**SUPPLEMENTARY DATA**

**METHODS**

**Participants’ demography**

The study comprised 149 participants, including two pathological groups (neurological and cardiac) with a total of 77 patients, as well as 72 controls matched for sex, age, and education (Supplementary Data Tables 1 and 2).

Neurological group: multiple diseases

*Alzheimer’s disease (AD) patients.* A selection of 11 patients was made, all diagnosed following NINCDS-ADRDA criteria (1, 2). In order to be included they must have presented memory and language deficits and early atrophy in the temporal lobes, parietal regions (3, 4), and, in some cases, in the insular cortex (5). Patients who presented logopenic progresssive aphasia and atypical forms of AD (e.g. posterior cortical atrophy), were not included.

*Behavioral variant fronto-temporal dementia (bvFTD) patients.* A group of 9 patients with bvFTD were selected for the study. They were diagnosed following current revised criteria (6). Patients who gave signs of other forms of dementia (e.g. primary progressive aphasia and amyotrophic lateral sclerosis) were discarded.

*Multiple sclerosis (MS) patients.* The study comprised 25 relapsing–remitting MS patients in early disease stages (fulfilling the McDonald’s criteria (7)), diagnosed by two experts and complemented with a clinical standard examination, magnetic resonance imaging, and lumbar puncture, when necessary. Additionally, patients were assessed with the Expanded Disability Status Scale (EDSS) (8) and the Multiple Sclerosis Severity Score (MSSS) (9).

*Stroke (ST) patients.* Seven patients with non-hemorrhagic, fronto-insular lesions provoked from a brain stroke were selected. All were evaluated at least six months post-stroke (the time needed for stability of the lesion and presentation of clinical symptoms).

As exclusion criteria, all neurological patients were screened for cardiac and psychiatric primary conditions. The latter were assessed following WHO’s ICD-10 diagnostic guidelines (10).

Cardiac group: hypertensive disease patients

Participants selected for the cardiac group were chronic outpatients of the Metabolic and Arterial, Hypertension Unit of the Favaloro Foundation Hospital, and they were diagnosed following current revised criteria (11); they presented blood pressure within Grade 1, which defines the hypertension range (12). To confirm the patient’s condition at the time of evaluation, office blood pressure (OBP) readings and ambulatory blood pressure monitoring (ABPM) were measured. To discard secondary ailments, the patients were checked for cognitive impairments, lacunar infarcts, white or grey matter lesions, or psychiatric diseases. Psychiatric features were assessed following WHO’s ICD-10 diagnostic guidelines (10). All patients which presented relevant symptoms were excluded from the study. A total of 25 patients with HTD were included.

**Supplementary Table 1.** Demographic data for both patient groups.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Group** | *N* | **Age** | **Education** |  | **Sex** |
| **Mean** | *SD* | *P-value* | **Mean** | *SD* | *P-value* | **(M:F)** | *P-value* |
| *Neurological* | 52 | 55.13 | 17.43 | <.01\* | 14.71 | 4.56 | NS | (15:37) | NS |
| *Cardiac* | 25 | 66.16 | 9.06 | 15.40 | 3.91 | (12:13) |
| Demographic data for both patient groups, mean and standard deviation (*SD*), *P-value* for the t-test between groups for age and education, and Fisher exact two-tailed for sex. Demographic data (age and education) was assessed using one-way T-tests to compare groups with the STATISTICA data analysis software system (StatSoft, Inc. 2011, version 10. www.statsoft.com.). The sex comparison was done with Fisher exact two tailed test. |

**Supplementary Table 2.** Demographic matching for both patient groups in each dimension.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Dimension** | **Cardiac group** |  |  | **Neurological group** |  |  |
| **Age** | **Education** | **Sex (M:F)** | **Total N** | **Age** | **Education** | **Sex (M:F)** | **Total N** |
| **C** | **P** | ***P-val*** | **C** | **P** | ***P-val*** | **C** | **P** | ***P-val*** | **C** | **P** | **C** | **P** | ***P-val*** | **C** | **P** | ***P-val*** | **C** | **P** | ***P-val*** | ***C*** | ***P*** |
| *HBD task Accuracy* | 63.86 (8.62) | 66.16 (9.06) | NS | 16.84 (2.32) | 15.40 (3.91) | NS | 11:33 | 12:13 | NS | 44 | 25 | 55.10 (15.78) | 56.10 (16.95) | NS | 16.05 (2.62) | 14.72 (4.62) | NS | 13:37 | 16:46 | NS | 50 | 62 |
| *Meta-cognition* | 64.52 (9.24) | 68.50 (8.07) | NS | 16.83 (2.10) | 15.63 (3.95) | NS | 06:17 | 09:07 | NS | 23 | 16 | 52.47 (16.21) | 54.40 (17.36) | NS | 16.02 (2.47) | 14.94 (4.46) | NS | 12:35 | 14:36 | NS | 47 | 50 |
| *HEP* | 63.02 (9.17) | 66.16 (9.06) | NS | 16.53 (2.54) | 15.40 (3.91) | NS | 12:35 | 12:13 | NS | 47 | 25 | 55.10 (15.78) | 56.06 (17.17) | NS | 16.05 (2.62) | 14.80 (4.66) | NS | 16:46 | 13:36 | NS | 62 | 49 |
| *Brain volume* | 59.60 (12.64) | 65.26 (8.88) | NS | 16.78 (2.25) | 15.87 (3.60) | NS | 11:34 | 11:12 | NS | 45 | 23 | 53.73 (18.00) | 55.12 (18.6) | NS | 15.71 (2.92) | 14.40 (4.41) | NS | 17:34 | 11:31 | NS | 51 | 42 |
| *Brain connectivity* | 61.11 (10.57) | 65.26 (8.88) | NS | 16.92 (2.28) | 15.87 (3.60) | NS | 09:29 | 11:12 | NS | 38 | 23 | 54.31 (17.01) | 54.83 (17.98) | NS | 16.02 (2.68) | 14.54 (4.60) | NS | 13:32 | 11:30 | NS | 45 | 41 |
|  | Demographic matching for both patient's group in each dimension: heart-beat detection (HBD) task accuracy, metacognition task, heart-evoked potential (HEP), brain grey matter volume and brain connectivity. Mean and standard deviation values are shown for each group. *P-value*s are shown for each comparison: Fisher two-tailed test for Sex and T-tests for age and education. C: Controls, P: Patients. *P-val*: *P-value*s. NS: non-significant. |  |

**Data acquisition and preprocessing**

HBD task: Accuracy

From the HBD task we derived the following indexes that measure different aspects of interoceptive behavioral performance (accuracy).

*Response-count-based accuracy.* A modified version of Schandry’s index (13) was calculated as the difference between actual heartbeats (as registered by the ECG) and total motor responses (instead of counted heartbeats, as in Schandy’s formula). The procedure is captured by the following equation:

$$Response-count-based Accuracy = 1- \frac{\left(Recorded hearbets-\sum\_{}^{}Total responses\right)}{Recorded heartbeats}$$

Note that the coefficient is displayed as a precision score (and not as an error-score), with higher values indicating better performance.

*Delay-based accuracy.* Following previous reports (14-21), we calculated this interoceptive accuracy index that considers two components: correct answers (which were estimated by comparing each participant’s motor response relative to the time window of the corresponding heartbeat); and the total of recorded heartbeats during each condition. In order to control for individual differences in heart rate (HR), each tapped response was time-locked to the ECG R-peak based on three different time-windows: 750 ms after the beat for a (HR) less than 69.76; 600 ms after the beat for a HR between 69.75 and 94.25; and 400 ms after the beat for a HR higher than 94.25. The ensuing score can vary between 0 and 1, with higher values indicating only small differences between correct answers and recorded heartbeats and thus better interoceptive performance. Accuracy was calculated with the following equation:

$$Delay Based Accuracy = 1- \frac{\left(Recorded hearbets-\sum\_{}^{}Correct Answers\right)}{Recorded heartbeats}$$

*Mean distance accuracy.* This index evaluates the synchronization between heartbeats and associated motor responses (22). Within each condition, each block is subdivided in overlapping windows starting at each individual heartbeat and extending for 10 seconds. Then, for each window, the absolute difference between cardiac frequency (CF, measured as 1/mean R-R) and response frequency (1/mean inter response intervals) is computed. The described process is represented in the following equation:

$$d\_{m,w }= \frac{\sum\_{i=1}^{N}|f\_{c}\_{i,w}-f\_{r}\_{i,w}|}{N}$$

Where CF is the average cardiac frequency in a window of w duration centered at time i, fr is the average response frequency in the same window and time, and N is the number of heartbeats in the block. To control for possible periods of the task where participants may have lost concentration, a coefficient of variation (CV) is estimated for the inter-response intervals inside each individual window. Only windows with a CV bigger than 1 are included to calculate the precision score. Finally, the absolute differences between cardiac and response frequencies are averaged across all the windows comprising each block of each condition. The minimum score for Dm index is 0, indicating a perfect match between motor responses and cardiac frequencies, with higher scores indicating higher distances, and thus, worse performance.

*D-prime accuracy.* This score is based on signal detection theory (SDT), a framework that allows distinguishing ambiguous stimuli as signal or noise (23, 24). In the context of the HBDT, a heartbeat is considered a signal, while its absence represents noise. Accordingly, participants’ responses are tagged as a ‘yes’ (when pressing the keyboard, a hit) or as a ‘no’ (when omitting, a miss). An affirmative response can only be considered correct if it occurs in a given window time-locked to the R-wave of the preceding heartbeat (the signal). Therefore, the absence of a response outside the window is a *correct rejection,* while a response outside the window is a *false alarm*.

The temporal extension of the window was determined for each participant according to his/her HR to control for potential inter-individual differences. The window was locked 750 ms after the beat for a HR below than 69.76; 600 ms after the beat for a HR between 69.75 and 94.25; and 400 ms after the beat for a HR higher than 94.25(25). We calculated the d’ index with the following equation:

$$d'=z\left(\frac{∑hit}{∑hit+∑miss}\right)-z\left(\frac{∑false alarm}{∑false alarm+∑correct rejection}\right)$$

Where z (p) is the inverse normal probability corresponding to cumulative probability. This formula weighs the strategy of the participant as it discriminates signal from noise, in order to eliminate successes by chance (e.g. a participant who always responds “yes” would get all the *hits)*. Obtaining higher values of d’ indicate better discrimination ability and thus better interoceptive accuracy.

HBD task: EEG-HEP

**Supplementary Table 3.** Mean and standard deviation number of epochs for each condition

|  |  |  |  |
| --- | --- | --- | --- |
| **Group** | **Interoceptive Accuracy** | **Post-feedback** | ***P-val*** |
|  | **Mean** | ***SD*** | **Mean** | ***SD*** |
| *Neurological* | 198.06 | 67.40 | 195.33 | 63.75 | NS |
| *Cardiac* | 199.80 | 59.36 | 188.20 | 67.52 | NS |
| Mean and standard deviation number of epochs for each condition (interoceptive accuracy, and post-feedback), and for each pathological group. An ANOVA factorial test was performed considering both conditions and their control group. |
|

Neuroimaging

**Supplementary Table 4.** Translation (right, forward, up) and rotation (pitch, yaw, roll) movement detail

|  |  |  |  |
| --- | --- | --- | --- |
| **Group** | **N** | **Translation (mm)** | **Rotation (°)** |
|  |  | **Mean** | ***SD*** | **Mean** | ***SD*** |
| *Neurological* | 41 | 0.11 | 0.07 | 0.10 | 0.09 |
| *Cardiac* | 23 | 0.10 | 0.05 | 0.07 | 0.06 |
| Mean and standard deviation translation (right, forward, up) and rotation (pitch, yaw, roll) movement for each pathological group. |

Classification data analysis

*Data normalization.* To avoid bias in sample variance introduced by outlier values, we first eliminated those that were 2.5 standard deviations above or below within each group and variable average score. Concerning the age gap between the two target groups (Supplementary Data Table 1), all scores of all levels of analysis were weighed by age (after outliers were eliminated), as done in previous research (26). To do this, we tested and ensured the validity of using a linear model with our data. Then, the normalization process was performed based on a multivariate linear model for each variable adjusted for the reference data (control group scores). Next, the difference between the dependent variable’s values and their corresponding prediction (given by the adjusted model) was calculated for the sample of testing data (the patient’s group score). To normalize these differences, their values were divided by the variability of the lineal model in an analogous way as in the standard calculus of z value. This procedure allows to compare the performance of the dependent variable by controlling for the influence of others independent variables. A weighted z-score close to zero corresponded to a performance similar to the base of the reference data, and lower/higher values corresponds to lower/higher values respect to base of the reference data. Following this procedure, we obtained weighted scores for each patient which also allowed us to compare both pathological groups in a classification analysis, since they both have different control group.

*Feature selection.* In order to choose the best features for each level of analysis, an attribute selection analysis was performed within each one. This analysis was performed with the Waikato Environment for Knowledge Analysis (WEKA) http://www.cs.waikato.ac.nz/ml/weka suite of ML software (27, 28), as done in previous research (29), using the “SVMAtributteSelection” function (an attribute evaluator). This function works by evaluating the importance of an attribute by using an SVM classifier. Attributes (variables) are ranked by the square of the weight assigned by the SVM (30). The filtering type used by the SVM was to normalize training data. The tolerance Parameter to pass to the SVM was 1.0E-10, the complexity parameter to pass to the SVM was 1 and the constant rate of attributes eliminated per iteration was 1, epsilon parameter to pass to the SVM was 1.0E-25 (as set in default). Feature search was performed with the ranker method, which ranks attributes by their individual evaluations in conjunction with attribute evaluators. The evaluation mode was 10-fold cross-validation.

*Classification methods.* Two models were run for classification analysis. The first one compared the two target pathologies (neurological vs cardiological), whereas the second one looked within the neurological group, testing between neurological diseases (AD, MS, ST and bvFTD). For the first analysis our objective was to determine which dimension (or combination) classified better between neurological pathologies in general, and another group of common old age disease (hypertension).

For the second analysis our intention was to analyze whether these same dimensions could differentiate between specific neurological diagnosis. We compared each pathological group with the others (OT) together. In the cases of AD, FTD and ST there was an imbalance between the OT and the target group. To balance the total of cases for both groups, we repeated each classification analysis (only for those target pathologies) 10 times with different randomized sub-samples of the OT group that were of the same size as the target group. In these cases, all classification results are the mean value of all the sub-sample repeated analysis.

Using the PredPsych package of R (31), we used the cross-validated Linear Discriminant Analysis function (LDA), to perform cross-validated Linear Discriminant Analysis. A 5-fold cross-validation analysis was performed for training and testing dataset, with a 70% fraction of data to keep for training data. These functions use an algorithm for classification-based analyses. Linear Discriminant Analysis works by building a model composed of several discriminant functions based on linear combinations of data features that provide the best discrimination between two or more conditions/classes. The goal of the statistical analysis is thus to combine the data features scores in a way that a single new composite variable, the discriminant function, is produced (for details see (32, 33)). All missing data were replaced with the mean value for each variable.

To determine whether the classification accuracy was biased by the sample size difference between the pathological groups, we decided to repeat the classification analysis 10 times with different randomized sub-samples of the neurological group that were of the same size as the cardiological group (25 patients).

*Feature relevance.* To clarify which dimension contributed most to both model’s classification accuracy, we performed a feature relevance analysis. To do this, we repeated the classification analysis that was performed by omitting one dimension, for each of the five features in the classification. If the resulting accuracy score decreased without a given feature, the feature can be considered relevant, conversely if these values increased, the feature can be considered to introduce unwanted noise. The differences in accuracy between the model with all the variables and each model omitting one dimensions determined the merit of each feature (the higher difference, the lower merit). Finally, classification was repeated by taking out one by one the features following the merit order (For further details see Suplementary Data 1.3.4.)

**RESULTS**

**Supplementary Table 5.** Feature selection.

|  |  |
| --- | --- |
| **Attribute** | **Average rank**  |
| 1. *Mean distance accuracy interoceptive accuracy condition*
 | 2.7 (1.19) |
| 1. *D-prime accuracy post condition*
 | 3.1 (1.51) |
| 1. *Delay-based accuracy interoceptive accuracy condition*
 | 3.6 (1.91) |
| 1. *Response-count-based accuracy interoceptive accuracy condition*
 | 4.1 (2.91) |
| 1. *Mean distance post condition*
 | 4.5 (2.38) |
| 1. *D-prime interoceptive accuracy condition*
 | 5 (1.26) |
| 1. *Delay-based post condition*
 | 5.6 (1.36) |
| 1. *Response-count-based accuracy post condition*
 | 7.4 (1.02) |
| Mean accuracy and standard deviation for attribute ranking in HBDT accuracy dimension. |
|

**Supplementary Table 6.** Attribute ranking in the metacognition dimension.

|  |  |
| --- | --- |
| **Attribute** | **Average rank**  |
| 1. *Mean Interoceptive sensibility for post-feedback condition*
 | 1 (0) |
| 1. *Interoceptive awareness for post-feedback condition with D-prime accuracy*
 | 2.3 (0.46) |
| 1. *Interoceptive awareness for post-feedback condition with Mean distance accuracy*
 | 2.7 (0.46) |
| 1. *Interoceptive awareness for interoceptive accuracy condition with D-prime accuracy*
 | 4.7 (0.9) |
| 1. *Interoceptive awareness for post condition with Delay-based accuracy*
 | 5.5 (1.02) |
| 1. *Interoceptive awareness for interoceptive accuracy condition with Mean distance accuracy*
 | 6.1 (1.37) |
| 1. *Interoceptive awareness for post-feedback condition with* *Response-count-based accuracy*
 | 7 (2.05) |
| 1. *Interoceptive awareness for interoceptive accuracy condition with Response-count-based accuracy*
 | 8.1 (0.94) |
| 1. *Interoceptive awareness for interoceptive accuracy condition with Delay-based accuracy*
 | 8.3 (1.1) |
| 1. *Mean Interoceptive sensibility for interoceptive accuracy condition*
 | 9.3 (1.19) |
| Mean accuracy and standard deviation for attribute ranking in the metacognition dimension. |
|

**Supplementary Table 7.** Attribute ranking in HEP dimension.

|  |  |
| --- | --- |
| **Attribute** | **Average rank**  |
| 1. *Central ROI in 200-300 window, interoceptive accuracy condition*
 | 1.2 (0.4) |
| 1. *Right ROI in 200-300 window, interoceptive accuracy condition*
 | 1.8 (0.4) |
| 1. *Central ROI in 200-300 window, pos-feedback condition*
 | 4.4 (1.43) |
| 1. *Right ROI in 300-400 window, post-feedback condition*
 | 5.4 (1.85) |
| 1. *Central ROI in 300-400 window, interoceptive accuracy condition*
 | 6.1 (2.07) |
| 1. *Right ROI in 200-300 window, post-feedback condition*
 | 6.7 (2.37) |
| 1. *Left ROI in 200-300 window, post-feedback condition*
 | 7.6 (2.5) |
| 1. *Left ROI in 300-400 window, interoceptive accuracy condition*
 | 8.3 (2.76) |
| 1. *Left ROI in 200-300 window, interoceptive accuracy condition*
 | 8.6 (2.76) |
| 1. *Right ROI in 300-400 window, interoceptive accuracy condition*
 | 8.8 (3.31) |
| 1. *Central ROI in 300-400 window, post-feedback condition*
 | 9 (1.84) |
| 1. *Left ROI in 300-400 window, post-feedback condition*
 | 10.1 (1.51) |
| Mean accuracy and standard deviation for attribute ranking in HEP dimension. |
|

**Supplementary Table 8.** Attribute ranking in brain grey-matter volume dimension

|  |  |
| --- | --- |
| **Attribute** | **Average rank** |
| 1. *Right Four ROIs*
 | * 1. (1.14)
 |
| 1. *Right anterior cingulate gyrus*
 | 3.7 (2) |
| 1. *Left anterior cingulate gyrus*
 | 3.7 (1.19) |
| 1. *Bilateral midcingulate área*
 | 5.5 (3.67) |
| 1. *Right insula*
 | 6.6 (3.29) |
| 1. *Left Four ROIs*
 | 6.8 (3.22) |
| 1. *Bilateral insula*
 | 6.9 (3.18) |
| 1. *Right midcingulate area*
 | 8.8 (2.75) |
| 1. *Left insula*
 | 9.1 (2.12) |
| 1. *Left midcingulate area*
 | 9.2 (4.96) |
| 1. *Bilateral Four ROIs*
 | 9.3 (3.74) |
| 1. *Bilateral anterior cingulate gyrus*
 | 11.3 (2.1) |
| 1. *Left postcentral gyrus*
 | 11.6 (3.1) |
| 1. *Right postcentral gyrus*
 | 11.8 (2.18) |
| 1. *Bilateral postcentral gyrus*
 | 13.6 (1.36) |
| Mean accuracy and standard deviation for attribute ranking in brain grey-matter volume dimension. |
|

**Supplementary Table 9**. Attribute ranking in brain connectivity dimension.

|  |  |
| --- | --- |
| **Attribute** | **Average rank** |
| 1. *Left anterior cingulate gyrus vs. Right postcentral gyrus*
 | 1 (0) |
| 1. *Right midcingulate area vs. Right postcentral gyrus*
 | 2.3 (0.46) |
| 1. *Left midcingulate area vs. Left postcentral gyrus*
 | 3.7 (2.05) |
| 1. *Left insula vs. Left anterior cingulate gyrus*
 | 5.1 (1.14) |
| 1. *Left insula vs. Right midcingulate area*
 | 7.6 (3.38) |
| 1. *Left insula vs. Right anterior cingulate gyrus*
 | 8.6 (4.45) |
| 1. *Right insula vs. Left midcingulate area*
 | 9.2 (2.93) |
| 1. *Left insula vs. Left midcingulate area*
 | 11.1 (6.49) |
| 1. *Left insula vs. Left postcentral gyrus*
 | 12 (6.02) |
| 1. *Right insula vs. Right anterior cingulate gyrus*
 | 12.9 (5.97) |
| 1. *Right anterior cingulate gyrus vs. Right postcentral gyrus*
 | 13.4 (6.86) |
| 1. *Left midcingulate area vs. Right midcingulate area*
 | 13.8 (5.6) |
| 1. *Right midcingulate area vs. Left postcentral gyrus*
 | 14 (5.37) |
| 1. *Right anterior cingulate gyrus vs. Right midcingulate area*
 | 15.5 (3.91) |
| 1. *Left anterior cingulate gyrus vs. Right anterior cingulate gyrus*
 | 15.7 (6.72) |
| 1. *Right insula vs. Right postcentral gyrus*
 | 16.8 (6.52) |
| 1. *Right anterior cingulate gyrus vs. Left postcentral gyrus*
 | 17.1 (4.89) |
| 1. *Right anterior cingulate gyrus vs. Left midcingulate area*
 | 18.4 (7.36) |
| 1. *Left insula vs. Right postcentral gyrus*
 | 18.9 (6.14) |
| 1. *Left midcingulate area vs. Right postcentral gyrus*
 | 19.2 (5.34) |
| 1. *Left postcentral gyrus vs. Right postcentral gyrus*
 | 19.6 (3.9) |
| 1. *Left insula vs. Right insula*
 | 19.7 (6.25) |
| 1. *Left anterior cingulate gyrus vs. Left midcingulate area*
 | 19.9 (4.78) |
| 1. *Left anterior cingulate gyrus vs. Left postcentral gyrus*
 | 20.9 (4.28) |
| 1. *Left anterior cingulate gyrus vs. Right midcingulate area*
 | 21.8 (7.18) |
| 1. *Right insula vs. Left anterior cingulate gyrus*
 | 22 (5.2) |
| 1. *Right insula vs. Left postcentral gyrus*
 | 22 (4.24) |
| 1. *Right insula vs. Right midcingulate area*
 | 23.8 (2.32) |
| Mean accuracy and standard deviation for attribute ranking in brain connectivity dimension. |
|

**Classification**

To determine whether the classification accuracy was biased by the sample size difference between the pathological groups, we decided to repeat the classification analysis 10 times with different randomized sub-samples of the neurological group that are of the same size as the cardiological group (25 patients). To perform this, we took random subsets of the participants keeping the same proportion of the original sample. This procedure was implemented using random permutations in Matlab software, obtaining subsets of the original sample. Afterwards LDA analysis was repeated for each sub-group. Overall classification accuracy was maintained when using different sub-samples of the neurological group with a mean 0.79 classification accuracy (0.07 standard deviation), 0.83 specificity (0.03 standard deviation), and 0.78 sensitivity (0.11 standard deviation).

**Control of motor and attentional deficits in pathological samples**

To control for potential motor and/or attentional deficits that could affect HBD task performance in the patient groups, we compared the performance of neurological and cardiac groups relative to their matched controls in the control (feedback) condition. These analyses we based on the mean distance index and sensibility (confidence) for the feedback condition, namely, the control variables of those that were first ranked by model 1 in the interoceptive accuracy and metacognitive dimensions, respectively. Two-tailed unpaired *t*-tests revealed no significant differences between patients and controls in any variable (Supplementary Table 10).

**Supplementary Table 10.** Behavioral performance in HBD task control variables.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable (feedback condition)** | **Group** | **Mean (*SD*)** | ***N*** | ***T*** | ***p*** |
| Mean distance accuracy | *Neurological* | 0.27 (0.31) | 50 | 0.38 | 0.70 |
| *Control* | 0.25 (0.30) | 61 |
| *Cardiac* | 0.24 (0.32) | 25 | 0.80 | 0.42 |
| *Control* | 0.18 (0.25) | 43 |
| Sensibility (confidence) | *Neurological* | 84.81 (17.96) | 46 | 1.74 | 0.09 |
| *Control* | 77.54 (22.28) | 49 |
| *Cardiac* | 85.42 (13.89) | 16 | 0.20 | 0.85 |
| *Control* | 84.34 (19.89) | 22 |
|  |

**Feature relevance**

**Supplementary Table 11.** Linear Discriminant Analysis classification feature relevance for the first model.

|  |  |  |  |
| --- | --- | --- | --- |
| **Dimension omitted** | **Accuracy** | **Specificity** | **Sensitivity** |
| *None* | 0.81 | 0.79 | 0.92 |
| *Metacognition* | 0.71 | 0.73 | 0.62 |
| *HBD task Accuracy* | 0.73 | 0.75 | 0.67 |
| *HEP* | 0.80 | 0.77 | 1.00 |
| *Brain volume* | 0.81 | 0.78 | 1.00 |
| *Brain connectivity* | 0.83 | 0.82 | 0.83 |
| **Dimension included** | **Accuracy** | **Specificity** | **Sensitivity** |
| *Metacognition* | 0.69 | 0.71 | 0.50 |
| *HBD task Accuracy* | 0.80 | 0.77 | 1.00 |
| *HEP* | 0.83 | 0.83 | 1.00 |
|  *Brain volume* | 0.83 | 0.82 | 0.83 |
| *Brain connectivity* | 0.81 | 0.79 | 0.92 |
| Linear Discriminant Analysis classification accuracy, specificity and sensitive values for the first model: by omitting only one variable or adding one by one.  |
|

**Supplementary Table 12.** Linear Discriminant Analysis classification accuracy for the second model.

|  |  |  |  |
| --- | --- | --- | --- |
| **Comparison** | **Accuracy** | **Specificity** | **Sensitivity** |
| AD vs. OT | 0.40 | 0.47 | 0.35 |
| bvFTD vs. OT | 0.50 | 0.49 | 0.50 |
| ST vs. OT | 0.69 | 0.71 | 0.72 |
| MS vs. OT | 0.53 | 0.59 | 0.48 |
| Mean | 0.53 | 0.56 | 0.52 |
| Linear Discriminant Analysis classification accuracy. specificity and sensitive values for each model adding education as sixth dimension. |
|
|

**Supplementary Table 13 and 14.** Linear Discriminant Analysis classification feature relevance for the second model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Comparison** | **Dimension omitted** | **Accuracy** | **Specificity** | **Sensitivity** |
| AD vs OT | *None* | 0.40 | 0.47 | 0.35 |
| bvFTD vs OT | 0.50 | 0.49 | 0.50 |
| ST vs OT | 0.69 | 0.71 | 0.72 |
| MS vs OT | 0.53 | 0.59 | 0.48 |
| AD vs OT | *HBDT Accuracy* | 0.37 | 0.51 | 0.31 |
| bvFTD vs OT | 0.62 | 0.71 | 0.53 |
| ST vs OT | 0.61 | 0.25 | 0.70 |
| MS vs OT | 0.33 | 0.39 | 0.25 |
| AD vs OT | *HEP* | 0.36 | 0.47 | 0.14 |
| bvFTD vs OT | 0.54 | 0.61 | 0.46 |
| ST vs OT | 0.74 | 0.48 | 0.79 |
| MS vs OT | 0.56 | 0.65 | 0.51 |
| AD vs OT | *Metacognition* | 0.34 | 0.47 | 0.16 |
| bvFTD vs OT | 0.52 | 0.65 | 0.43 |
| ST vs OT | 0.79 | 0.64 | 0.80 |
| MS vs OT | 0.53 | 0.59 | 0.48 |
| AD vs OT | *Brain volume* | 0.40 | 0.51 | 0.25 |
| bvFTD vs OT | 0.5 | 0.64 | 0.42 |
| ST vs OT | 0.77 | 0.59 | 0.82 |
| MS vs OT | 0.47 | 0.52 | 0.43 |
| AD vs OT | *Brain connectivity* | 0.34 | 0.45 | 0.15 |
| bvFTD vs OT | 0.46 | 0.50 | 0.41 |
| ST vs OT | 0.77 | 0.68 | 0.80 |
| MS vs OT | 0.64 | 0.78 | 0.57 |
| Linear Discriminant Analysis classification mean accuracy, specificity and sensitive values for the second model: by omitting only one variable. For the AD, FTD and ST groups a mean of the 10 subsamples is reported. |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Comparison** | **Dimension included** | **Accuracy** | **Specificity** | **Sensitivity** |
| AD vs. OT | *Metacognition* | 0.45 | 0.58 | 0.39 |
| *Brain connectivity* | 0.40 | 0.58 | 0.39 |
| *HBDT Accuracy* | 0.42 | 0.52 | 0.25 |
| *Brain volume* | 0.36 | 0.47 | 0.14 |
| *HEP* | 0.40 | 0.47 | 0.35 |
| bvFTD vs. OT | *Brain connectivity* | 0.69 | 0.76 | 0.62 |
| *Brain volume* | 0.68 | 0.77 | 0.58 |
| *Metacognition* | 0.67 | 0.77 | 0.57 |
| *HEP* | 0.62 | 0.71 | 0.53 |
| *HBDT Accuracy* | 0.50 | 0.49 | 0.50 |
| ST vs. OT | *Brain volume* | 0.79 | 0.96 | 0.79 |
| *HBDT Accuracy* | 0.94 | 1.00 | 0.93 |
| *HEP* | 0.80 | 0.74 | 0.81 |
| *Brain connectivity* | 0.79 | 0.64 | 0.8 |
| *Metacognition* | 0.69 | 0.71 | 0.72 |
| MS vs. OT | *HBDT Accuracy* | 0.64 | 0.75 | 0.57 |
| *Brain volume* | 0.67 | 0.83 | 0.59 |
| *Metacognition* | 0.64 | 0.78 | 0.57 |
| *HEP* | 0.64 | 0.78 | 0.57 |
| *Brain connectivity* | 0.53 | 0.59 | 0.48 |
| Linear Discriminant Analysis classification accuracy, specificity, and sensitive values for the first model: by adding one by one. For the AD. FTD and ST groups a mean of the 10 subsamples is reported. |
|
|

Even though there were no differences between pathological groups in educational years (or between each group and its control group) we wanted to test whether it would have an importance in the model’s classification. We repeated the classification analysis for each model but adding education as a sixth variable. As seen in the table adding the education variable to the model only lowers the accuracy for both models.

**Supplementary Table 15.** Linear Discriminant Analysis classification detail for adding education as sixth dimension.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Dimension included** | **Accuracy** | **Specificity** | **Sensitivity** | **Difference** |
| NEU vs. CAR | *All* | 0.81 | 0.79 | 0.92 | 0.00 |
| *All+Education* | 0.80 | 0.79 | 0.86 | 0.01 |
| Linear Discriminant Analysis classification accuracy, specificity and sensitive values for each model adding education as sixth dimension. |
|

**Neuroimaging feature relevance**

In order to determine the relevance of the neuroimaging features (brain volume and connectivity) separated from the other 3 features considered, we repeated the Linear Discriminant Analysis classification only with them but comparing each pathological group separately with the cardiological one. The comparison between the AD group and the cardiological group yielded the bigger classification accuracy (0.86), followed by ST (0.73) and bvFTD (0.74) and lastly MS (0.58). This is consistent with the level and location of atrophy for each neurological disease.

**REFERENCES**

1. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack Jr CR, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R. The diagnosis of dementia due to Alzheimer’s disease: Recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & dementia. 2011;7(3):263-9.

2. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS‐ADRDA Work Group\* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology. 1984;34(7):939-.

3. Braak H, Braak E. Neuropathological stageing of Alzheimer-related changes. Acta neuropathologica. 1991;82(4):239-59.

4. Naggara O, Oppenheim C, Rieu D, Raoux N, Rodrigo S, Dalla Barba G, Meder J-F. Diffusion tensor imaging in early Alzheimer's disease. Psychiatry Research: Neuroimaging. 2006;146(3):243-9.

5. Bonthius DJ, Solodkin A, Van Hoesen GW. Pathology of the insular cortex in Alzheimer disease depends on cortical architecture. Journal of neuropathology and experimental neurology. 2005;64(10):910-22.

6. Rascovsky K, Hodges JR, Knopman D, Mendez MF, Kramer JH, Neuhaus J, van Swieten JC, Seelaar H, Dopper EG, Onyike CU, Hillis AE, Josephs KA, Boeve BF, Kertesz A, Seeley WW, Rankin KP, Johnson JK, Gorno-Tempini ML, Rosen H, Prioleau-Latham CE, Lee A, Kipps CM, Lillo P, Piguet O, Rohrer JD, Rossor MN, Warren JD, Fox NC, Galasko D, Salmon DP, Black SE, Mesulam M, Weintraub S, Dickerson BC, Diehl-Schmid J, Pasquier F, Deramecourt V, Lebert F, Pijnenburg Y, Chow TW, Manes F, Grafman J, Cappa SF, Freedman M, Grossman M, Miller BL. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. Brain : a journal of neurology. 2011;134(Pt 9):2456-77.

7. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, Fujihara K, Havrdova E, Hutchinson M, Kappos L, Lublin FD, Montalban X, O'Connor P, Sandberg-Wollheim M, Thompson AJ, Waubant E, Weinshenker B, Wolinsky JS. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Annals of neurology. 2011;69(2):292-302.

8. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology. 1983;33(11):1444-52.

9. Roxburgh RH, Seaman SR, Masterman T, Hensiek AE, Sawcer SJ, Vukusic S, Achiti I, Confavreux C, Coustans M, le Page E, Edan G, McDonnell GV, Hawkins S, Trojano M, Liguori M, Cocco E, Marrosu MG, Tesser F, Leone MA, Weber A, Zipp F, Miterski B, Epplen JT, Oturai A, Sorensen PS, Celius EG, Lara NT, Montalban X, Villoslada P, Silva AM, Marta M, Leite I, Dubois B, Rubio J, Butzkueven H, Kilpatrick T, Mycko MP, Selmaj KW, Rio ME, Sa M, Salemi G, Savettieri G, Hillert J, Compston DA. Multiple Sclerosis Severity Score: using disability and disease duration to rate disease severity. Neurology. 2005;64(7):1144-51.

10. santé Omdl, Organization WH, Staff W, WHO. The ICD-10 Classification of Mental and Behavioural Disorders: Clinical descriptions and diagnostic guidelines: World Health Organization; 1992.

11. Sanchez RA, Ramos F, Giannone C, Fischer P, Masnatta L, Baglivo HP, Ramirez AJ, Hollenberg NK. Parallel renal and extremity blood supply abnormalities in nonmodulation: responses to ACE inhibition. Hypertension. 2003;41(4):919-24.

12. Mancia G, De Backer G, Dominiczak A, Cifkova R, Fagard R, Germano G, Grassi G, Heagerty AM, Kjeldsen SE, Laurent S, Narkiewicz K, Ruilope L, Rynkiewicz A, Schmieder RE, Boudier HA, Zanchetti A, Vahanian A, Camm J, De Caterina R, Dean V, Dickstein K, Filippatos G, Funck-Brentano C, Hellemans I, Kristensen SD, McGregor K, Sechtem U, Silber S, Tendera M, Widimsky P, Zamorano JL, Erdine S, Kiowski W, Agabiti-Rosei E, Ambrosioni E, Lindholm LH, Viigimaa M, Adamopoulos S, Bertomeu V, Clement D, Farsang C, Gaita D, Lip G, Mallion JM, Manolis AJ, Nilsson PM, O'Brien E, Ponikowski P, Redon J, Ruschitzka F, Tamargo J, van Zwieten P, Waeber B, Williams B. 2007 Guidelines for the Management of Arterial Hypertension: The Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). Journal of hypertension. 2007;25(6):1105-87.

13. Schandry R. Heart beat perception and emotional experience. Psychophysiology. 1981;18(4):483-8.

14. Yoris A, García A, Traiber L, Santamaría-García H, Esteves S, Martorell M, Alifano F, Kichic R, Moser J, Cetkovich M, Manes F, Ibanez A, Sedeno L. The inner world of overactive monitoring: neural markers of interoception in OCD. Psychological Medicine. 2017;Accepted.

15. Garcia-Cordero I, Esteves S, Mikulan EP, Hesse E, Baglivo FH, Silva W, Garcia MDC, Vaucheret E, Ciraolo C, Garcia HS, Adolfi F, Pietto M, Herrera E, Legaz A, Manes F, Garcia AM, Sigman M, Bekinschtein TA, Ibanez A, Sedeno L. Attention, in and Out: Scalp-Level and Intracranial EEG Correlates of Interoception and Exteroception. Frontiers in neuroscience. 2017;11:411.

16. Garcia-Cordero I, Sedeno L, de la Fuente L, Slachevsky A, Forno G, Klein F, Lillo P, Ferrari J, Rodriguez C, Bustin J, Torralva T, Baez S, Yoris A, Esteves S, Melloni M, Salamone P, Huepe D, Manes F, Garcia AM, Ibanez A. Feeling, learning from and being aware of inner states: interoceptive dimensions in neurodegeneration and stroke. Philos Trans R Soc Lond B Biol Sci. 2016;371(1708).

17. Yoris A, Abrevaya S, Esteves S, Salamone P, Lori N, Martorell M, Legaz A, Alifano F, Petroni A, Sanchez R, Sedeno L, Garcia AM, Ibanez A. Multilevel convergence of interoceptive impairments in hypertension: New evidence of disrupted body-brain interactions. Human brain mapping. 2018;39(4):1563-81.

18. Melloni M, Sedeno L, Couto B, Reynoso M, Gelormini C, Favaloro R, Canales-Johnson A, Sigman M, Manes F, Ibanez A. Preliminary evidence about the effects of meditation on interoceptive sensitivity and social cognition. Behav Brain Funct. 2013;9:47.

19. Couto B, Adolfi F, Velasquez M, Mesow M, Feinstein J, Canales-Johnson A, Mikulan E, Martinez-Pernia D, Bekinschtein T, Sigman M, Manes F, Ibanez A. Heart evoked potential triggers brain responses to natural affective scenes: A preliminary study. Auton Neurosci. 2015;193:132-7.

20. Yoris A, Esteves S, Couto B, Melloni M, Kichic R, Cetkovich M, Favaloro R, Moser J, Manes F, Ibanez A, Sedeno L. The roles of interoceptive sensitivity and metacognitive interoception in panic. Behav Brain Funct. 2015;11:14.

21. Couto B, Adolfi F, Sedeno L, Salles A, Canales-Johnson A, Alvarez-Abut P, Garcia-Cordero I, Pietto M, Bekinschtein T, Sigman M, Manes F, Ibanez A. Disentangling interoception: insights from focal strokes affecting the perception of external and internal milieus. Frontiers in psychology. 2015;6:503.

22. de la Fuente A, Sedeno L, Vignaga SS, Ellmann C, Sonzogni S, Belluscio L, Garcia-Cordero I, Castagnaro E, Boano M, Cetkovich M, Torralva T, Canepa ET, Tagliazucchi E, Garcia AM, Ibanez A. Multimodal neurocognitive markers of interoceptive tuning in smoked cocaine. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology. 2019.

23. Stanislaw H, Todorov N. Calculation of signal detection theory measures. Behavior research methods, instruments, & computers. 1999;31(1):137-49.

24. Stanislaw H, Todorov N. Calculation of signal detection theory measures. Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc. 1999;31(1):137-49.

25. Salamone PC, Esteves S, Sinay VJ, Garcia-Cordero I, Abrevaya S, Couto B, Adolfi F, Martorell M, Petroni A, Yoris A, Torquati K, Alifano F, Legaz A, Cassara FP, Bruno D, Kemp AH, Herrera E, Garcia AM, Ibanez A, Sedeno L. Altered neural signatures of interoception in multiple sclerosis. Human brain mapping. 2018;39(12):4743-54.

26. Chung J, Yoo K, Lee P, Kim CM, Roh JH, Park JE, Kim SJ, Seo SW, Shin JH, Seong JK, Jeong Y. Normalization of cortical thickness measurements across different T1 magnetic resonance imaging protocols by novel W-Score standardization. Neuroimage. 2017;159:224-35.

27. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA data mining software: an update. ACM SIGKDD explorations newsletter. 2009;11(1):10-8.

28. Witten IH, Frank E, Hall MA, Pal CJ. Data Mining: Practical machine learning tools and techniques: Morgan Kaufmann; 2016.

29. Santamaria-Garcia H, Reyes P, Garcia A, Baez S, Martinez A, Santacruz JM, Slachevsky A, Sigman M, Matallana D, Ibanez A. First Symptoms and Neurocognitive Correlates of Behavioral Variant Frontotemporal Dementia. Journal of Alzheimer's disease : JAD. 2016;54(3):957-70.

30. Guyon I, Weston J, Barnhill S, Vapnik V. Gene selection for cancer classification using support vector machines. Machine learning. 2002;46(1-3):389-422.

31. Koul A, Becchio C, Cavallo A. PredPsych: A toolbox for predictive machine learning-based approach in experimental psychology research. Behavior research methods. 2018;50(4):1657-72.

32. Fisher RA. The use of multiple measurements in taxonomic problems. Annals of eugenics. 1936;7(2):179-88.

33. Rao CR. The utilization of multiple measurements in problems of biological classification. Journal of the Royal Statistical Society Series B (Methodological). 1948;10(2):159-203.