Accepted manuscript to appear in IJNS

Accepted Manuscript

International Journal of Neural Systems

Article Title:	Multivariate pattern analysis techniques for electroencephalography data to study Flanker interference effects
Author(s):	David Lopez-Garcia, Alberto Sobrado, Jose M. G. Penalver, Juan Manuel Gorriz, Maria Ruz
DOI:	10.1142/S0129065720500240
Received:	02 March 2020
Accepted:	03 March 2020
To be cited as:	David Lopez-Garcia *et al.*, Multivariate pattern analysis techniques for electroencephalography data to study Flanker interference effects, *International Journal of Neural Systems*, doi: 10.1142/S0129065720500240
Link to final version:	https://doi.org/10.1142/S0129065720500240

This is an unedited version of the accepted manuscript scheduled for publication. It has been uploaded in advance for the benefit of our customers. The manuscript will be copyedited, typeset and proofread before it is released in the final form. As a result, the published copy may differ from the unedited version. Readers should obtain the final version from the above link when it is published. The authors are responsible for the content of this Accepted Article.

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

World Scientific
www.worldscientific.com

Accepted manuscript to appear in IJNS

International Journal of Neural Systems, Vol. 0, No. 0 (2019) 1–16
© World Scientific Publishing Company

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

Multivariate pattern analysis techniques for electroencephalography data to study Flanker interference effects

David López-García, Alberto Sobrado, José M. G. Peñalver
Mind, Brain and Behavior Research Center.
University of Granada, Granada, 18071 Spain.
E-mail: dlopez@ugr.es

Juan Manuel Górriz
Department of Signal Theory, Telematics and Communications,
University of Granada, Granada, 18071 Spain.
E-mail: gorriz@ugr.es

María Ruz
Mind, Brain and Behavior Research Center.
Department of Experimental Psychology.
University of Granada, Granada, 18071 Spain.
E-mail: mruz@ugr.es

A central challenge in cognitive neuroscience is to understand the neural mechanisms that underlie the capacity to control our behavior according to internal goals. Flanker tasks, which require responding to stimuli surrounded by distracters that trigger incompatible action tendencies, are frequently used to measure this conflict. Even though the interference generated in these situations has been broadly studied, multivariate analysis techniques can shed new light into the underlying neural mechanisms. The current study is an initial approximation to adapt an interference Flanker paradigm embedded in a Demand-Selection Task to a format that allows measuring concurrent high-density electroencephalography. We used multivariate pattern analysis (MVPA) to decode conflict-related electrophysiological markers associated with congruent or incongruent target events in a time-frequency resolved way. Our results replicate findings obtained with other analysis approaches and offer new information regarding the dynamics of the underlying mechanisms, which show signs of reinstantiation. Our findings, some of which could not had been obtained with classic analytical strategies, open novel avenues of research.

Keywords: multivariate pattern analysis; electroencephalography; classification; support vector machine; demand-selection task.

1. Introduction

Cognitive control comprises a set of mechanisms that allow humans to behave according to their internal goals while ignoring distracting information.[1] The Flanker task,[2] where participants respond to the direction of an arrow surrounded by other distracting arrows, is among the most used in the field. The main result of this task is the so-called interference or conflict effect, where responses are slower and less accurate in incongruent (when the direction of the distracters is opposite to the target) vs. congruent trials. In the current study, we employed the described Flanker task in the context of effort avoidance.[3] Cognitive control involves effort, which is costly and partly aversive, and thus humans usually avoid it if given the chance. In Demand- Selection Tasks (DST),[3] participants tend to choose the easy option over the hard one. The tendency to avoid

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

the hard option seems partly due to the cost of overcoming the increased cognitive control required when responding to incongruent situations. However, the neural underpinnings of this effect are not well understood.

The majority of Electroencephalography (EEG) studies of the interference effect have analyzed Event-Related Potentials (ERPs), focusing on the N2 component. Besides, studies employing frequency analyses have shown Theta and Delta band involvement. Other authors[4] have proposed a link between the ERPs and modulations in the Delta-Theta band of frequency. These univariate approaches have been the gold standard in the EEG literature for years, not only to study the interference effects but several cognitive processes.

In recent years, newer Multivariate Pattern Analysis (MVPA) techniques based on Machine Learning algorithms, in conjunction with neuroimaging techniques such as functional Magnetic Resonance Imaging (fMRI), Electroencephalography or Magnetoencephalography (MEG), have gained popularity in Cognitive Neuroscience.[5–10] These supervised Machine Learning algorithms, particularly Linear Support Vector Machines (LSVM)(Vapnik, 1979)[11,12] have been also widely applied in clinical settings such as computer-aided diagnosis of Alzheimer's disease,[13–16] automatic sleep stages classification[17,18] or automatic detection of sleep disorders.[19]

One of the most remarkable advantages of these multivariate over univariate approaches is its sensitivity in detecting subtle changes in the patterns of activations associated with specific information content.[20] When applied to fMRI data, the poor temporal resolution of the signal prevents an accurate study of how cognitive processes unfold in time. In contrast, when applied to M/EEG signals,[21] MVPA has been useful to uncover the neural dynamics of face detection,[22] the process of memory retrieval,[23] the representational dynamics of task and object processing in humans[24] or the representation of spoken words in bilingual listeners.[25] In the same line, time-resolved MVPA presents an opportunity to categorize the temporal sequence of the neural processes underlying the interference effect. Furthermore, the relationship between these and Theta frequency modulations reported in previous studies[26] can be better understood using this approach.

This study is an extension of previous work[27] which adapted a DST to a format that allows measuring concurrent high-density electroencephalography. Our main goal is to present a set of methodological MVPA tools that allow to study and decode the conflict-related neural processes underlying interference effects, in a time-frequency resolved way.

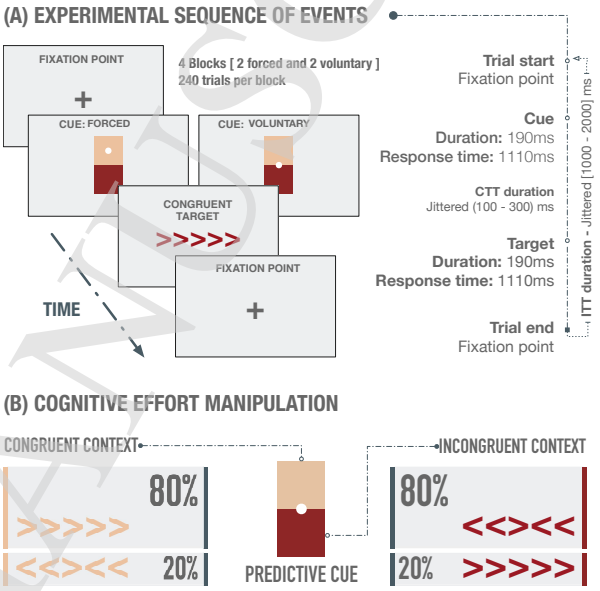

Figure 1. **(A)** Experimental sequence of events in case of a correct response on both cue and flanker stimuli. Each trial started with a fixation point, followed by a cue, which acted as a selector of the difficulty of the upcoming Flanker target. Participants had to choose (freely or forced, depending on the block type) the posible color of the upcoming target stimulus, which was associated with either high (difficult) or low (easy) probability of incongruent trials. Finally, after a variable time interval (100-300ms) the target stimulus appeared and participants had to respond to the orientation of the central arrow. Another variable time interval appeared before the beginning of the next trial. The cue and the target stimuli remained on screen for 190ms. **(B)** Cognitive effort was manipulated through the percentage of congruent and incongruent trials. Each cue color was associated with the high or low conflict contexts.

2. Materials and methods

2.1. *Paradigm and data acquisition*

2.1.1. *Participants*

Thirty-two healthy individuals (21 females, 29 right-handed, mean age = 24.65, SD = 4.57) were recruited for the experiment. The sex imbalance reflects the usual distribution of gender in the student pool (Psy-

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

chology) where participant recruitment took place. Participants had normal or corrected-to-normal vision and no neurological or psychiatric disorders. All of them provided informed, written consent before the beginning of the experiment and received a 10-euro payment or course credits in exchange for their participation. The experiment was approved by the Ethics Committee of the University of Granada.

2.1.2. *Stimuli and apparatus*

Stimuli presentation and behavioral data collection were carried out using MATLAB (MathWorks) in conjunction with Phychtoolbox-3.[28] The visual stimuli were presented in an LCD screen (Benq, 1920x1080 resolution, 60 Hz refresh rate) and placed 68.31±5.37 cm away of participants' Glabella, in a magnetically shielded room. Using a photodetector, the stimuli onset lag was measured at 8ms, which corresponds to half of the refresh rate of the monitor. Triggers were sent from the presentation computer to the EEG recording system through an 8-bit parallel port and using a custom MATLAB function in conjunction with inpoutx64 driver,[29] a C++ extension (mex-file) that uses native methods to access low-level hardware in MATLAB (I/O parallel ports).

Cues consisted of two squares of two different colors (red/green and yellow/blue, in different blocks) stacked and presented at the center of the screen (visual angle ∼5 degree). In forced blocks, a small white indicator (circle 50% or square 50%) appeared on top of the color that had to be chosen. In voluntary blocks, this indicator appeared between the two colored squares (see Figure 1). Each target stimulus consisted of five arrows pointing left or rightwards, which were displayed at the center of the screen (visual angle ∼6 degree). The color of the target stimulus was the same as the cue previously selected.

2.1.3. *Procedure*

The Color-Based Demand-Selection Task (DST) (Figure 1 a), modified from,[3] consisted of a cue-target sequence arranged in four blocks (2 forced and 2 voluntary). In voluntary blocks, participants were required to freely choose one of the two colors available, which indicated the difficulty of the upcoming task. In forced blocks, a small white indicator appeared on top of the color that had to be chosen. The color of the target stimulus was the same as the cue previously selected and participants were required to discriminate the orientation (right or left) of a central arrow target surrounded by arrows pointing at the same (compatible distractors) or opposite (incompatible distractors) directions.

Our task was built following a 3-way factorial design, containing the following within-subjects independent variables: (1) Stimulus type (congruent/incongruent); (2) Block type (forced/voluntary) and (3) Context (easy/difficult). The task difficulty manipulation was based on the proportion of congruent and incongruent trials, with the easy contexts presenting 80% of congruent and 20% of incongruent trials, and the difficult task context the opposite proportion. Within forced blocks, half of the trials corresponded to the easy context and the remaining to the difficult one (maintaining, within each condition, the proportion of congruent and incongruent trials). On voluntary blocks, however, participants freely chose the context and no experimental control could be exerted upon this variable.

Participants were instructed to respond as fast and accurately as possible, and to not choose color based on personal preference. They were unaware of the cognitive effort manipulation. To preserve the signals as clean as possible and remove the least number of trials, participants were encouraged to remain as still and relaxed as possible, avoiding face muscle activity and eye movements, but blinking normally. The order of the blocks, cue colors, response keys and color-conflict context mappings were counterbalanced across participants. There were 4 blocks, 240 trials per block, and the total recording session lasted ∼90min. Before the experimental session, participants performed a brief practice to familiarize themselves with the task (4 blocks, 20 trials per block, practice duration ∼20min). To reduce fatigue, there were rest periods between blocks, with a variable duration depending on participants' choice. During this period, participants were asked to remain seated and rest their eyes and posture before continue with the task. Additionally, block order was counterbalanced across participants, and within each block, trial order was randomized, which effectively prevents confounds due to differential levels of fatigue across conditions.

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

2.1.4. *EEG acquisition and preprocessing*

High-density electroencephalography was recorded from 65 electrodes mounted on an elastic cap (actiCap slim, Brain Products) at the Mind, Brain, and Behavior Research Center (CIMCYC, University of Granada, Spain). The TP9 and TP10 electrodes were used to record the electrooculogram (EOG) and were placed below and next to the left eye of the participant. Impedances were kept below $5k\Omega$, as recommended by the amplifiers manufacturer. EEG activity was referenced online to the FCz electrode and signals were digitized at a sampling rate of 1KHz.

Electroencephalography recordings were average referenced, downsampled to 256Hz, and digitally filtered using a low-pass FIR filter with a cutoff frequency of 120Hz, preserving phase information. The recording amplifiers have an intrinsic lower cutoff frequency of 0.016Hz (time constant $\tau = 10$s).

No channel was interpolated for any participant. EEG recordings were epoched [-1000, 2000ms centered at onset of the target arrows] and baseline corrected [-200, 0ms], and data were extracted only from correct trials. To remove blinks from the remaining data, Independent Component Analysis (ICA) was computed using the *runica* algorithm in EEGLAB,[30] excluding TP9 and TP10 channels. Artifactual components were rejected by visual inspection of raw activity of each component, scalp maps and power spectrum. Then, an automatic trial rejection process was performed, pruning the data from non-stereotypical artifacts. The trial rejection procedure was based on (1) abnormal spectra: the spectrum should not deviate from baseline by ±50dB in the 0-2 Hz frequency window, which is optimal for localizing any remaining eye movements, and should not deviate by -100dB or +25dB in 20-40Hz, useful for detecting muscle activity (∼1% of the total sample was rejected); (2) improbable data: the probability of occurrence of each trial was computed by determining the probability distribution of values across the data epochs. Trials were thresholded, in terms of $\pm 6SD$, and automatically rejected (∼6% of the total sample); (3) extreme values: all trials with amplitudes in any electrode out of a $\pm 150\mu$V range were automatically rejected (∼3% of the total sample). See[31–33] for similar EEG preprocessing routines.

2.1.5. *Final dataset description*

The final dataset for our binary classification problem is shown in Table 1, where N is the initial number of trials per participant and class, N_r represents the number of remaining correct trials after the trial rejection stage and N_b is the final number of trials used for classification per participant (after downsampling the majority class in order to get balanced datasets).

Table 1. Number of observations of the final dataset

Observations per participant	N	N_r	N_b
Congruent class	480	426±49	359±52
Incongruent class	480	368±59	359±52
Total number of observations		N_{Tr}	N_{Tb}
Congruent class		13644	11505
Incongruent class		11782	11505

2.1.6. *Behavioral data analysis*

Reaction time (RT) and error rates were registered for each participant. Before the statistical analysis, the first trial of each block, trials with choice errors and trials after errors were filtered out.[34] Finally, RT outliers were also rejected using a ±2.5 SD threshold, calculated individually per participant and condition. To analyze behavioral data (accuracy and reaction times) we conducted repeated-measures ANOVAs in SPSS Statistics Software (v.20). Post hoc tests were carried out on the significant interactions using a Bonferroni correction for multiple comparisons.

2.2. *Multivariate pattern analysis*

The MVPA for the decoding analysis was performed in MATLAB by a custom-developed set of linear Support Vector Machines, trained to discriminate between congruent and incongruent target stimuli. To avoid skewed classification results, the datasets were strictly balanced, by downsampling the majority class to match the size of the minority one. In addition, class size was set as a factor of k, the total number of folds in the cross-validation stage. Accordingly, each fold was composed by exactly the same number of observations, avoiding any kind of bias in

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

the results. The rest of the classification parameters remained by default.

2.2.1. *Feature extraction*

To obtain the classification performance in a time-resolved way, the feature vectors were extracted as shown in Figure 2. The classification procedure, for each participant, ran as follows: (1) For each timepoint and trial, we generated two feature vectors (one for each condition or class) consisting of the raw potential measured in all electrodes (excluding EOG electrodes: TP9 and TP10). (2) Each individual feature vector, containing raw potential values were normalized (z-score, $\mu = 0, \sigma = 1$).

2.2.2. *Supertrial generation*

Due to the noisy nature of the EEG signal, a trial averaging approach was carried out during the feature extraction stage. This approach increases the signal-to-noise ratio (SNR),[35] improves the overall decoding performance and also reduces the computational load. Each participant's dataset was reduced by randomly averaging a number of trials t_a belonging to the same condition. The value of t_a is a trade-off between an increased classification performance (due to an increased SNR) and the variance in the classifier performance, since reducing number of trials per condition typically increases the variance in (within-participant) classifier performance.[36] Therefore, the optimal number of trials to average depends on the dataset, taking into account that averaging more trials does not increment the decoding performance linearly.

2.2.3. *Feature selection*

As mentioned in section 2.2.1, $\mathrm{X}_{\mathrm{n}\times\mathrm{p}}$ datasets are generated for each participant and timepoint, where n is the number of trials (observations) and p the total number of electrodes (variables or features). In machine learning, feature selection techniques, also known as dimension reduction, are a common practice to reduce the number of variables in high-dimensional datasets (Figure 3). Principal Component Analysis (PCA) is probably the most popular multivariate statistical technique and it is used in almost all scientific disciplines,[37] including Neuroscience.[38]

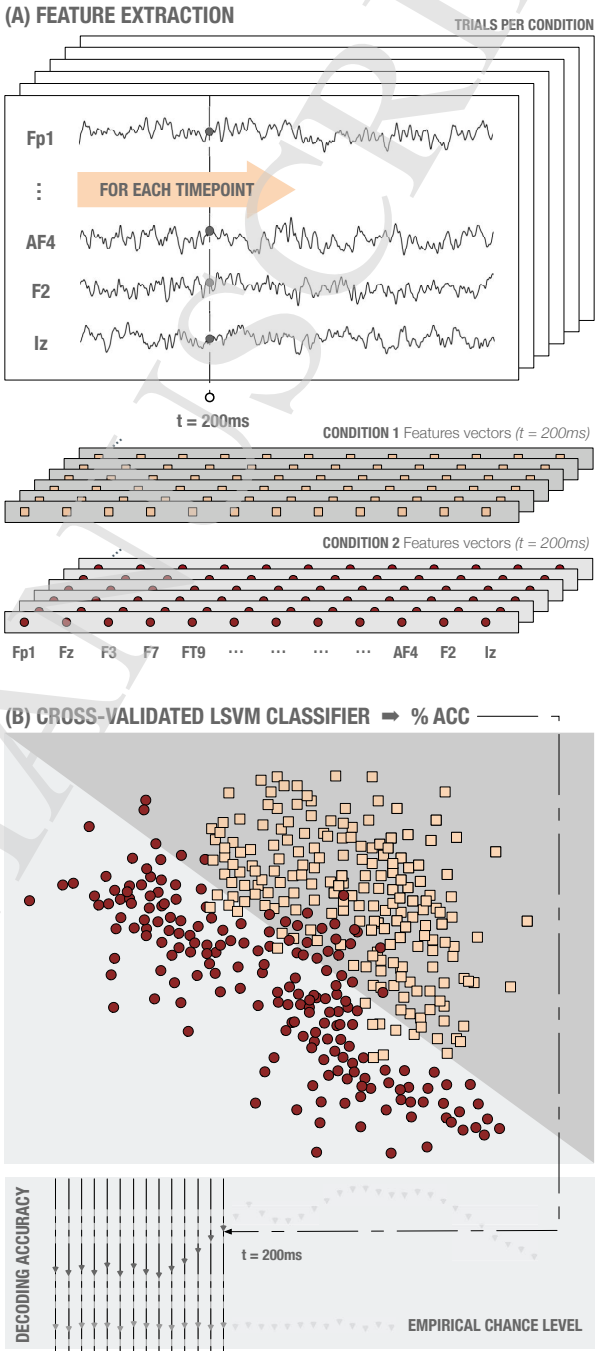

Figure 2. **(A)** Feature extraction process in simulated data. The feature vectors of each condition and time point consisted of a z-scored voltage array for all the scalp electrodes. For an improved SNR, several trials were averaged before feature extraction. **(B)** Cross-validated LSVM classifier. For each time point, an LSVM was trained and tested (stratified k-fold cross-validation, $k = 5$). Chance level was calculated by permuting the labels.

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

PCA is a linear transformation of the original dataset in an orthogonal coordinate system in which axis coordinates (principal components) correspond to the directions of highest variance sorted by importance. To compute this transformation,[39] each row vector $\mathbf{x}_i$ of the original dataset $\mathbf{X}$ is mapped to a new vector of principal components $\mathbf{t}_i = (t_1, ..., t_l)$, also called *scores*, using a p-dimensional *coefficient* vector $\mathbf{w}_j = (w_1, ..., w_p)$. For dimension reduction, $l < p$.

$$\mathbf{t}_i = \mathbf{x}_i \cdot \mathbf{w}_j \quad i = 1, ..., n \quad j = 1, ..., l \tag{1}$$

To maintain the model's performance as fair as possible, in our study PCA was computed only for training sets $\mathbf{X}_{\text{training}}$, independently for each fold inside the cross-validation procedure. Once PCA for the corresponding training set was computed and the model was trained, the exact same transformation was applied to the test set $\mathbf{X}_{\text{test}}$ (including centering, μ_{training}). In other words, the test set was projected onto the reduced feature space obtained during the training stage. According to equation (1), this projection is computed as follows:

$$\mathbf{T}_{\text{test}} = \frac{\mathbf{X}_{\text{test}} - \mu_{\text{training}}}{\mathbf{W'}_{\text{training}}} \tag{2}$$

Feature selection techniques such PCA usually imply an intrinsic loss of spatial information, e.g. data projected from the sensor space onto the reduced PCA features space. Therefore, PCA presents a trade-off between dimension reduction and results' interpretation. If PCA is computed, the spatial information of each electrode is lost, which means that, for example, we cannot directly analyze which electrodes are contributing more to decoding performance.

2.2.4. *Model's performance evaluation*

To evaluate classification models in neuroscience, performance is usually measured employing mean accuracy.[40] However, mean accuracy may generate systematic biases in situations with very skewed sample distributions, and overfitting one single class should be avoided. Therefore, nonparametric and criterion-free estimates, such as the Area Under the ROC Curve (AUC) have been proved as a better measure of generalization in these situations.[41] The AUC provides a way to evaluate the performance of a classification model. The larger the area, the more accurate the classification model is, and it is computed as follows:

$$\mathbf{AUC} = \int_0^1 \mathbf{ROC}(s)ds \tag{3}$$

The ROC curve is one of the most suitable evaluation criteria, as it shows how capable the model is in distinguishing between conditions, by facing the sensitivity (True Positive Rate, TPR) against 1-specificity (False Positive Rate, FPR). In this study, we employed both methods, the mean accuracy, to replicate a common approach in literature, and ROC curves and AUC, to provide a more informative measure.

To evaluate the performance of our model, LSVMs were trained and validated, resulting in a single performance value for each timepoint and participant. The classification performance at the group level was calculated by averaging these values across participants. The chance level was calculated following the former analysis but using randomly permuted labels for each trial.

The generalization ability of our model was estimated through a Cross-Validation (CV) approach (stratified k-fold, $k = 5$), which is a well-established and a widely implemented technique to preserve complex models from overfitting.

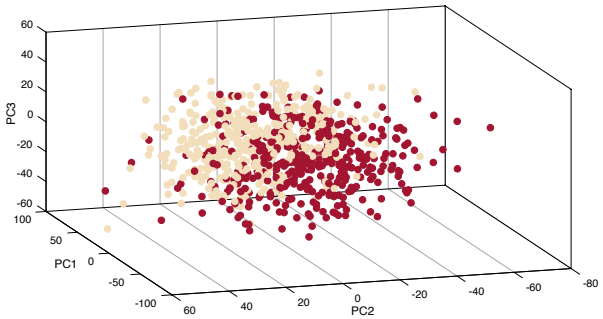

Figure 3. Dimension reduction in real datasets. 3D representation of the three first PCA components for congruent vs. incongruent trials [example participant, t = 421ms after Flanker stimulus onset]

Moreover, some important aspects are worth being highlighted. The use of CV approaches often leads (particularly in Neuroscience) to small sample

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

sizes and a high level of heterogeneity when conditions are splitted into each fold, causing among other things a large classification variability.[42] To address these problems, recent studies[43,44] considered the use of the resubstitution error estimate when using LSVM (in small sample sizes and low dimensional scenarios), proposing a novel analytic expression for the upper bound on the actual risk $\gamma_{emp}(l,d)$ for a range of sample sizes l, dimensions d and any significance $\eta < 0.05$ (Figure 8). Therefore, the difference between the actual error and the resubstitution error is bounded by the actual risk γ_{emp}, which is computed as follows:

$$\gamma_{emp} \leq \sqrt{\frac{1}{2l} \ln \frac{N(l,d)}{\eta}} \tag{4}$$

where N is defined in[43] as:

$$N(l,d) = 2 \sum_{k=0}^{d-1} \binom{l-1}{k} \tag{5}$$

Resubstitution has been proved competitive in some heterogeneous-data scenarios with CV approaches not only in terms of accuracy but also in computational load.[45] The proposed solution has been recently applied in clinical settings studying autistic patterns[46] or Alzheimers Disease.[47] The scenario previously mentioned (linear classifiers, small sample size and low dimensional space) seems to fit perfectly with our study setup, therefore, we also used the resubstitution error estimate to evaluate the classification performance.

2.2.5. *Optimization of SVM hyperparameters*

A search-grid based optimization of the misclassification cost parameter C was carried out using five-fold cross-validation on the training set:

$$||\beta||^2 + C \sum_{i=1}^{l} \xi_i \tag{6}$$

where C is a constant which modulates the trade-off between the training error and the complexity of the model and the vector β contains the coefficients that define an orthogonal vector to the hyperplane.

2.3. *Temporal generalization matrix*

Temporal generalization analyses are used to evaluate the stability of the brain patterns along time, by training the model in one temporal point and testing its ability to discriminate between conditions in the remaining temporal window. This process is repeated for every timepoint. In our study, classification performance was assessed through a cross-validation technique (stratified k-fold cross-validation, $k = 5$). For each timepoint, the classifier was trained with $\mathbf{X}_{\text{training}}$ dataset and tested with $\mathbf{X}_{\text{test}}$ in the remaining points of the temporal window. This process was repeated k times, obtaining the final decoding accuracy.

An above-chance discrimination rate outside the diagonal of the matrix suggests that the same activity pattern is sustained in time. However, if there is no evidence of temporal generalization, different patterns of activity can be assumed.[24]

2.4. *Multivariate cross-classification*

The ability of MVPA to detect subtle differences in brain activity patterns can be used to study how these patterns are similar across different cognitive contexts. In other words, the consistency of the information across different sets of data can be analyzed. To this end, classification algorithms are trained with one set of data and the consistency is assessed by testing the model with another dataset, belonging to a different experimental condition. This technique is called Multivariate Cross-Clasification[20](MVCC) and is growing in popularity in recent years.[48–50]

The fact that the training and test sets are different eliminates the need to use cross-validation techniques. However, the classification direction have to be taken into account, that is, which set is used for training and which one for testing. The result of the classification could differ if, for instance, the signal-to-noise ratio is quite different across datasets, that is to say, differences in classes separability across datasets and an asymmetry in the generalization direction.[51] For this reason, reporting results in both directions is highly recommended.

In this study, MVCC was used to analyze if the neural patterns associated with the congruency effect are similar across voluntary and forced blocks. For that, classifiers were trained with data of forced blocks and tested in voluntary blocks, and *vice versa*.

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

Int. J. Neur. Syst. Downloaded from www.worldscientific.com by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

In addition, a temporal generalization matrix was also computed to study the similarity across block types and time. Feature selection in MVCC analysis also requires some additional considerations, as features selected for the training set could not be optimal for the test set. To avoid possible skewed results, no feature selection was computed for MVCC analyses.

2.5. *Statistical analysis*

Applying t-test statistics on multivariate results is an unsuitable approach to draw statistical inferences at the group level.[52] For that reason, the use of cluster-based non-parametric permutation methods is widespread, not only in fMRI[53–56] but more recently also in M/EEG studies.[57–60] In our study, a non-parametric cluster-based permutation approach, proposed in[52] for fMRI data, was adapted and implemented for the statistical analysis.

We thresholded the decoding accuracy obtained with an empirical accuracy null distribution, calculated by means of a combined permutation and bootstrapping technique. First, at the single-subject level, 100 randomly permuted accuracy maps were generated. To draw statistical inferences at the group level, we randomly drew one of the previously calculated accuracy maps for each participant. This selection was group-averaged and the procedure was repeated 10^5 times, generating 10^5 permuted group accuracy maps.

Next, for each timepoint we estimated the chance distribution of accuracy values and determined the accuracy threshold (99.9th percentile of the right-tailed area of the distribution), which corresponds to a very low probability to obtain significant results by chance.

Then, we searched and collected clusters of timepoints exceeding the previously calculated threshold in all the 10^5 permuted accuracy maps, generating the normalized null distribution of cluster sizes. Finally, we applied a correction for multiple comparisons (FDR, False Discovery Rate) at a cluster level to obtain the smallest cluster size to be considered significant.[61–64]

2.6. *Frequency contribution analysis*

The contribution of each frequency band to the overall decoding performance was assessed through an exploratory sliding filter approach. We designed a band-stop FIR filter using `pop_firws` EEGLab function (2Hz bandwidth, 0.2Hz transition band, 2816 filter order, Blackman window) and pre-filtered the EEG data (120 overlapped frequency bands, between 0-120Hz and linearly-spaced steps) producing 120 filtered versions of the original EEG dataset. The former time-resolved decoding analysis (congruent vs. incongruent, ta = 10) was repeated for each filtered version and the importance of each filtered-out band was quantified computing the difference maps in decoding performance between the filtered and the original decoding results. Significant clusters were found applying the proposed cluster-based permutation test to filtered-out datasets, generating accuracy null distributions for each time-frequency point.

With the purpose of obtaining better frequency resolution in lower bands, the previous analysis was repeated for frequencies between 0-40Hz in 120 overlapped and logarithmically spaced steps.

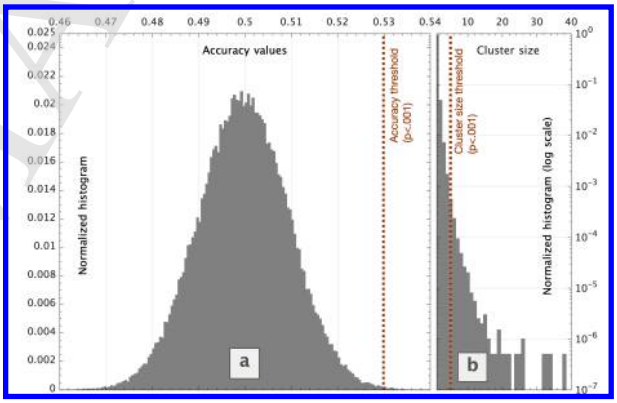

Figure 4. Accuracy **(a)** and cluster size **(b)** null distributions. The vertical dotted line represents the threshold corresponding to a very low probability to obtain significant results by chance. This threshold correspond to a p-value below 0.001 for both distributions.

3. Results

3.1. *Behavioral results*

The behavioral results of reaction times replicate well-known conflict effects linked to context-dependent congruency,[3,34] with a significant interaction of Context × Stimulus Type ($F_{(1,31)} = 26.285$, $p < .004$, $\eta_p^2 = .459$). Planned comparisons showed significant differences between congruent and incongruent trials for both the easy ($F_{(1,31)} = 272.707$, $p < .001$, $\eta_p^2 = .885$) and the difficult contexts

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

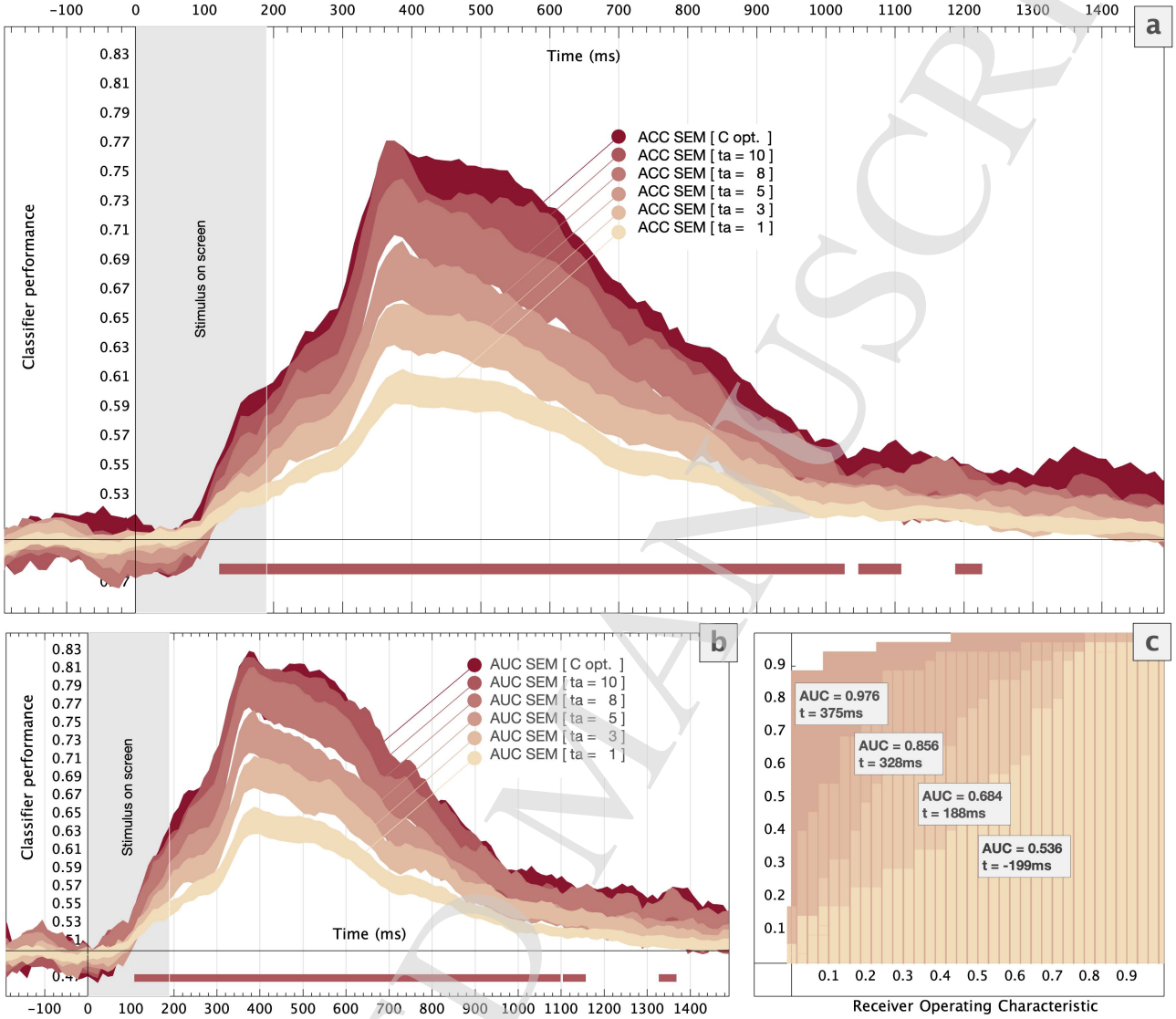

Figure 5. Group level MVPA results. Time-resolved classifier performance when different number of trials were averaged. The standard error of the **(a)** classification accuracy and **(b)** the Area Under the Curve are represented using colored areas. Significant windows ($t_a = 10$) obtained via Stelzer permutation test are highlighted using horizontal bold lines. The stimulus screen time $[0-190]$ms is shaded. **(c)** Receiver Operating Characteristic curves for different timepoints [example participant, $t_a = 10$, t =-199ms, 188ms, 328ms and 375ms].

($F_{(1,31)} = 183.109$, $p < .001$, $\eta_p^2 = .855$) with larger differences in reaction times in the easy (congruent trials: $M = 0.465$, $SD = 0.13$; incongruent trials: $M = 0.560$, $SD = 0.15$), compared to the difficult context (congruent trials: $M = 0.474$, $SD = 0.13$; incongruent trials: $M = 0.553$, $SD = 0.14$). The effort-avoidance effect, as expected, was also observed in voluntary decision blocks (percentage of choice of easy 57,11% $SEM = 2.93$ vs difficult 42,88% $SEM = 2.93$ contexts; $t = 2.42$, $p = .021$).

3.2. *Electrophysiological results*

The electrophysiological analyses (Figure 5a) show significant differences ($p <.001$, cluster corrected) in activity patterns for congruent vs. incongruent trials, peaking at 375ms after the stimulus onset. At this point, the classifier accurately predicted ($> 80\%$) if participants were responding to congruent or incon-

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

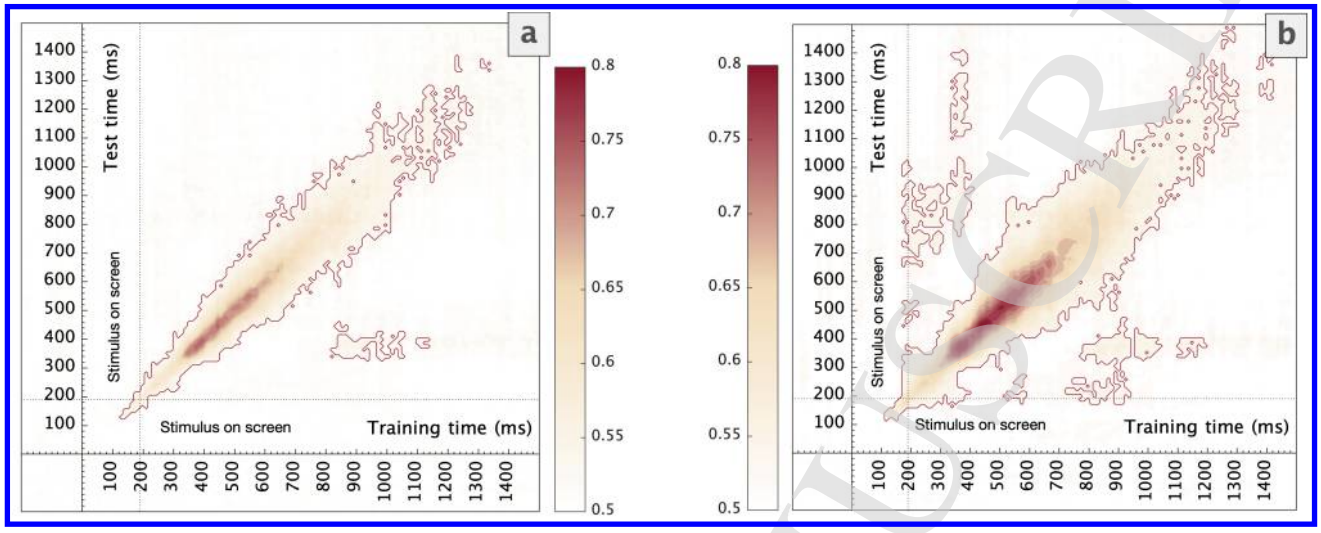

Figure 6. Group level temporal generalization results for congruent vs. incongruent trials ($t_a = 10$). Accuracy **(a)** and AUC **(b)** values when the model was trained and tested in each time point of the whole time window. Significant clusters obtained via Stelzer permutation tests are highlighted using red lines.

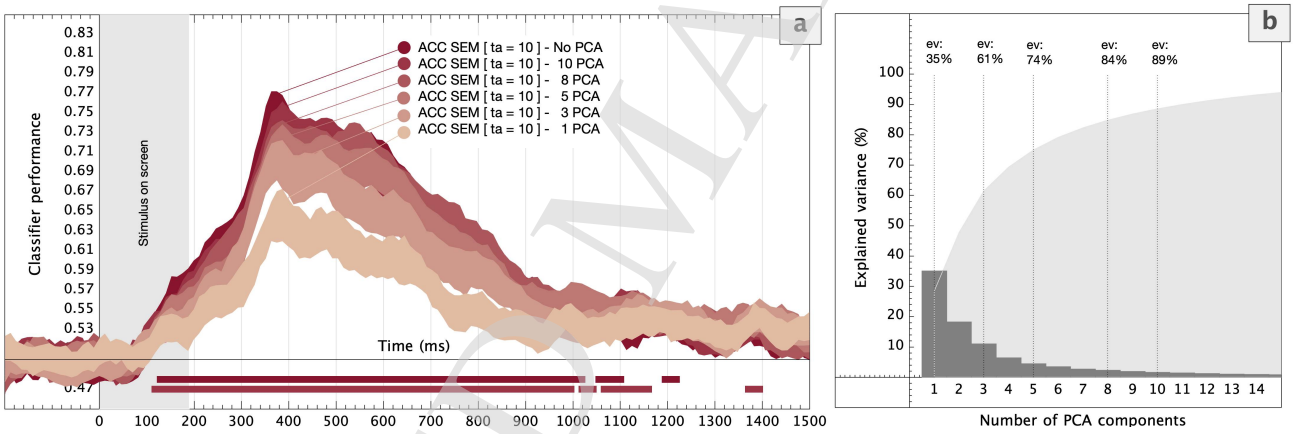

Figure 7. Group level MVPA. **(a)** Time-resolved classifier performance (ACC, congruent vs. incongruent trials) for different number of PCA components. Colored areas represent the ACC standard error. Statistically significant time windows ($t_a = 10$, for both 10 PCA components and when PCA was not computed) are highlighted using horizontal bold lines. The stimulus screen time [0-190]ms is shaded. **(b)** Explained variance for different numbers of PCA components [example participant].

gruent trials. Table 2 reports the variations in classification performance for averages of different number of trials. The SVM hyper-parameter C was optimized, slightly increasing the decoding performance; however, the computation time required increased significantly.

When ten trials were averaged to generate supertrials, the statistically significant regions extended from 130ms after stimulus onset to 1200ms afterwards. As Figure 5 shows, before the stimulus onset the classification accuracy remained at chance levels (0.5).

The temporal generalization analysis is shown in Figure 6. First, the AUC proved to be a more sensitive measure. The AUC temporal generalization matrix (Fig 6b) shows a distinct pattern of generalization. Clusters appearing only alongside the diagonal have been associated with a succession of different mechanisms. That is to say, the neural information that allows the classifier to tell apart congruent and incongruent trials is likely the result of a series of distinct events. Moreover, Figure 6b shows a clus-

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

ter of homogeneous AUC between 200 and 400ms, which theoretically suggests the operation of a single cognitive process maintained in time.[41] Such mechanism apparently reappears at ∼800-1000ms after the target onset, posterior to the mean RT (513ms).

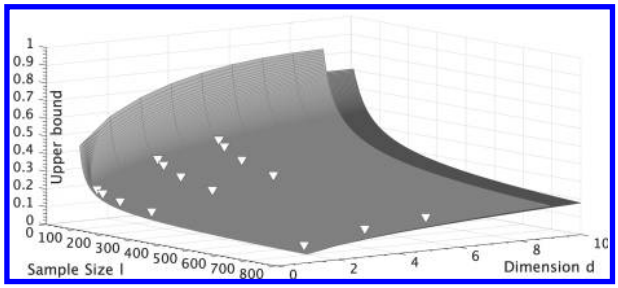

Figure 8. Different upper bound estimations via the procedure found in[43] for LSVM across dimension and sample size at a 95% confidence level ($\eta = 0.05$). White markers represent the upper bound values for the experimental conditions tested in our study.

The actual risk estimation for different sample sizes and dimensions γ_{emp} is shown in Figure 8. The difference between the actual error and the resubstitution error is bounded by γ_{emp}. White markers represent different experimental configurations for both the sample size (l) and the number of PCA components (d) analyzed in our study. Performance results obtained by resubstitution (C-optimized, $t = 375$ms) for these experimental configurations are shown in Table 3. The classification accuracy remained above chance despite the conservative estimation of the upper bound of the actual error, preserving our classification model for overfitting and proving that both conditions (congruent and incongruent) are representative of the different underlying activity patterns associated with congruent and incongruent stimuli.

The cross-classification results (Figure 9a,b) showed smaller clusters compared to the MVPA time generalization (Figure6a, b). However, the main diagonal cluster in the matrix indicates a series of different events that occur in cascade, but shared between both contexts.[41] This mechanism could reflect the interference process itself, previous to the response.

Table 2. LSVM model peak classification performance [$t = 375ms$] at the group level. The mean accuracy and AUC are reported for different values of t_a and different numbers of PCA components.

No. of averaged trials (t_a)	ACC ± SD	AUC ± SD
$t_a = 1$	.60±.05	.65±.07
$t_a = 3$	.65±.07	.70±.08
$t_a = 5$	.69±.10	.74±.12
$t_a = 8$	.74±.10	.79±.12
$t_a = 10$	.76±.11	.80±.13
$t_a = 10$ - C optimized	.76±.10	.81±.13
No. of PCA components ($t_a = 10$)		
First Component	.64±.14	.66±.17
3 first components	.71±.11	.76±.13
5 first components	.72±.11	.78±.12
8 first components	.73±.12	.80±.13
10 first components	.74±.11	.81±.13

Table 3. Classification performance and the actual risk γ_{emp} for a C-optimized LSVM model obtained by the resubstitution approach. [example participant, $t = 375$ms]

ACC($\gamma_{emp}(l, d)$)	$d = 1$PCA	$d = 3$PCA	$d = 5$PCA
$l = 790$ ($ta = 1$)	.55(.04)	.63(.10)	.65(.13)
$l = 260$ ($ta = 3$)	.58(.08)	.68(.16)	.71(.20)
$l = 150$ ($ta = 5$)	.58(.11)	.79(.20)	.81(.25)
$l = 90$ ($ta = 8$)	.72(.14)	.80(.24)	.83(.30)
$l = 70$ ($ta = 10$)	.81(.15)	.88(.27)	.92(.34)

3.3. *Frequency contribution results*

A sliding bandstop filter approach was followed to study the contribution of each frequency band to the overall decoding accuracy. Results show that the interference effect observed relies on neural processes operating in the Delta and Theta frequency bands. Figure 10a shows how decoding accuracy significantly drops when frequencies up to 8Hz were filtered-out.

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

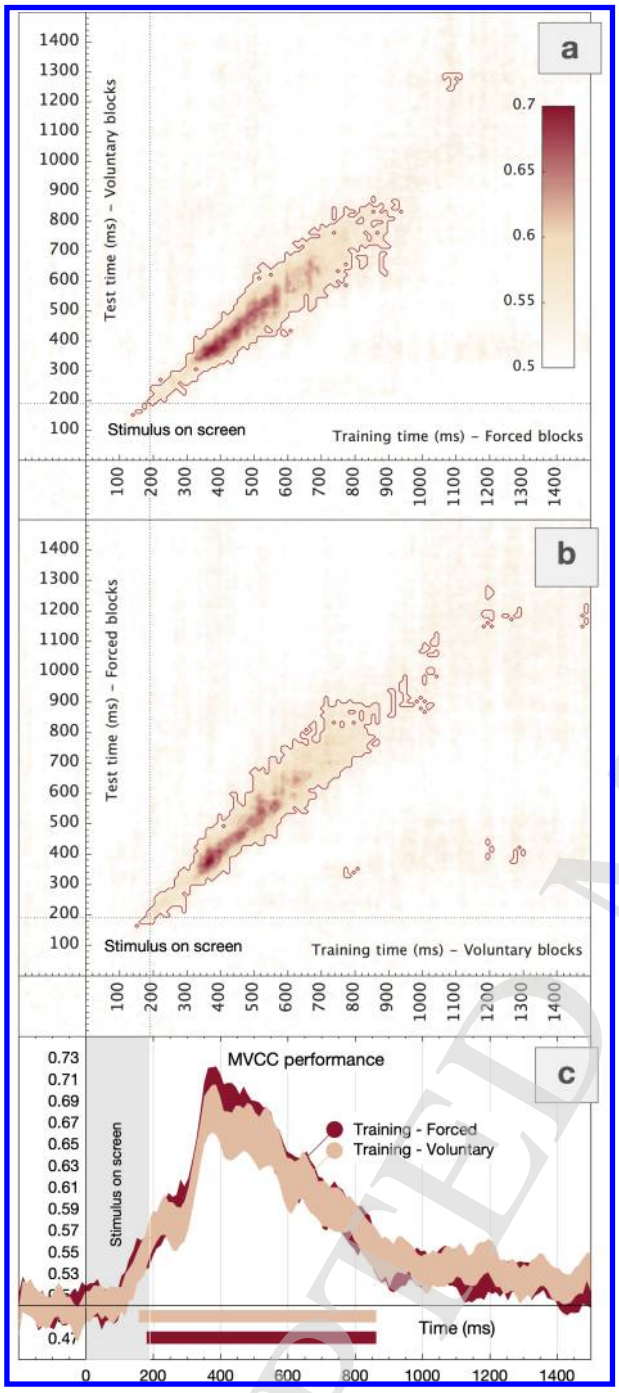

Figure 9. MVCC results. **(a)** Temporal generalization results when the model was trained with forced blocks and tested in voluntary blocks and **(b)** *vice versa*. **(c)** Classifier performance (acc) for the former analyses. Colored areas represent the standard error. Significant windows calculated via Stelzer permutation test are highlighted using horizontal lines.

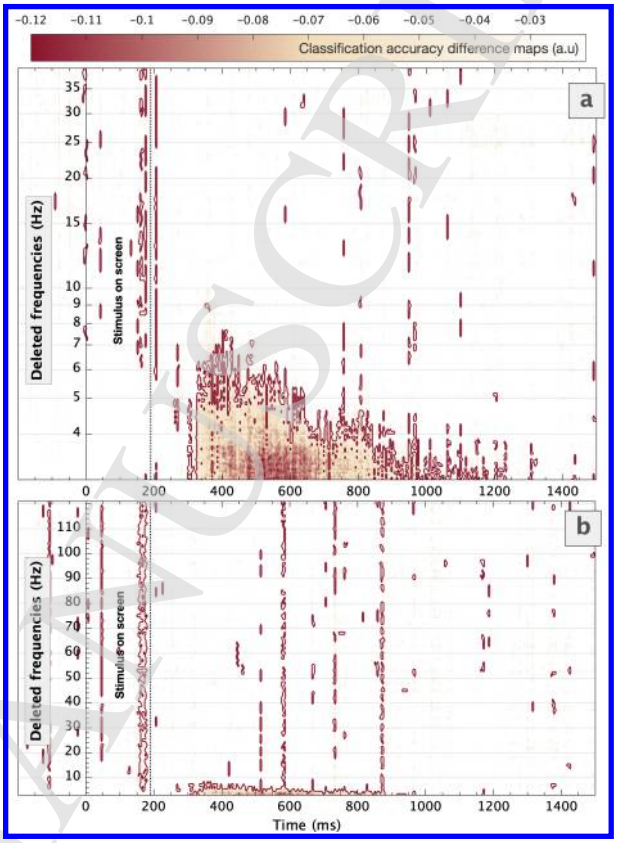

Figure 10. Results of frequency contribution analysis. Classification accuracy differences when a specific frequency band is filtered-out. **(a)** [0-40]Hz logarithmically spaced and **(b)** [0-120]Hz linearly spaced sliding filter approach. Significant clusters obtained via Stelzer permutation test are highlighted using red lines.

4. Discussion

In the current study, we present a set of multivariate pattern analysis techniques for EEG data. Overall, we effectively classified interference-related electrophysiological signals from a Demand-Selection Task in both a time and frequency-resolved manner.

Previous studies on cognitive control and more specifically, interference processing, have found that slow rhythms (i.e. Theta, Delta) are associated with communication between distant brain regions.[65] Our results are in line with those studies, showing that Theta and Delta oscillations are relevant for the brain activity underlying performance in an interference task. Moreover, previous results (e.g. Cohen[26]) show the relevance of Theta in the first instances of target processing, which changed to Delta after the participants' response. These results are supported

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

by the present study, which shows Theta and Delta to be crucial for classification right after the target onset, which evolves into a single Delta-based classification around and after the response time. The meaning of the change from one frequency band to another along time could be due to neuronal activity on the Theta band preventing the distractors to be processed. Once the target is selected, Delta, which arises later, could reflect inhibition of competing and erroneous motor responses.[66] Our results also indicate the existence of a particular brain process involved in the interference effect that intervenes in the initial stages of target processing during an extended time window, and reappears after the behavioral response is given. Interestingly, this is the same temporal window where classic Event-Related Potential studies[67] have repeatedly observed the N2 potential, which is taken as the reflection of interference processing. The indication that the same underlying mechanism reappears after the response could reflect the reinstantiation of the interference episode, perhaps reflecting trial event boundaries.[68] This finding, which could not have been obtained with classic analytical strategies, opens novel avenues of research and aids to better characterize a commonly used neural correlate. Further research will be needed to clarify and extend this phenomenon. To summarize, our behavioral and electrophysiological results add up to existing literature, pointing to an overall robust effect of interference and conflict avoidance, which can be observed in various environments and demographic samples.

Future lines of investigation should address these findings to complement the results found in the current investigation. In addition, to increase our understanding of preparation processes and conflict effects, it would be of interest to continue analyzing the current dataset, focusing not only on the target stimulus, but also on the neural activity triggered by the cues. Further detailed analyses should be carried out to study the activation differences between forced and voluntary blocks or high and low congruency contexts. The use of newer classification algorithms, such as Spiking Neural Networks[69–72] should be considered in related studies. They have been demonstrated to be more powerful in some scenarios than linear SVM.[73,74] Nonetheless, in the present study, given the small sample size and our main goal (effective discrimination in time and frequency regardless of the actual accuracy value obtained), we decided to use less complex algorithms, which lead to more easily interpretable results[75–77] and prevent model overfitting. For this, SVM-based multivariate techniques represent an opportunity to study the neural basis of complex psychological processes. In addition, the resubstitution error estimate proved that, even in the worst case scenario when the estimated actual risk is maximum, the classification performance remained over the chance level. This method, which is suitable for small sample sizes and low dimensional scenarios is worthy of consideration in Cognitive Neuroscience studies, opening up a new path that could lead to promising results.

5. Conclusion

The current study is an initial approximation to adapt a DST to a format that allows measuring concurrent high-density electroencephalography. While most of previous studies categorize the interference effect through ERP markers such as the N2 potential,[67] we successfully used multivariate pattern analysis (MVPA) to decode conflict-related neural processes associated with congruent or incongruent events in a time-frequency resolved way. Our results replicate findings obtained with other analysis approaches and offer new information regarding the dynamics of the underlying mechanisms.

Acknowledgments.

This research was supported by the Spanish Ministry of Economy and Business under the TEC2015-64718-R and PSI2016-78236-P grants. The first author of this work is supported by a scholarship from the Spanish Ministry of Economy and Business (BES-2017-079769).

Bibliography

1. M. M. Botvinick, C. S. Carter, T. S. Braver, D. M. Barch and J. D. Cohen, Conflict monitoring and cognitive control, *Psychological Review* **108**(3) (2001) 624–652.
2. B. A. Eriksen and C. W. Eriksen, Effects of noise letters upon the identification of a target letter in a nonsearch task, *Perception & Psychophysics* **16**(1) (1974) 143–149.
3. W. Kool, J. T. McGuire, Z. B. Rosen and M. M. Botvinick, Decision Making and the Avoidance of Cogntive Demand, *Journal of Experimental Psychology: General* **139**(4) (2010) 665–682.

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

Int. J. Neur. Syst. Downloaded from www.worldscientific.com by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

4. J. Harper, S. M. Malone and E. M. Bernat, Theta and delta band activity explain N2 and P3 ERP component activity in a go/no-go task, *Clinical Neurophysiology* **125**(1) (2014) 124–132.
5. N. Kriegeskorte and P. Bandettini, Analyzing for information, not activation, to exploit high-resolution fMRI, *NeuroImage* **38**(4) (2007) 649–662.
6. K. A. Norman, S. M. Polyn, G. J. Detre and J. V. Haxby, Beyond mind-reading: multi-voxel pattern analysis of fMRI data, *Trends in Cognitive Sciences* **10**(9) (2006) 424–430.
7. J. V. Haxby, A. C. Connolly and J. S. Guntupalli, Decoding Neural Representational Spaces Using Multivariate Pattern Analysis, *Annual Review of Neuroscience* **37**(1) (2014) 435–456.
8. M. Misaki, Y. Kim, P. A. Bandettini and N. Kriegeskorte, Comparison of multivariate classifiers and response normalizations for pattern-information fMRI, *NeuroImage* **53**(1) (2010) 103–118.
9. J. D. Johnson, S. G. McDuff, M. D. Rugg and K. A. Norman, Recollection, Familiarity, and Cortical Reinstatement: A Multivoxel Pattern Analysis, *Neuron* **63**(5) (2009) 697–708.
10. J. A. Etzel, N. Valchev and C. Keysers, The impact of certain methodological choices on multivariate analysis of fMRI data with support vector machines, *NeuroImage* **54**(2) (2011) 1159–1167.
11. B. E. Boser, I. M. Guyon and V. N. Vapnik, A training algorithm for optimal margin classifiers, *Proceedings of the fifth annual workshop on Computational learning theory - COLT '92*, (ACM Press, New York, New York, USA, 1992), pp. 144–152.
12. C. Cortes and V. Vapnik, Support-vector networks, *Machine Learning* **20**(3) (1995) 273–297.
13. J. Ramírez, J. Górriz, D. Salas-Gonzalez, A. Romero, M. López, I. Álvarez and M. Gómez-Río, Computer-aided diagnosis of Alzheimer's type dementia combining support vector machines and discriminant set of features, *Information Sciences* **237** (2013) 59–72.
14. R. Chaves, J. Ramírez, J. Górriz, M. López, D. Salas-Gonzalez, I. Álvarez and F. Segovia, SVM-based computer-aided diagnosis of the Alzheimer's disease using t-test NMSE feature selection with feature correlation weighting, *Neuroscience Letters* **461**(3) (2009) 293–297.
15. D. Salas-Gonzalez, J. M. Górriz, J. Ramírez, M. López, I. Álvarez, F. Segovia, R. Chaves and C. G. Puntonet, Computer-aided diagnosis of Alzheimer's disease using support vector machines and classification trees, *Physics in Medicine and Biology* **55**(10) (2010) 2807–2817.
16. I. Álvarez, J. Górriz, J. Ramírez, D. Salas-Gonzalez, M. López, C. Puntonet and F. Segovia, Alzheimer's diagnosis using eigenbrains and support vector machines, *Electronics Letters* **45**(7) (2009) p. 342.
17. B. Koley and D. Dey, An ensemble system for automatic sleep stage classification using single channel EEG signal, *Computers in Biology and Medicine* **42**(12) (2012) 1186–1195.
18. K. a. I. Aboalayon, H. T. Ocbagabir and M. Faezipour, Efficient sleep stage classification based on EEG signals, *IEEE Long Island Systems, Applications and Technology (LISAT) Conference 2014* (2014) 1–6.
19. D. López-García, M. Ruz, J. Ramírez and J. M. Górriz, Automatic detection of sleep disorders: Multi-class automatic classification algorithms based on Support Vector Machines, *International Conference on Time Series and Forecasting (ITISE 2018)*, **3**2018, pp. 1270–1280.
20. J. T. Kaplan, K. Man and S. G. Greening, Multivariate cross-classification: applying machine learning techniques to characterize abstraction in neural representations., *Frontiers in human neuroscience* **9** (2015) p. 151.
21. L. Su, E. Fonteneau, W. Marslen-Wilson and N. Kriegeskorte, Spatiotemporal searchlight representational similarity analysis in EMEG source space, *Proceedings - 2012 2nd International Workshop on Pattern Recognition in NeuroImaging, PRNI 2012* (2012) 97–100.
22. M. Cauchoix, G. Barragan-Jason, T. Serre and E. J. Barbeau, The Neural Dynamics of Face Detection in the Wild Revealed by MVPA, *The Journal of Neuroscience* **34**(3) (2014) 846–854.
23. C. Kerrén, J. Linde-Domingo, S. Hanslmayr and M. Wimber, An Optimal Oscillatory Phase for Pattern Reactivation during Memory Retrieval., *Current biology : CB* **28**(21) (2018) 3383–3392.
24. M. N. Hebart, B. B. Bankson, A. Harel, C. I. Baker and R. M. Cichy, The representational dynamics of task and object processing in humans, *eLife* **7** (2018) p. e32816.
25. J. M. Correia, B. Jansma, L. Hausfeld, S. Kikkert and M. Bonte, EEG decoding of spoken words in bilingual listeners: From words to language invariant semantic-conceptual representations, *Frontiers in Psychology* **6** (2015) p. 71.
26. M. X. Cohen and T. H. Donner, Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior, *Journal of Neurophysiology* **110**(12) (2013) 2752–2763.
27. D. López-García, A. Sobrado, J. M. González-Peñalver, J. M. Górriz and M. Ruz, Multivariate Pattern Analysis of Electroencephalography Data in a Demand-Selection Task, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2019) 403–411.
28. M. Kleiner, D. Brainard, D. Pelli, A. Ingling, R. Murray, C. Broussard and Others, What's new in Psychtoolbox-3, *Perception* **36**(14) (2007).
29. L. P. Gibbons, InpOut32 is an open source windows DLL and Driver to give direct access to hardware ports.

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

30. A. Delorme and S. Makeig, EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, *Journal of neuroscience methods* **134**(1) (2004) 9–21.
31. M. Chaumon, D. V. Bishop and N. A. Busch, A practical guide to the selection of independent components of the electroencephalogram for artifact correction, *Journal of Neuroscience Methods* **250** (2015) 47–63.
32. J. Gross, S. Baillet, G. R. Barnes, R. N. Henson, A. Hillebrand, O. Jensen, K. Jerbi, V. Litvak, B. Maess, R. Oostenveld, L. Parkkonen, J. R. Taylor, V. van Wassenhove, M. Wibral and J. M. Schoffelen, Good practice for conducting and reporting MEG research, *NeuroImage* **65** (2013) 349–363.
33. A. Keil, S. Debener, G. Gratton, M. Junghöfer, E. S. Kappenman, S. J. Luck, P. Luu, G. A. Miller and C. M. Yee, Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography, *Psychophysiology* **51**(1) (2014) 1–21.
34. N. Schouppe, J. Demanet, C. N. Boehler, K. R. Ridderinkhof and W. Notebaert, The Role of the Striatum in Effort-Based Decision-Making in the Absence of Reward, *Journal of Neuroscience* **34**(6) (2014) 2148–2154.
35. L. Isik, E. M. Meyers, J. Z. Leibo and T. Poggio, The dynamics of invariant object recognition in the human visual system, *Journal of Neurophysiology* **111**(1) (2014) 91–102.
36. T. Grootswagers, S. G. Wardle and T. A. Carlson, Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data, *Journal of Cognitive Neuroscience* **29**(4) (2017) 677–697.
37. H. Abdi and L. J. Williams, Principal component analysis, *Wiley Interdisciplinary Reviews: Computational Statistics* **2**(4) (2010) 433–459.
38. M. N. Hebart, B. B. Bankson, A. Harel, C. I. Baker and R. M. Cichy, The representational dynamics of task and object processing in humans, *eLife* **7** (2018) 1–21.
39. I. Jolliffe, *Principal component analysis* (Springer, 2011).
40. E. Combrisson and K. Jerbi, Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy, *Journal of Neuroscience Methods* **250** (2015) 126–136.
41. J.-R. King and S. Dehaene, Characterizing the dynamics of mental representations: the temporal generalization method, *Trends in Cognitive Sciences* **18**(4) (2014) 203–210.
42. G. Varoquaux, Cross-validation failure: Small sample sizes lead to large error bars, *NeuroImage* **180** (2018) 68–77.
43. J. M. Górriz, J. Ramirez and J. Suckling, On the computation of distribution-free performance bounds: Application to small sample sizes in neuroimaging, *Pattern Recognition* **93** (2019) 1–13.
44. J. M. Gorriz, S. Group and C. Neuroscience, Statistical Agnostic Mapping: a Framework in Neuroimaging based on Concentration Inequalities, *bioRxiv* (2019).
45. U. Braga-Neto, R. Hashimoto, E. R. Dougherty, D. V. Nguyen and R. J. Carroll, Is cross-validation better than resubstitution for ranking genes?, *Bioinformatics* **20**(2) (2004) 253–258.
46. J. M. Górriz, J. Ramírez, F. Segovia, F. J. Martínez, M. C. Lai, M. V. Lombardo, S. Baron-Cohen and J. Suckling, A machine learning approach to reveal the neurophenotypes of autisms, *International Journal of Neural Systems* **29**(7) (2019) 1–22.
47. F. J. Martinez-Murcia, A. Ortiz, J.-M. Gorriz, J. Ramirez and D. Castillo-Barnes, Studying the Manifold Structure of Alzheimer's Disease: A Deep Learning Approach Using Convolutional Autoencoders, *IEEE Journal of Biomedical and Health Informatics* **24**(1) (2020) 17–26.
48. J. A. Etzel, V. Gazzola and C. Keysers, Testing Simulation Theory with Cross-Modal Multivariate Classification of fMRI Data, *PLoS ONE* **3**(11) (2008) p. e3690.
49. N. N. Oosterhof, S. P. Tipper and P. E. Downing, Crossmodal and action-specific: neuroimaging the human mirror neuron system, *Trends in Cognitive Sciences* **17**(7) (2013) 311–318.
50. N. N. Oosterhof, A. J. Wiggett, J. Diedrichsen, S. P. Tipper and P. E. Downing, Surface-Based Information Mapping Reveals Crossmodal VisionAction Representations in Human Parietal and Occipitotemporal Cortex, *Journal of Neurophysiology* **104**(2) (2010) 1077–1089.
51. J. van den Hurk and H. P. O. de Beeck, Generalization asymmetry in multivariate cross-classification: When representation A generalizes better to representation B than B to A, *bioRxiv* (2019).
52. J. Stelzer, Y. Chen and R. Turner, Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control, *NeuroImage* **65** (2013) 69–82.
53. T. E. Nichols and A. P. Holmes, Nonparametric permutation tests for functional neuroimaging: a primer with examples., *Human brain mapping* **15**(1) (2002) 1–25.
54. P. Díaz-Gutiérrez, J. E. Arco, S. Alguacil, C. González-García and M. Ruz, Neural representation of social expectations during interpersonal decisions, *bioRxiv* (2018).
55. J. E. Arco, C. González-García, P. Díaz-Gutiérrez, J. Ramírez and M. Ruz, Influence of activation pattern estimates and statistical significance tests in fMRI decoding analysis, *Journal of Neuroscience Methods* **308** (2018) 248–260.

ACCEPTED MANUSCRIPT

Accepted manuscript to appear in IJNS

56. A. F. Palenciano, C. González-García, J. E. Arco, L. Pessoa and M. Ruz, Representational Organization of Novel Task Sets during Proactive Encoding, *The Journal of Neuroscience* **39**(42) (2019) 8386–8397.
57. R. Koenig-Robert and J. Pearson, Decoding the contents and strength of imagery before volitional engagement, *Scientific Reports* **9**(1) (2019) p. 3504.
58. J. Hubbard, A. Kikumoto and U. Mayr, EEG Decoding Reveals the Strength and Temporal Dynamics of Goal-Relevant Representations, *Scientific Reports* **9**(1) (2019) p. 9051.
59. F. Meconi, J. Linde-Domingo, C. S. Ferreira, S. Michelmann, B. Staresina, I. Apperly and S. Hanslmayr, Autobiographical memory reactivation in empathy, *bioRxiv* (2019).
60. E. Maris and R. Oostenveld, Nonparametric statistical testing of EEG- and MEG-data, *Journal of Neuroscience Methods* **164**(1) (2007) 177–190.
61. Y. Benjamini and Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, *Journal of the Royal Statistical Society: Series B (Methodological)* **57**(1) (1995) 289–300.
62. Y. Benjamini and W. Liu, A step-down multiple hypotheses testing procedure that controls the false discovery rate under independence, *Journal of Statistical Planning and Inference* **82**(1-2) (1999) 163–170.
63. A. Eklund, T. E. Nichols and H. Knutsson, Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates, *Proceedings of the National Academy of Sciences* **113**(28) (2016) 7900–7905.
64. C. M. Bennett, G. L. Wolford and M. B. Miller, The principled control of false positives in neuroimaging, *Social Cognitive and Affective Neuroscience* 4(4) (2009) 417–422.
65. P. Fries, A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence, *Trends in Cognitive Sciences* **9**(10) (2005) 474–480.
66. T. Harmony, The functional significance of delta oscillations in cognitive processing, *Frontiers in Integrative Neuroscience* **7** (2013) 1–10.
67. V. Van Veen and C. S. Carter, The anterior cingulate as a conflict monitor: FMRI and ERP studies, *Physiology and Behavior* **77**(4-5) (2002) 477–482.
68. I. Sols, S. DuBrow, L. Davachi and L. Fuentemilla, Event Boundaries Trigger Rapid Memory Reinstatement of the Prior Events to Promote Their Representation in Long-Term Memory, *Current Biology* **27**(22) (2017) 3499–3504.
69. M. Bernert and B. Yvert, An Attention-Based Spiking Neural Network for Unsupervised Spike-Sorting, *International Journal of Neural Systems* **29**(08) (2019).
70. R. Hu, Q. Huang, H. Wang, J. He and S. Chang, Monitor-Based Spiking Recurrent Network for the Representation of Complex Dynamic Patterns, *International Journal of Neural Systems* **29**(08) (2019) p. 1950006.
71. S. Ghosh-Dastidar and H. Adeli, A new supervised learning algorithm for multiple spiking neural networks with application in epilepsy and seizure detection, *Neural Networks* **22**(10) (2009) 1419–1431.
72. S. Ghosh-Dastidar and H. Adeli, Improved spiking neural networks for EEG classification and epilepsy and seizure detection, *Integrated Computer-Aided Engineering* **14**(3) (2007) 187–212.
73. M. Ahmadlou and H. Adeli, Enhanced probabilistic neural network with local decision circles: A robust classifier, *Integrated Computer-Aided Engineering* **17**(3) (2010) 197–210.
74. M. H. Rafiei and H. Adeli, A New Neural Dynamic Classification Algorithm, *IEEE Transactions on Neural Networks and Learning Systems* **28**(12) (2017) 3074–3083.
75. F. Pereira, T. Mitchell and M. Botvinick, Machine learning classifiers and fMRI: A tutorial overview, *NeuroImage* **45**(1) (2009) 199–209.
76. M. N. Hebart, Y. Schriever, T. H. Donner and J.-D. Haynes, The Relationship between Perceptual Decision Variables and Confidence in the Human Brain, *Cerebral Cortex* **26**(1) (2016) 118–130.
77. K. Korjus, M. N. Hebart and R. Vicente, An Efficient Data Partitioning to Improve Classification Performance While Keeping Parameters Interpretable., *PloS one* **11**(8) (2016) p. e0161788.

Int. J. Neur. Syst. Downloaded from www.worldscientific.com
by UNIVERSIDAD DE GRANADA on 05/14/20. Re-use and distribution is strictly not permitted, except for Open Access articles.

ACCEPTED MANUSCRIPT