenberg BD et al. 2010 Deep brain stimulation of the ventral internal capsule/ventral striatum for obsessive-compulsive disorder: worldwide experience Mol. Psychiatry 15 64–79 [PubMed: 18490925]

[12]. Dougherty DD et al. 2015 A randomized sham-controlled trial of deep brain stimulation of the ventral capsule/ventral striatum for chronic treatment-resistant depression Biol. Psychiatry 78 240–8 [PubMed: 25726497]

[13]. Widge AS and Dougherty DD 2015 Deep brain stimulation for treatment-refractory mood and obsessive-compulsive disorders Curr. Behav. Neurosci. Rep 2 187–97

[14]. Holtzheimer PE et al. 2017 Subcallosal cingulate deep brain stimulation for treatment-resistant depression: a multisite, randomised, sham-controlled trial Lancet Psychiatry 4 839–49 [PubMed: 28988904]

[15]. Widge AS et al. 2017 Treating refractory mental illness with closed-loop brain stimulation: progress towards a patient-specific transdiagnostic approach Exp. Neurol 287 461–72 [PubMed: 27485972]

[16]. Provenza NR. et al. 2019; The case for adaptive neuromodulation to treat severe intractable mental disorders. Front. Neurosci. 13:152. [PubMed: 30890909]

[17]. Amemori KI and Graybiel AM 2012 Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making Nat. Neurosci 15 776–85 [PubMed: 22484571]

[18]. Williams ZM and Eskandar EN 2006 Selective enhancement of associative learning by microstimulation of the anterior caudate Nat. Neurosci 9 562–8 [PubMed: 16501567]

[19]. Katnani HA, Patel SR, Kwon C-S, Abdel-Aziz S, Gale JT and Eskandar EN 2016 Temporally coordinated deep brain stimulation in the dorsal and ventral striatum synergistically enhances associative learning Sci. Rep 6 18806 [PubMed: 26725509]

[20]. Ezzyat Y et al. 2017 Direct brain stimulation modulates encoding states and memory performance in humans Curr. Biol 27 1251–8 [PubMed: 28434860]

[21]. Ezzyat Y. et al. 2018; Closed-loop stimulation of temporal cortex rescues functional networks and improves memory. Nat. Commun. 9:365. [PubMed: 29410414]

[22]. Grisanzio KA, Goldstein-Piekarski AN, Wang MY, Rashed Ahmed AP, Samara Z and Williams LM 2017 Transdiagnostic symptom clusters and associations with brain, behavior, and daily function in mood, anxiety, and trauma disorders JAMA Psychiatry 94305 1–9

[23]. Robbins TW, Gillan CM, Smith DG, de Wit S and Ersche KD 2012 Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry Trends Cogn. Sci 16 81–91 [PubMed: 22155014]

[24]. McTeague LM, Huemer J, Carreon DM, Jiang Y, Eickhoff SB and Etkin A 2017 Identification of common neural circuit disruptions in cognitive control across psychiatric disorders Am. J. Psychiatry 174 676–85 [PubMed: 28320224]

[25]. Etkin A, Patenaude B, Song YJC, Usherwood T, Rekshan W, Schatzberg AF, Rush AJ and Williams LM 2015 A cognitive-emotional biomarker for predicting remission with antidepressant medications: a report from the iSPOT-D trial Neuropsychopharmacology 40 1332–42 [PubMed: 25547711]

[26]. Melcher T, Falkai P and Gruber O 2008 Functional brain abnormalities in psychiatric disorders: neural mechanisms to detect and resolve cognitive conflict and interference Brain Res. Rev 59 96–124 [PubMed: 18621078]

[27]. Trivedi MH and Greer TL 2014 Cognitive dysfunction in unipolar depression: Implications for treatment J. Affect. Disord 152–4 19–27

[28]. Fineberg NA, Potenza MN, Chamberlain SR, Berlin HA, Menzies L, Bechara A, Sahakian BJ, Robbins TW, Bullmore ET and Hollander E 2010 Probing compulsive and impulsive behaviors, from animal models to endophenotypes: a narrative review Neuropsychopharmacology 35 591–604 [PubMed: 19940844]

[29]. Etkin A, Egner T and Kalisch R 2011 Emotional processing in anterior cingulate and medial prefrontal cortex Trends Cogn. Sci 15 85–93 [PubMed: 21167765]

[30]. Cavanagh JF, Wiecki TV, Cohen MX, Figueroa CM, Sherman SJ and Frank MJ 2011 Influence over decision threshold Nat. Neurosci 14 1462–7 [PubMed: 21946325]

[31]. Hochberg LR, Serruya MD, Friehs GM, Mukand JA, Saleh M, Caplan AH, Branner A, Chen D, Penn RD and Donoghue JP 2006 Neuronal ensemble control of prosthetic devices by a human with tetraplegia Nature 442 164–71 [PubMed: 16838014]

[32]. Hochberg LR. et al. 2013; Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. Nature. 485:372.

[33]. Swann NC, de Hemptinne C, Thompson MC, Miocinovic S, Miller AM, Gilron R, Ostrem JL, Chizeck HJ and Starr PA 2018 Adaptive deep brain stimulation for Parkinson's disease using motor cortex sensing J. Neural Eng 15 046006 [PubMed: 29741160]

[34]. Smith EH, Banks GP, Mikell CB, Cash SS, Patel SR, Eskandar EN and Sheth SA 2015 Frequency-dependent representation of reinforcement-related information in the human medial and lateral prefrontal cortex J. Neurosci 35 15827–36 [PubMed: 26631465]

[35]. Pelli DG 1997 The VideoToolbox software for visual psychophysics: transforming numbers into movies Spat. Vis 10 437–42 [PubMed: 9176953]

[36]. Brainard D 1997 The psychophysics toolbox Spat. Vis 10 433–6 [PubMed: 9176952]

[37]. Kleiner M, Brainard D, Pelli D, Ingling A, Murray R and Broussard C 2007 What's new in psychtoolbox-3 Perception 36 1–16

[38]. Bush G and Shin LM 2006 The Multi-source interference task: An fMRI task that reliably activates the cingulo-frontal-parietal cognitive/attention network Nat. Protocols 1 308–13 [PubMed: 17406250]

[39]. Bush G, Shin LM, Holmes J, Rosen BR and Vogt BA 2003 The Multi-Source Interference Task: validation study with fMRI in individual subjects Mol. Psychiatry 8 60–70 [PubMed: 12556909]

[40]. Etkin A, Prater KE, Hoeft F, Menon V and Schatzberg AF 2010 Failure of anterior cingulate activation and connectivity with the amygdala during implicit regulation of emotional processing in generalized anxiety disorder Am. J. Psychiatry 167 545–54 [PubMed: 20123913]

[41]. Etkin A, Egner T, Peraza DM, Kandel ER and Hirsch J 2006 Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala Neuron 51 871–82 [PubMed: 16982430]

[42]. Etkin A and Schatzberg AF 2011 Common abnormalities and disorder-specific compensation during implicit regulation of emotional processing in generalized anxiety and major depressive disorders Am. J. Psychiatry 168 968–78 [PubMed: 21632648]

[43]. Davey CG, Yücel M, Allen NB and Harrison BJ 2012 Task-related deactivation and functional connectivity of the subgenual cingulate cortex in major depressive disorder Frontiers Psychiatry 3 1–8

[44]. Cocchi L, Harrison BJ, Pujol J, Harding IH, Fornito A, Pantelis C and Yücel M 2012 Functional alterations of large-scale brain networks related to cognitive control in obsessive-compulsive disorder Hum. Brain Mapp 33 1089–106 [PubMed: 21612005]

[45]. Dykstra AR, Chan AM, Quinn BT, Zepeda R, Keller CJ, Cormier J, Madsen JR, Eskandar EN and Cash SS 2012 Individualized localization and cortical surface-based registration of intracranial electrodes NeuroImage 59 3563–70 [PubMed: 22155045]

[46]. Desikan RS et al. 2006 An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest NeuroImage 31 968–80 [PubMed: 16530430]

[47]. Fischl B 2012 FreeSurfer NeuroImage 62 774–81 [PubMed: 22248573]

[48]. Oostenveld R, Fries P, Maris E and Schoffelen JM 2011 FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data Comput. Intell. Neurosci 2011 156869 [PubMed: 21253357]

[49]. Bastos AM and Schoffelen J-M 2016 A tutorial review of functional connectivity analysis methods and their interpretational pitfalls Frontiers Syst. Neurosci 9 175

[50]. Fair DA, Schlaggar BL, Cohen AL, Miezin FM, Dosenbach NUF, Wenger KK, Fox MD, Snyder AZ, Raichle ME and Petersen SE 2007 A method for using blocked and event-related fMRI data to study 'resting state' functional connectivity NeuroImage 35 396–405 [PubMed: 17239622]

[51]. Chang C-C and Lin C-J 2011 Libsvm ACM Trans. Intell. Syst. Technol 2 1–27

[52]. Galar M, Fernandez A, Barrenechea E, Bustince H and Herrera F 2012 A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches IEEE Trans. Syst. Man Cybern. C 42 463–84

[53]. Brodersen KH, Ong CS, Stephan KE and Buhmann JM 2010 The balanced accuracy and its posterior distribution Int. Conf. on Pattern Recognition 3125–8

[54]. Gillan CM, Kosinski M, Whelan R, Phelps EA and Daw ND 2016 Characterizing a psychiatric symptom dimension related to deficits in goal-directed control eLife 5 e11305 [PubMed: 26928075]

[55]. Widge A, Zorowitz S, Basu I, Paulk A, Cash S, Eskandar E, Deckersbach T, Miller E and Dougherty D 2019 Deep brain stimulation of the internal capsule enhances human cognitive control and prefrontal cortex function Nat. Commun 10 1536 [PubMed: 30948727]

[56]. Yousefi A, Paulk AC, Basu I, Mirsky JL, Dougherty DD, Eskandar EN, Eden UT and Widge AS 2019 COMPASS: an open-source, general-purpose software toolkit for computational psychiatry Front. Neurosci 12 957 [PubMed: 30686965]

[57]. Yousefi A, Basu I, Paulk A, Peled N, Eskandar E, Dougherty D, Cash S, Widge A and Eden U Decoding hidden cognitive states from behavior and physiology using a Bayesian approach Neural Comput (accepted)

[58]. Stephen EP, Lepage KQ, Eden UT, Brunner P, Schalk G, Brumberg JS, Guenther FH and Kramer MA 2014 Assessing dynamics, spatial scale, and uncertainty in task-related brain network analyses Frontiers Comput. Neurosci 8 31

[59]. Dvorak D, Shang A, Abdel-Baki S, Suzuki W and Fenton AA 2018 Cognitive behavior classification from scalp EEG signals IEEE Trans. Neural Syst. Rehabil. Eng 26 729–39 [PubMed: 29641377]

[60]. Kirkby LA, Luongo FJ, Lee MB, Nahum M, Van Vleet TM, Rao VR, Dawes HE, Chang EF and Sohal VS 2018 An amygdala-hippocampus subnetwork that encodes variation in human mood Cell 175 1688–700.e14 [PubMed: 30415834]

[61]. Badre D and Wagner AD 2007 Left ventrolateral prefrontal cortex and the cognitive control of memory Neuropsychologia 45 2883–901 [PubMed: 17675110]

[62]. Elliott R, Dolan RJ and Frith CD 2000 Dissociable functions in the medial and lateral orbitofrontal cortex: evidence from human neuroimaging studies Cereb. Cortex 10 308–17 [PubMed: 10731225]

[63]. Rao VR et al. 2018 Direct electrical stimulation of lateral orbitofrontal cortex acutely improves mood in individuals with symptoms of depression article direct electrical stimulation of lateral orbitofrontal cortex acutely improves mood in individuals with symptoms of depr Curr. Biol 28 3893–902.e4 [PubMed: 30503621]

[64]. Nigbur R, Ivanova G and Stürmer B 2011 Theta power as a marker for cognitive interference Clin. Neurophysiol 122 2185–94 [PubMed: 21550845]

[65]. Cohen MX and Donner TH 2013 Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior J. Neurophysiol 110 2752–63 [PubMed: 24068756]

[66]. Cohen MX and Cavanagh JF 2011 Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict Frontiers Psychol 2 1–12

[67]. Cavanagh JF and Frank MJ 2014 Frontal theta as a mechanism for cognitive control Trends Cogn. Sci 18 414–21 [PubMed: 24835663]

[68]. Greicius MD, Krasnow B, Reiss AL and Menon V 2003 Functional connectivity in the resting brain: a network analysis of the default mode hypothesis Proc. Natl Acad. Sci. USA 100 253–8 [PubMed: 12506194]

[69]. Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA and Shulman GL 2001 A default mode of brain function Proc. Natl Acad. Sci. USA 98 676–82 [PubMed: 11209064]

[70]. Tyagi H et al. 2019 Archival report a randomized trial directly comparing ventral capsule and anteromedial subthalamic nucleus stimulation in obsessive-compulsive disorder : clinical and imaging evidence for dissociable effects Biol. Psychiatry 85 726–34 [PubMed: 30853111]

[71]. Swann NC et al. 2018 Chronic multisite brain recordings from a totally implantable bidirectional neural interface: experience in 5 patients with Parkinson's disease J. Neurosurgery 128 605–16

[72]. Herron JA, Thompson MC, Brown T, Chizeck HJ, Ojemann JG and Ko AL 2017 Cortical brain–computer interface for closed-loop deep brain stimulation IEEE Trans. Neural Syst. Rehabil. Eng 25 2180–7 [PubMed: 28541211]

[73]. Ramirez-Zamora A. et al. 2018; Evolving applications, technological challenges and future opportunities in neuromodulation: Proceedings of the fifth annual deep brain stimulation think tank. Frontiers Neurosci. 11:734.

[74]. Stanslaski S, Afshar P, Cong P, Giftakis J, Stypulkowski P, Carlson D, Linde D, Ullestad D, Avestruz AT and Denison T 2012 Design and validation of a fully implantable, chronic, closed-loop neuromodulation device with concurrent sensing and stimulation IEEE Trans. Neural Syst. Rehabil. Eng 20 410–21 [PubMed: 22275720]

[75]. Lo MC and Widge AS 2017 Closed-loop neuromodulation systems: next-generation treatments for psychiatric illness Int. Rev. Psychiatry 29 191–204 [PubMed: 28523978]

[76]. Basu I et al. 2019 Brain stimulation consistent linear and non-linear responses to invasive electrical brain stimulation across individuals and primate species with implanted electrodes Brain Stimul 12 877–92 [PubMed: 30904423]

[77]. Sani OG, Yang Y, Lee MB, Dawes HE, Chang EF and Shanechi MM 2018 Mood variations decoded from multi-site intracranial human brain activity Nature Biotech 36 954–61

[78]. Michelson NJ, Vazquez AL, Eles JR, Salatino JW, Purcell EK, Williams JJ, Cui XT and Kozai TDY 2018 Multi-scale, multi-modal analysis uncovers complex relationship at the brain tissue-implant neural interface: new emphasis on the biological interface J. Neural Eng 15 033001 [PubMed: 29182149]

[79]. Sellers EW, Vaughan TM and Wolpaw JR 2010 A brain-computer interface for long-term independent home use Amyotroph. Lateral Scler 11 449–55 [PubMed: 20583947]

[80]. Zeitlin D, Wolpaw JR and Vaughan TM 2014 Brain–computer interface (BCI) evaluation in people with amyotrophic lateral sclerosis Amyotroph. Lateral Scler 15 207

[81]. Vansteensel MJ et al. 2016 Fully implanted brain–computer interface in a locked-in patient with ALS New Engl. J. Med 375 2060–6 [PubMed: 27959736]

[82]. Brandman DM, Cash SS and Hochberg LR 2017 Review: human intracortical recording and neural decoding for brain–computer interfaces IEEE Trans. Neural Syst. Rehabil. Eng 25 1687–96 [PubMed: 28278476]

[83]. Dangi S, Orsborn AL, Moorman HG and Carmena JM 2013 Design and analysis of closed-loop decoder adaptation algorithms for brain-machine interfaces Neural Comput 25 1693–731 [PubMed: 23607558]

[84]. Jarosiewicz B. Virtual typing by people with tetraplegia using a stabilized, self-calibrating intracortical brain-computer interface. IEEE Brain Gd. Challenges Conf; Washington, DC. 2014. 1–11.

[85]. Brandman D. et al. 2018; Rapid calibration of an intracortical brain computer interface for people with tetraplegia. J. Neural Eng. 15:026007. [PubMed: 29363625]

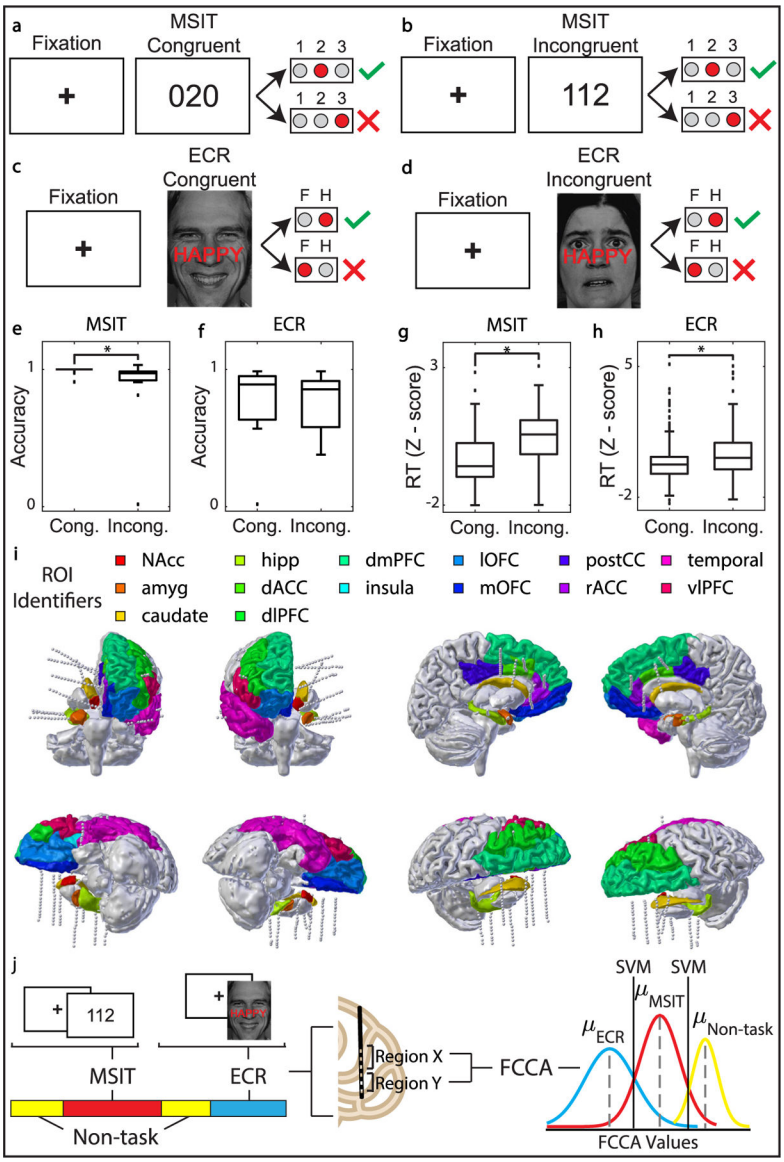

Figure 1.

Behavioral setup/outcomes, neural recording montage, and classification pipelines for real-time task detection. (A) MSIT congruent trial. The value of the target matches its position, and the value of the distractor, '0', is not a response option. (B) MSIT incongruent trial. The value of the target does not match its position, and the value of the distractor, '1', is a response option. (C) ECR congruent trial. The overlaid text, 'HAPPY', matches the emotion on the face. (D) ECR incongruent trial. The overlaid text, 'HAPPY', does not match the emotion on the face. (E) MSIT accuracy decreased during cognitive interference ($P = 7.2 * 10^{-3}$, Wilcoxon signed rank test). (F) ECR accuracy did not change during cognitive interference. (G) MSIT reaction time increased during cognitive interference ($P = 4.1 * 10^{-5}$, Wilcoxon signed rank test). (H) ECR reaction time increased during cognitive interference ($P = 2.1 * 10^{-3}$, Wilcoxon signed rank test). (I) Reconstruction of electrode placement in 13 regions of interest (ROIs) for Participant 14 (NAcc: nucleus accumbens, amyg: amygdala,

caudate, hipp: hippocampus, dACC: dorsal anterior cingulate cortex, dlPFC: dorsolateral prefrontal cortex, dlPFC: dorsomedial prefrontal cortex, insula, lOFC: lateral orbitofrontal cortex, mOFC: medial orbitofrontal cortex, postCC: posterior cingulate cortex, temporal lobe, vlPFC: ventral lateral prefrontal cortex). (J) Classification pipeline. Neural recording is labeled as either non-task, MSIT, or ECR. Functional connectivity between regions is assessed via fixed canonical correlation analysis (FCCA) and Fixed Canonical Coherence Analysis (FCHA). FCCA and FCHA features are computed across the entire recording on windowed neural data. A support vector machine (SVM) is trained to predict whether the participant was engaged in MSIT, ECR, or non-task.

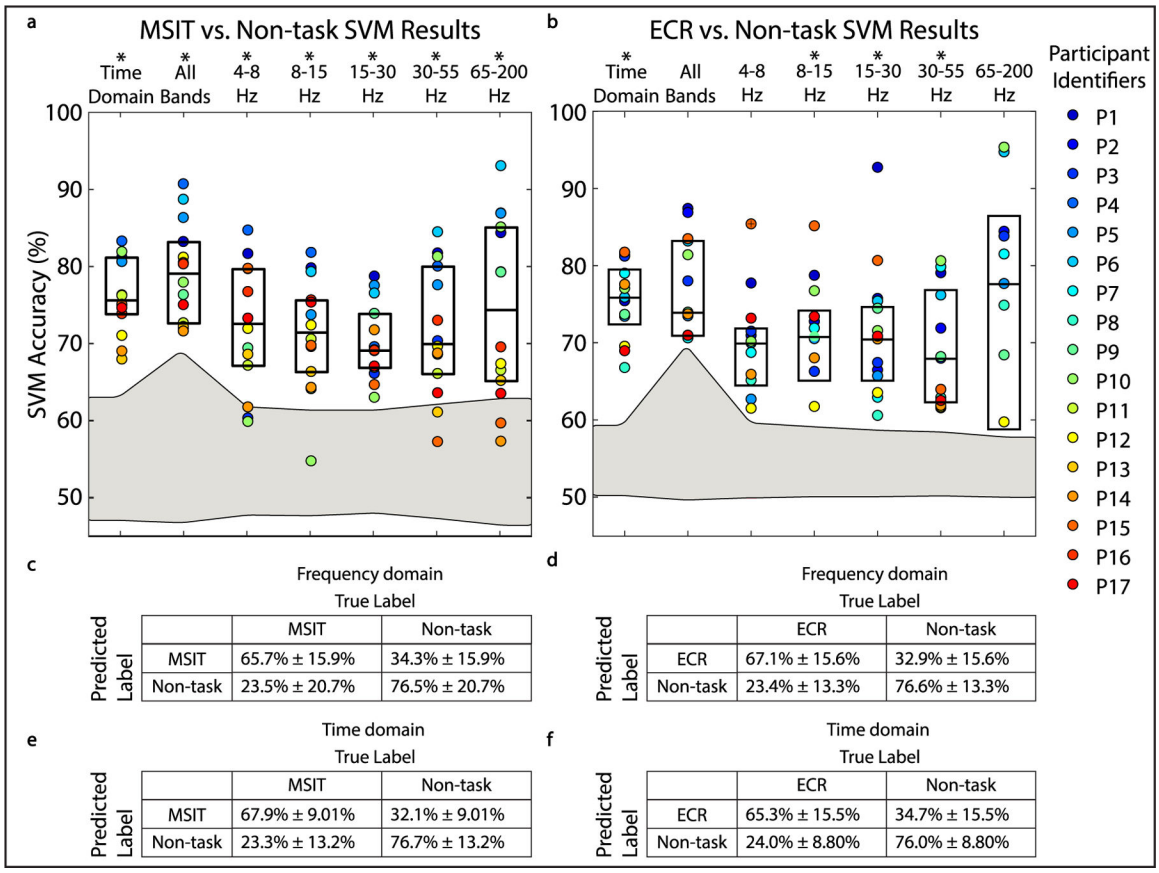

Figure 2.

Canonical correlation and coherence features enable task versus non-task classification for the MSIT and ECR task. (A) and (B) Boxplots show decoder performance in terms of accuracy for MSIT versus non-task classification and ECR versus non-task classification for true labels and shuffled labels. Colored points show individual participant results. Gray shaded area represents the confidence interval for chance classification performance. Asterisks indicate bands where classification significantly exceeded chance after correcting for multiple comparisons. SVM classifier performance was assessed by training the classifier with FCHA features from different frequency bands of interest (4–8 Hz, 8–15 Hz, 15–30, 30–55 Hz, 65–200 Hz), features from all the frequency bands of interest, and time-domain FCCA features. *P*-values from significance testing are included in tables S4 and S5. (C)–(F) Diagnostic tables show median true positive rate, false positive rate, true negative rate, and false negative rate for MSIT versus non-task classification and ECR versus non-task classification using FCHA features from all frequency bands (C) and (D) and FCCA (E) and (F).

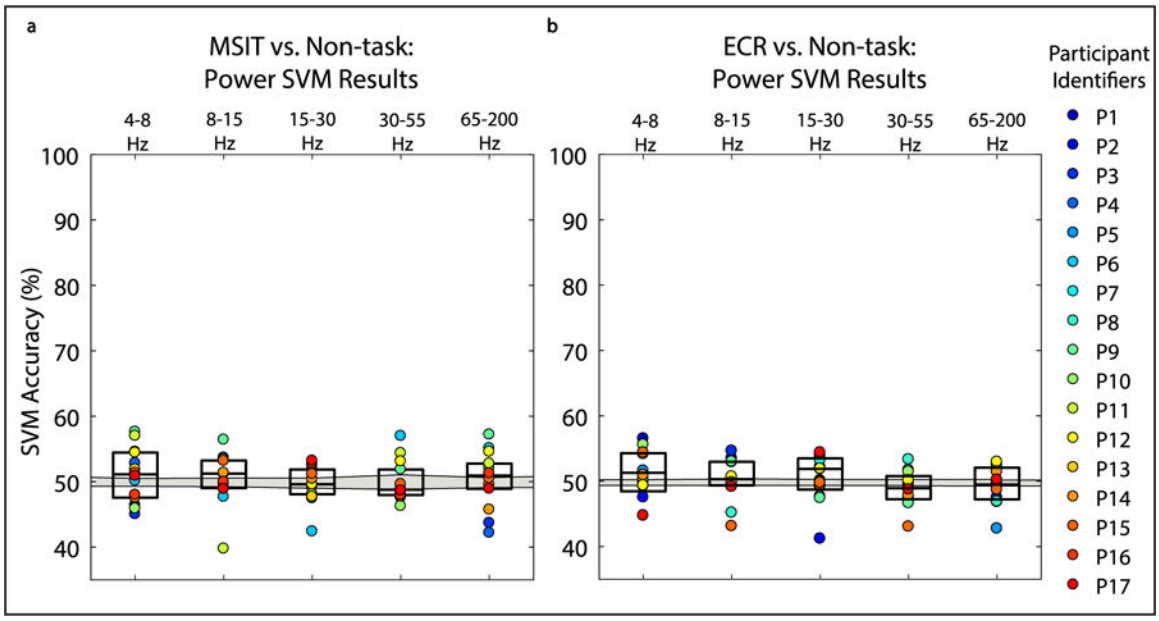

Figure 3.

Power-frequency feature inputs do not distinguish task versus non-task engagement for either MSIT or ECR. (A) and (B) Boxplots show decoder performance in terms of accuracy for MSIT versus non-task classification and ECR versus non-task classification for true labels and shuffled labels. Colored points show individual participant results. Gray shaded area represents the confidence interval for chance classification performance. Classification was not significantly different from chance after correcting for multiple comparisons. SVM classifier performance was assessed by training the classifier with power-frequency features at each individual electrode from different frequency bands of interest (4–8 Hz, 8–15 Hz, 15–30, 30–55 Hz, 65–200 Hz).

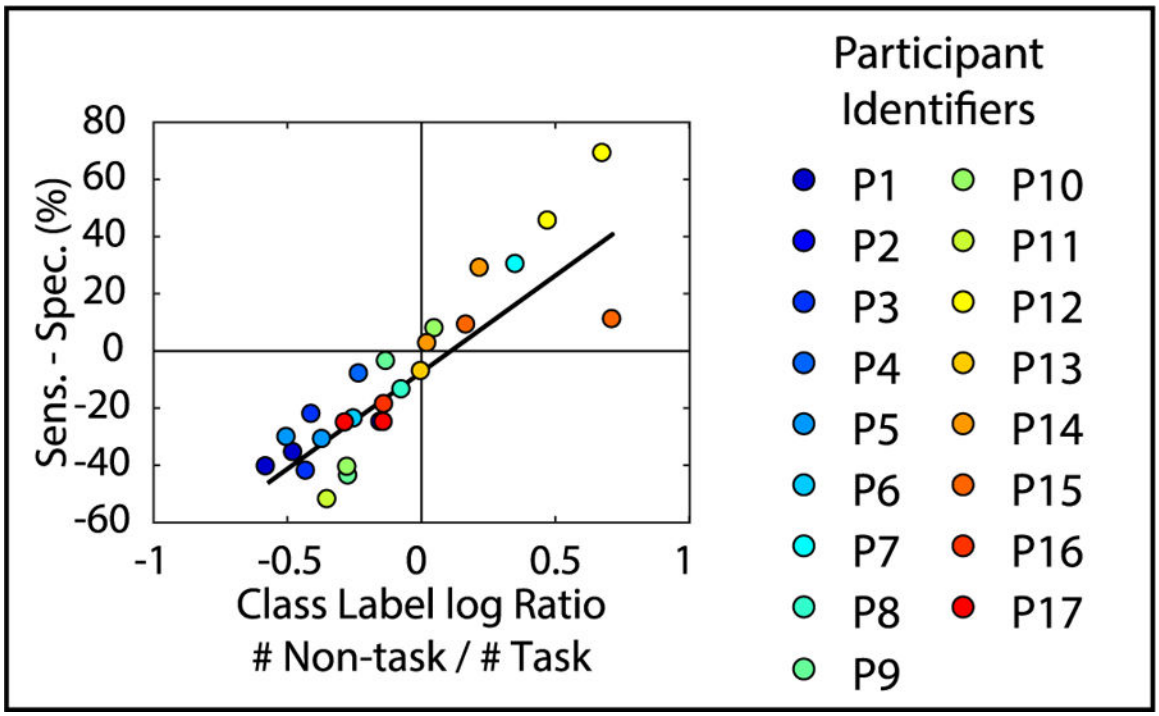

Figure 4.

Difference between specificity and sensitivity has a linear relationship with the ratio of task to non-task observations present in the original dataset. Difference between sensitivity and specificity for MSIT versus non-task and ECR versus non-task frequency domain classification results is plotted across participants against the ratio between the number of non-task to task observations present in the dataset before class balancing (linear regression: $R^2 = 0.72$).

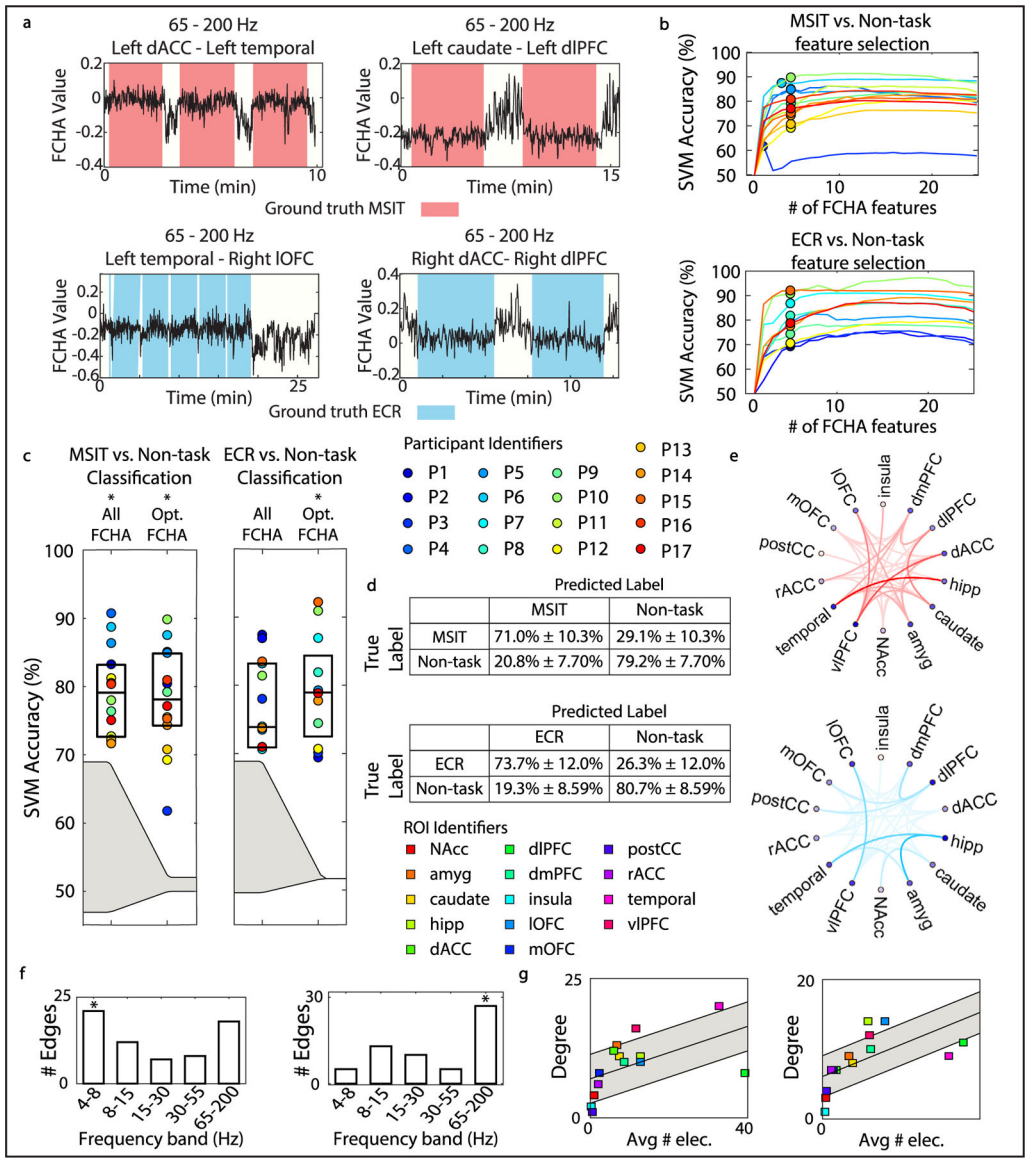

Figure 5.

Optimal feature selection maintains decoder performance across participants. (A) Example optimal FCHA feature traces are plotted over time. Ground truth (MSIT, ECR, and non-task labels) is indicated by background color on the plot. (B) Curves show SVM accuracy for individual participants as features are dropped. Up to five frequency domain features for each participant were used to achieve optimal decoder performance, indicated by points. (C) Boxplots show classification accuracy for MSIT versus non-task classification and ECR versus non-task classification for true labels and shuffled labels using all FCHA features and optimal FCHA features. Formatting is identical to figures 2(A) and (B). (D) Diagnostic tables show optimal classification performance. Formatting is identical to figures 2(C)–(F). (E) Optimal region pairs for MSIT versus non-task classification (red) and ECR versus non-task classification (blue) across participants are shown via connected nodes on the network graph. Opacity of the edges corresponds to the number of times each region pair occurred as

an optimal feature across participants. (F) Bar plot shows the number edges that corresponded to activity in each frequency band of interest for MSIT versus non-task classification (left) (4–8 Hz: $P = 7.6 * 10^{-3}$, Binomial test) and ECR versus non-task classification (right) (65–200 Hz: $P = 3.0 * 10^{-6}$, Binomial test). (G) Plot shows connectivity degree of each ROI node versus the average number of electrodes implanted in each ROI across participants for MSIT versus non-task classification (left) and ECR versus non-task classification (right). Gray shaded area shows one standard deviation above and below the line of best fit. Linear trend line is shown in black (left: $R^2 = 0.54$, right: $R^2 = 0.68$).

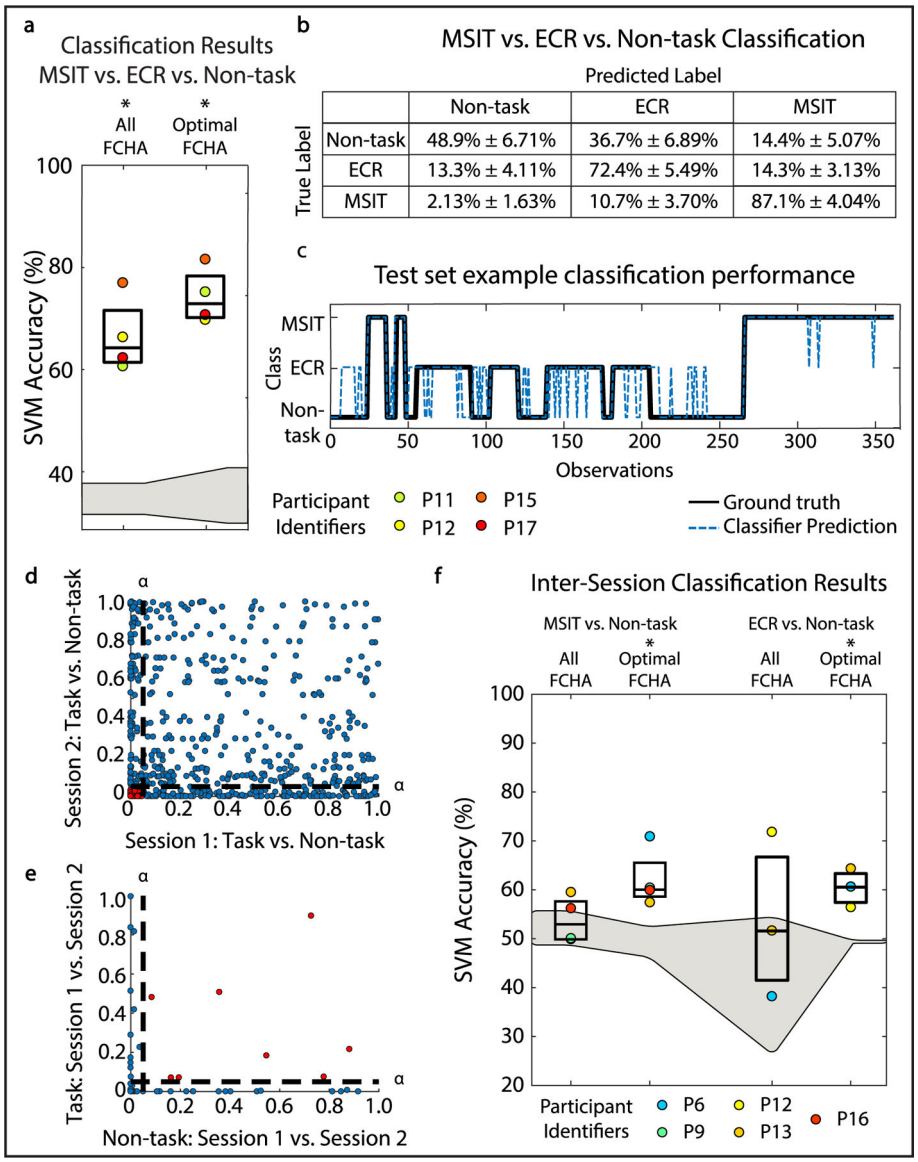

Figure 6.

Canonical coherence can distinguish between multiple task states and is stable over time. (A) Boxplots show decoder performance in terms of accuracy for MSIT versus ECR versus non-task classification using all FCHA features and up to 25 optimal FCHA features for true labels and shuffled labels. Formatting is identical to figures 2(A) and (B). (B) Diagnostic table shows classification performance for multi-task classification. Formatting is identical to figures 2(C)–(F). (C) Example multi-task classification performance on test set. Black solid line shows the ground truth state, while blue dotted line shows the classifier prediction. (D) *P*-value between task and non-task FCHA distributions from session two plotted against that of session one. Dotted lines indicate significance threshold, $\alpha = 0.05$. Points indicate individual features. Red points indicate feature distributions that are significantly different from null distributions in both session one and session two. Red points are carried over into panel (F). (E) *P*-value between task session one and task session two FCHA distributions,

plotted against the similar *P*-value for non-task FCHA distributions of the two sessions. Dotted lines indicate significance threshold, $\alpha = 0.05$. Red points indicate feature distributions that are not significantly different across sessions for task and non-task, and thus remain highly discriminating between sessions. (F) Inter-session classification results for MSIT versus non-task classification and ECR versus non-task classification using all features and optimal features. Formatting is identical to figures 2(A) and (B).