

## P3b amplitude as a signature of cognitive decline in the older population: An EEG study enhanced by Functional Source Separation

Camillo Porcaro<sup>a,b,c,d,1</sup>, Joshua Henk Balsters<sup>b,e,f,\*</sup>, Dante Mantini<sup>b,c,g</sup>, Ian H. Robertson<sup>e</sup>, Nicole Wenderoth<sup>b,c</sup>

<sup>a</sup> Institute of Cognitive Sciences and Technologies (ISTC) - National Research Council (CNR), Rome, Italy

<sup>b</sup> Neural Control of Movement Lab, Department of Health Sciences and Technology ETH Zurich, Switzerland

<sup>c</sup> Research Center for Motor Control and Neuroplasticity, KU Leuven, Leuven, Belgium

<sup>d</sup> S. Anna Institute and Research in Advanced Neurorehabilitation (RAN) Crotona, Italy

<sup>e</sup> Global Brain Health Institute, Trinity College Institute of Neuroscience, Trinity College, Dublin, Ireland

<sup>f</sup> Department of Psychology, Royal Holloway University of London, Egham, Surrey, UK

<sup>g</sup> Functional Neuroimaging Laboratory, IRCCS San Camillo Hospital Foundation, Venice, Italy

### ARTICLE INFO

#### Keywords:

P300  
P3a/P3b  
Aging  
Cognitive impairment  
Functional Source Separation (FSS)  
Electroencephalography (EEG)  
Support vector machine learning

### ABSTRACT

With the greying population, it is increasingly necessary to establish robust and individualized markers of cognitive decline. This requires the combination of well-established neural mechanisms, and the development of increasingly sensitive methodologies. The P300 event-related potential (ERP) has been one of the most heavily investigated neural markers of attention and cognition, and studies have reliably shown that changes in the amplitude and latency of the P300 ERP index the process of aging. However, it is still not clear whether either the P3a or P3b sub-components additionally index levels of cognitive impairment. Here, we used a traditional visual three-stimulus oddball paradigm to investigate both the P3a and P3b ERP components in sixteen young and thirty-four healthy elderly individuals with varying degrees of cognitive ability. EEG data extraction was enhanced through the use of a novel signal processing method called Functional Source Separation (FSS) that increases signal-to-noise ratio by using a weighted sum of all electrodes rather than relying on a single, or a small sub-set, of EEG channels. Whilst clear differences in both the P3a and P3b ERPs were seen between young and elderly groups, only P3b amplitude differentiated older people with low memory performance relative to IQ from those with consistent memory and IQ. A machine learning analysis showed that P3b amplitude (derived from FSS analysis) could accurately categorise high and low performing elderly individuals (78% accuracy). A comparison of Bayes Factors found that differences in cognitive decline within the elderly group were 87 times more likely to be detected using FSS compared to the best performing single electrode (Cz). In conclusion, we propose that P3b amplitude could be a sensitive marker of early, age-independent, episodic memory dysfunction within a healthy older population. In addition, we advocate for the use of more advanced signal processing methods, such as FSS, for detecting subtle neural changes in clinical populations.

### 1. Introduction

The P300 was one of the first event-related potentials (ERPs) to be reported (Sutton et al., 1965), and its discovery is largely credited for the application of EEG methods to cognitive neuroscience (Polich, 2012). The P300 ERP presents as a large positive waveform with a centro-parietal topography peaking approximately 300 ms after stimulus onset. Although the term P300 (or P3) is often used in the literature, it is

important to clarify that the P300 encompasses at least two distinct subcomponents; the P3a and P3b (Squires et al., 1975). Whilst the P3b has a parietal topography, the P3a presents with a more anterior focus peaking around central electrodes most likely indicating that the P3a and P3b ERP components are underpinned by distinct neural sources (Linden, 2005; Polich, 2003). In addition, the morphology of P3a and P3b ERPs also differ. Whilst the peak of the P3a is relatively consistent across trials, the P3b peak is response-locked and as such it changes from trial-to-trial

\* Corresponding author. Department of Psychology, Royal Holloway University of London, Egham, Surrey, TW20 0EX, UK.

E-mail address: [Joshua.balsters@rhul.ac.uk](mailto:Joshua.balsters@rhul.ac.uk) (J.H. Balsters).

<sup>1</sup> equal contribution.

<https://doi.org/10.1016/j.neuroimage.2018.09.057>

Received 5 July 2018; Received in revised form 3 September 2018; Accepted 20 September 2018

Available online 21 September 2018

1053-8119/© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

depending on the speed of the response (Hillyard et al., 1971; Squires et al., 1973). When P3a and P3b trials are averaged across subjects, this often means the grand average P3a presents with a clear peak around 300 ms (due to the consistent single trial peaks), whereas the P3b has a wide spread morphology ranging between 300 and 600 ms. These differences in scalp topography and ERP morphology suggest that the P3a and P3b might sub-serve different cognitive processes (Linden, 2005; Polich, 2007, 2003). The P3a is often driven by rare non-target (distractor) stimuli, and as such it is believed that the P3a ERP represents stimulus-driven frontal attention mechanisms (Linden, 2005; Polich, 2007). The P3b is often elicited by target detection paradigms (most famously the oddball paradigm), however there is still ongoing debate about whether the P3b encodes allocation of attentional resources (Nieuwenhuis et al., 2011), context and memory updating (Donchin and Coles, 1988; Polich, 2007), the uncertainty of the stimulus (Mars et al., 2008), or the accumulation of sensory evidence (Kelly and O'Connell, 2015).

Along with being one of the most investigated ERPs in typical populations, the P300 has proven to be a useful marker of clinical pathology. Perturbations of the P300 have been shown in ADHD, Autism, Depression, Obsessive Compulsive Disorder, and Parkinson's disease to name but a few (Polich and Criado, 2006). However, the largest number of P300 studies have focussed on how the P300 ERP changes with age and whether this might be a marker of cognitive decline seen in Mild Cognitive Impairment (MCI) and Alzheimer's disease (AD) (Rossini et al., 2007). Polich (1997) previously showed that increases in P300 latency (which we assume to be the P3b based on the paradigm) and decreases in P300 amplitude correlate with age (20–80yrs;  $n = 120$ , 10 male and 10 female per decade). Walhovd and Fjell (Walhovd and Fjell, 2001) expanded on this by using a combination of two- and three-stimulus oddball paradigms to distinguish the P3a and P3b ERPs. They similarly found that both P3a and P3b peak latency increase with age and P3a and P3b peak amplitude decrease with age. Studies by Polich and Corey-Bloom (2005) and Van Deursen et al. (van Deursen et al., 2009) have shown that the amplitude of P300 responses further decreased whilst peak latency further increased in both MCI and AD compared to matched elderly controls.

Compensatory neural mechanisms that occur with aging, along with increasing frontalization of brain responses, have been formalized in the Compensation-Related Utilization of Neural Circuits Hypothesis (CRUNCH; (Reuter-Lorenz and Cappell, 2008)) and Posterior-Anterior Shift in Aging (PASA; (Davis et al., 2008)) models respectively. Some studies have suggested that this increased frontal activity in elderly individuals is important for task performance (Davis et al., 2008; Goh and Park, 2009), however studies by Fabiani et al. (Fabiani et al., 1998) and West et al. (2010) have also shown that elderly individuals who show a larger frontal P300 response perform more poorly on neuropsychological tests of executive functions. Although it is not clear whether the increased frontal activity that occurs with aging is a help or a hindrance, it is clear that understanding age-related changes and cognitive decline requires investigating both frontal and parietal mechanisms in parallel rather than just exploring single electrodes in isolation.

EEG source reconstruction suffers from the fact that multiple potential source configurations result in the same measured EEG potentials. This issue is especially problematic when a very low number of electrodes are used, which is often the case in clinical applications. For this reason, EEG studies investigating the P300 in aging, MCI, and AD have avoided source localization and relied upon one single electrode or a localised subset of electrodes. However, the generators of EEG activity cannot be reliably inferred on the basis of a-priori selected single channels, or a limited group of channels, due to the electric/magnetic field propagation problem (Rusiniak et al., 2013; Siegel et al., 2012). Moreover, using information coming from only one electrode can be misleading especially when the activated network is spread among the entire scalp, as in the case of the P300. As such, this traditional approach of using a limited number of electrodes to investigate age-related changes may lack the

sensitivity to detect the neural changes underpinning cognitive decline and thus only pick up on the much clearer age-related changes in ERP activity. In order to overcome this issue, we propose using a novel signal processing method - Functional Source Separation (FSS - (Porcaro and Tecchio, 2014; Porcaro et al., 2009a, 2008; Tecchio et al., 2007). FSS overcomes an important limitation of most previous studies by using the best combination of all available electrodes to detect the P300 generators. This approach has already been applied to extract EEG specific features for primary motor (Porcaro et al., 2008), primary sensory (Porcaro et al., 2009a,b; 2008), and primary visual (Porcaro et al., 2011, 2010) areas. We believe this approach will provide much richer information than traditional EEG approaches, especially given the larger network of brain regions involved in producing the P3a and P3b responses.

Here for the first time, we have used FSS as a tool to investigate the P3a and P3b ERP responses in young and older individuals. Here we propose two complementary hypotheses: 1) differences in P300 ERP responses will correspond with individual differences in cognitive decline (based on neuropsychological measurements) in the older population, and 2) the FSS approach (utilizing a weighted combination of all electrodes) will be more sensitive to these differences than traditional approaches based on individual electrodes.

## 2. Materials and methods

### 2.1. Participants

Sixteen young (18–28 years old; mean age,  $22.4 \pm 3.28$  years) and thirty-four elderly individuals (65–78 years old; mean age,  $70.22 \pm 4.31$  years) were recruited for this study. Data from two elderly participants had to be discarded due to technical problems and poor EEG data quality, and one young and one additional elderly participant were excluded for poor performance on the task ( $d\text{-prime} < 1$ ). Table 1 shows demographic details for young and old participants. All participants were right-handed according to the Edinburgh Handedness Questionnaire (Oldfield, 1971). Participants gave written informed consent before the study, which was approved by the Trinity College Dublin School of Psychology Ethics Committee. On a day separate from the EEG testing (elderly group:  $163.43 \pm 118$  days; young group:  $161 \pm 81$  days), participants also underwent a neuropsychological battery consisting of the Mini Mental State Examination (MMSE; (Folstein et al., 1975)), the National Adult Reading Test (NART; estimate of intelligence; (Nelson, 1982)), the Stroop test, category fluency (animal), the Logical Memory subtest of the Wechsler Memory Scale III (WMS; (Wechsler, 1997)), and the Hospital Anxiety and Depression Scale (Zigmond and Snaith, 1983). Participants who scored more than 8 on either the anxiety or depression subscales of the Hospital Anxiety and Depression Scale were excluded from the study. Participants were not taking any psychiatric or neurological medications at the time of testing.

We subdivided old participants into cognitively high performing (HP) and low performing (LP) individuals as suggested by previous research (Dockree et al., 2015). To do so, logical memory delayed recall scores and NART estimated IQ scores were converted into Z scores. LP individuals were identified as having a logical memory score one standard deviation below their NART estimated IQ scores. This approach of baselining memory scores using NART IQ has been shown to be more sensitive than using raw values to categorize individuals (Dockree et al., 2015).

### 2.2. EEG recording

EEG recordings were acquired with a 32-channel BrainAmp system (BrainProducts, Munich, Germany). Thirty-three EEG electrodes were placed on the scalp, including the reference electrode positioned at FCz and the ground electrode placed at position AFz. One external electrode was applied to the subjects back to acquire the electrocardiogram (ECG). Electrode impedances were maintained at less than 10 k $\Omega$ . Data were

Table 1

**Participant demographics for young compared to elderly individuals.** T values were generated using Students two sample *t*-test (no variables showed a significant difference in Levene's test of equality of variance). Significant differences between young and old individuals are highlighted in bold ( $p < 0.05$ ). Effects size (Cohen's *d*) and Bayes Factor are also reported to highlight the magnitude of the effect. Cohen's *d* values between 0.2 and 0.5 are considered small, 0.5–0.8 medium, and greater than 0.8 are considered large. Bayes Factors between 1 and 3 are considered anecdotal, 3–10 moderate, and greater than 10 are considered strong.

	Young (n = 15)		Old (n = 31)		t	df	p	Cohen's d	BF <sub>10</sub>
	mean	SD	mean	SD					
Gender	4/15 female		18/31 female						
Age (yrs)	22.4	3.418	70	4.195	38.173	44	<.001	12.006	3.923e +31
Years of Education	16.357	2.951	14.5	3.012	-1.896	40	0.065	-0.621	1.265
HADS anxiety	4.143	2.248	3.667	2.905	-0.541	42	0.591	-0.175	0.353
HADS depression	1.357	1.737	2.367	1.79	1.758	42	0.086	0.569	1.043
MMSE	29.2	0.862	28	1.713	-2.551	44	<b>0.014</b>	-0.802	3.743
Log Memory (Zscore)	0.573	0.914	0.511	0.979	-0.207	44	0.837	-0.065	0.313
NART (Zscore)	0.952	0.35	1.253	0.467	2.207	44	<b>0.033</b>	0.694	2.03
LogMem – NART	-0.379	1.063	-0.742	0.83	-1.268	44	0.211	-0.399	0.58
<b>Oddball performance</b>									
Hits (%)	84.388	13.251	88.94	12.706	1.124	44	0.267	0.353	0.507
False Alarm (%)	3.513	4.516	7.772	10.165	1.544	44	0.13	0.486	0.785
D-Prime	3.151	0.89	2.902	0.756	-0.988	44	0.328	-0.311	0.453
RT (ms)	485.684	58.566	509.605	78.189	1.049	44	0.3	0.33	0.475
RT variability	0.245	0.077	0.3	0.115	1.669	44	0.102	0.525	0.917

recorded on a laptop computer using Brain Recorder v1.04 software (BrainProducts, Munich, Germany) at a sampling rate of 5 kHz with a band-pass filter of 0.016–250 Hz.

### 2.3. Experimental procedure

Participants performed a 3-stimulus visual oddball task previously described in O'Connell et al. (2012a). Every 2075 ms a stimulus appeared on the screen for 75 ms. Standard stimuli consisted of a 3.5-cm diameter purple circle and appeared on 80% of trials. Target stimuli were a slightly larger purple circle (4 cm diameter) and appeared on 10% of trials. Distractor stimuli were a black and white checkerboard and appeared on 10% of trials. Participants were asked to make a speeded response to target stimuli using the response box placed in their right hand. The stimulus array was pseudo randomly designed such that between 3 and 5 standard stimuli were presented after any target or distractor stimulus. Thus, the minimum interval between a P300 eliciting events was 8300 ms but because targets and distractors were randomly interspersed throughout the task the interval between 2 targets or between 2 distractors could be as high as 64 s. The average was approximately 20 s, and approximately 70% of trials occurred between 8 and 22 s. All stimuli were presented on a grey background and participants were asked to maintain fixation on a white cross-presented at the centre of the screen.

### 2.4. EEG pre-processing

The data were re-referenced to common average and down-sampled to 512 Hz. The data were bandpass filtered (0.5–30 Hz) prior to the off-line analysis. Whilst it has been argued that a highpass filter  $> 1$  Hz is preferable for ICA, this filter setting significantly reduces the amplitude of the P300 ERP, which has an optimal highpass filter of 0.1 Hz (Kapenman & Luck, 2010). For this reason, a highpass filter setting of 0.5 Hz was chosen, as in previous studies (Murphy et al., 2011). The first step of the analysis we have employed is a semiautomatic independent component analysis (ICA)-based procedure to identify and remove cardiac and/or ocular artefacts without rejecting the contaminated epochs (Barbati et al., 2004; Porcaro et al., 2015).

### 2.5. Functional source separation

Functional source separation (FSS – (Barbati et al., 2008, 2006, Porcaro et al., 2009a,b; 2008; Porcaro et al., 2011, 2010; Tecchio et al., 2007) is a semi-blind source separation method (Porcaro and Tecchio, 2014) which uses some well-known distinctive features of

electrophysiological signals. The aim of FSS is to enhance the separation of relevant signals by exploiting a priori knowledge without renouncing the advantages of using only information contained in original signal waveforms. FSS, analogous to ICA, models the set of EEG recorded signals  $\mathbf{x}$  as a linear combination of an equal number of sources  $\mathbf{s}$  via a mixing matrix  $\mathbf{A}$ . Differing from other constrained ICA models (Khan et al., 2012; Lu and Rajapakse, 2005; Wang and James, 2007), FSS identifies a single source at a time, building a contrast function for that source that exploits *fingerprint* information associated to the neuronal pool to be identified (Porcaro and Tecchio, 2014). In general, FSS starts from the original EEG data matrix  $\mathbf{x}$  for each source, and returns one functional source (FS) with the required functional property. This scheme gives us the possibility to extract the FS that maximizes the functional behaviour in agreement with the functional constraint (Porcaro and Tecchio, 2014). A modified cost function (with respect to standard ICA) is defined as:  $\mathbf{F} = \mathbf{J} + \lambda \mathbf{R}$  where  $\mathbf{J}$  is the statistical constraint normally used in ICA, while  $\mathbf{R}$  accounts for the a priori information known about the sources. The relative weight of these two parameters can be adjusted via  $\lambda$  ((Porcaro et al., 2008) - Appendix A).  $\lambda$  has been chosen to both minimize computational time and maximize the functional constraint  $\mathbf{R}$ . Moreover, the FSS contrast function  $\mathbf{F}$  is optimized by means of simulated annealing (Kirkpatrick et al., 1983), thus allowing prior information about the FS to be described by a non-differentiable function.

Our study aimed to investigate the activity on P300 in both young and old participants. Thus, we first identified the functional source underlying the P300 processes maximizing the P300 response (around the following time window [320–480 ms] for both young and old participants) named  $\text{FS}_{\text{P300}}$ . The functional constraints used were defined as follows:

$$R_{\text{FS}_{\text{P300}}} = \sum_{t_1}^{t_2} |EA(t)| - \sum_{t=-500}^0 |EA(t)|$$

with the evoked activity, EA, computed by averaging signal epochs of the source  $\text{FS}_{\text{P300}}$ , triggered on the visual stimulation ( $t = 0$ );  $t_1$  is the lower bound (320 ms) and  $t_2$  is the upper bound (480 ms) of the window chosen to maximize the temporal *fingerprint* of the source. The baseline was computed in the time interval from  $-500$  to 0 ms.

### 2.6. Functional sources behaviour

Once the sources ( $\text{FS}_{\text{P300}}$ ) were extracted, ERP analysis and topographic distribution were calculated in both young and old participants to characterize and validate the source extracted. In particular, ERP

analysis was performed on P3a and P3b to investigate different behaviours in the two groups (Young vs. Old) and ‘High Performing’ vs. ‘Low Performing’ (HP vs. LP) Elderly groups. Target trials (P3b) and distractor trials (P3a) were epoched (–500 to 1000 ms) separately and were baseline corrected for the interval from –500 to 0 ms. For each subject, single trials were averaged excluding error trials (i.e. missing response for Target trials, or a response for Distractor trials). We performed pointwise statistical analysis on the averaged target (P3b) and distractor (P3a) waveforms conducting two-sample permutation *t*-tests (10,000 permutations) on every point of the ERP waveform (0–1000 ms). We used false discovery rate (FDR) to correct for multiple comparisons. In addition, peak amplitude and latency values were extracted and analysed using a combination of frequentist and Bayesian approaches. These analyses were conducted in JASP (The JASP Team, 2017) v0.8.2 - <https://jasp-stats.org>.

### 2.7. Support vector machine (SVM) learning

We trained separate sets of classifiers to predict from EEG data whether an individual belonged to the Young vs HP, Young vs LP, or HP vs LP groups. Each set of classifiers used peak amplitude and latency as the only features, but these features were derived from either the P3a or P3b ERP which were extracted from (i) FSS data, or from the single electrodes (ii) Fz, (iii) Cz, or (iv) Pz. In total we implemented eight separate sets of classifiers (P3a vs P3b x FSS vs Fz vs Cz vs Pz) using linear support vector machines (SVM, default parameters of Matlab R2016a). Accuracies were determined using a leave-one-subject-out cross-validation scheme. For example, for the LP versus HP classification we trained the classifier on a test dataset consisting of thirty participants and validated it on one participant who has been left out. This procedure was repeated 31 times (i.e. leaving out each of the participants) and the overall accuracy was calculated as the percentage of participants which were accurately classified during the validation step. Feature permutation was used to assess the significance of the classifier performance under the null-hypothesis of independence between features and labels. The *p*-value was calculated as the fraction of classification accuracies from the permuted dataset larger than accuracies obtained from the original dataset, here based on 1000 such permutations (Ojala and Garriga, 2009). Subsequently, all *p*-values were corrected for multiple comparisons using the false discovery rate to account for the fact that we used multiple classifiers.

**Table 2**

**Participant demographics for HP compared to LP individuals.** *T* values were generated using Students two sample *t*-test (an asterisk indicates a significant difference in Levene's test of equality of variance in which case Welch's *t*-test was used). Significant differences between HP and LP individuals are highlighted in bold ( $p < 0.05$ ). Effect sizes (Cohen's *d*) and Bayes Factor are also reported to highlight the magnitude of the effect. Cohen's *d* values between 0.2 and 0.5 are considered small, 0.5–0.8 medium, and greater than 0.8 are considered large. Bayes Factors between 1 and 3 are considered anecdotal, 3–10 moderate, and greater than 10 are considered strong.

	HP (n = 17)		LP (n = 14)		<i>t</i>	df	<i>p</i>	Cohen's <i>d</i>	BF <sub>10</sub>
	mean	SD	mean	SD					
Gender	8/17 female		10/14 female						
Age (yrs)	71.235	4.25	68.5	3.737	–1.881	29	0.07	–0.679	1.263
Years of Education	15.059	2.609	13.636	3.501	–1.232	26	0.229	–0.477	0.627
HADS anxiety	3.176	2.789	4.308	3.038	1.059	28	0.299	0.39	0.528
HADS depression	2.059	1.784	2.769	1.787	1.08	28	0.289	0.398	0.537
MMSE *	28.471	1.068	27.429	2.174	–1.638	18.1	0.119	–0.608	1.054
Log Memory (Zscore)	1.21	0.505	–0.338	0.694	–7.178	29	<b>&lt;.001</b>	–2.591	138051.179
NART (Zscore)	1.296	0.431	1.2	0.519	–0.56	29	0.58	–0.202	0.385
LogMem – NART	–0.086	0.428	–1.538	0.345	–10.245	29	<b>&lt;.001</b>	–3.698	1.421e +8
<b>Oddball performance</b>									
Hits (%)	89.174	9.565	88.656	16.119	–0.111	29	0.912	–0.04	0.343
False Alarm (%) *	5.862	4.665	10.09	14.181	1.069	15.32	0.302	0.401	0.567
D-Prime	2.917	0.728	2.884	0.816	–0.119	29	0.906	–0.043	0.343
RT (ms)	512.146	87.254	506.52	68.695	–0.196	29	0.846	–0.071	0.346
RT variability *	0.262	0.062	0.346	0.147	1.998	16.79	0.062	0.745	1.833

## 3. Results

### 3.1. Oddball performance and neuropsychological differences

Table 1 summarizes demographic information, results from neuropsychological tests and the oddball task performance for young and elderly individuals. In all cases frequentist and Bayesian two sample *t*-tests were performed to establish group differences. Young subjects had significantly more years of education and performed better on the MMSE while old individuals performed significantly better on the NART compared to young individuals. However, a complementary Bayesian analysis found that these results were in the anecdotal range ( $BF_{10} < 3$ ). There were no significant differences in oddball performance (all  $p \geq 0.1$ ;  $BF < 1$ ). In addition, when the old group was split into HP and LP individuals, there were no significant differences in oddball performance (Table 2). The only significant between-group differences were performance on the logical memory subscale of the WMS III, and the difference between the logical memory subscale and the NART.

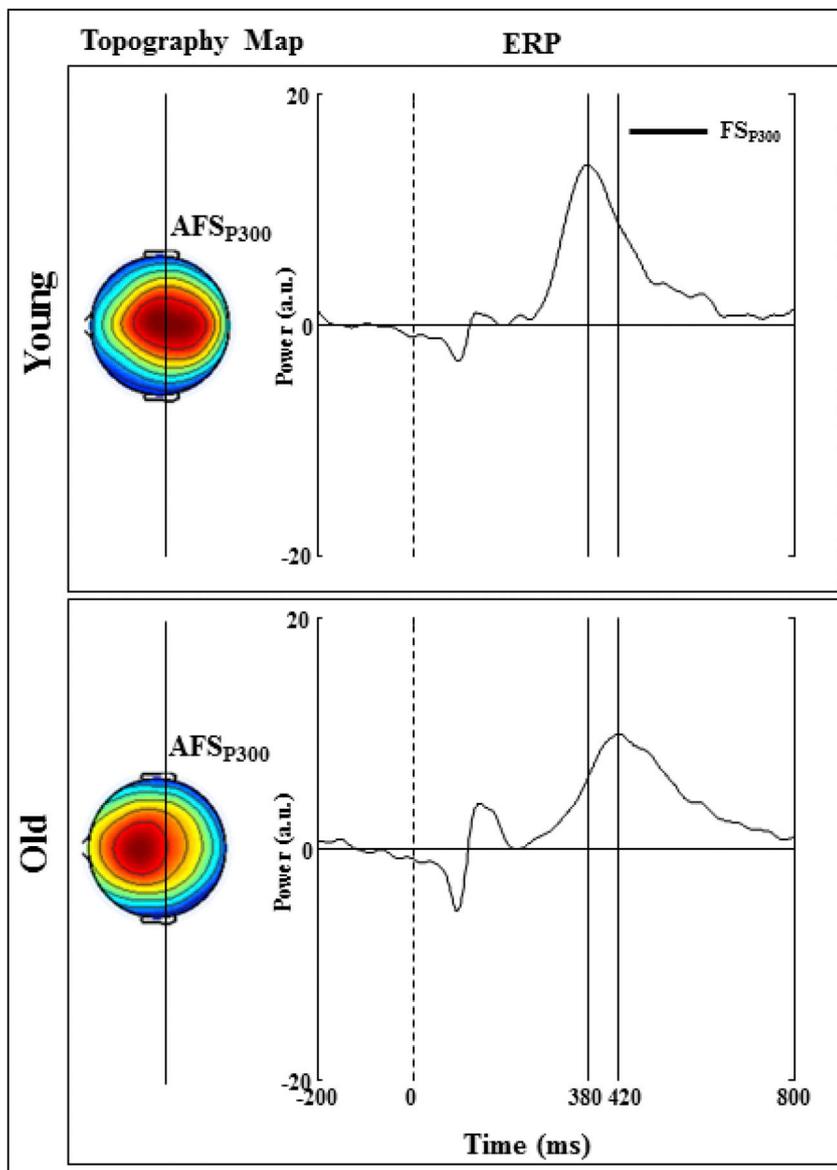
### 3.2. Electrophysiological signals

#### 3.2.1. Functional source behaviour validations

FSS successfully extracted P300 ERP responses in both old (HP and LP) and young participants (FS<sub>P300</sub>). The outcomes of the FSS approach mirror a number of well-established properties of the P300 in aging. First, topographic maps show a posterior to anterior shift in the old group with respect to the young participants (Fig. 1, left column). Second, there was an apparent delay in the peak latency of the FS<sub>P300</sub> ERPs in the old group ( $420 \pm 44$  ms; mean  $\pm$  standard deviation) compared to the young group ( $380 \pm 30$  ms). All of these findings are well established in previous electrophysiology studies of aging and the P300 (Balsters et al., 2013; O'Connell et al., 2012a).

#### 3.2.2. Discrepancy quality check

In order to validate the quality of the FS<sub>P300</sub> source extraction we performed the ERP P300 on Raw Data, reconstructed data by FS<sub>P300</sub> and discrepancy (i.e. the ERP Raw Data minus the ERP obtained reconstructing the data using the source FS<sub>P300</sub> extracted). As shown in Fig. 2, for both groups we were able to identify all the electrical activity at around 380 ms and 420 ms for the P3a and P3b peaks respectively. The discrepancy topography map also showed that the residual activity of the P300 peak is near to zero for both P3a and P3b.



**Fig. 1.** FSS extracted signal between 300 and 600 ms in elderly and young participants.

The topographic map and Event Related Potential (ERP) are shown for the  $FSP_{300}$  source extracted (young top row; old bottom row). In the ERP plot, the vertical dashed line indicates the stimulus onset and the continuous lines indicate the maximum peaks for the P3a and P3b respectively. The vertical continuous line in topographic map indicate maximum topographic peak to emphasize the shift between the two groups.

### 3.2.3. Group differences in P3a and P3b derived using FSS

Fig. 3 shows P3a and P3b ERPs and topographic maps for each group separately (Young, HP, LP). To assess whether there were any significant group differences we began by running a one-way ANOVA on peak latency and amplitude in the P3a and P3b (see Fig. 4). The P3a showed significant main effect of group in both latency ( $F(2,43) = 21.01$ ,  $p < 0.001$ , partial eta squared = 0.494) and amplitude ( $F(2,43) = 5.785$ ,  $p = 0.006$ , partial eta squared = 0.212). In both cases group differences were driven by the young group compared to both HP and LP. Post-hoc t-tests confirmed that young individuals had faster peak latencies in the P3a compared to HP group ( $t(30) = 5.717$ ,  $pFDR < 0.001$ ) and the LP group ( $t(27) = 5.531$ ,  $pFDR < 0.001$ ). There were no significant differences between the HP and LP group in P3a peak latency ( $t(29) = 0.084$ ,  $pFDR = 0.934$ ). Post-hoc t-tests also confirmed that young individuals had larger P3a peak amplitudes compared to HP group ( $t(30) = 2.797$ ,  $pFDR = 0.015$ ) and the LP group ( $t(27) = 3.085$ ,  $pFDR = 0.011$ ). There were no significant differences between the HP and LP group in P3a peak amplitude ( $t(29) = 0.431$ ,  $pFDR = 0.669$ ).

Similar to the P3a, there was a significant main effect of group in both P3b latency ( $F(2,43) = 6.849$ ,  $p = 0.003$ , partial eta squared = 0.242) and P3b amplitude ( $F(2,43) = 12.49$ ,  $p < 0.001$ , partial eta

squared = 0.368). As with P3a peak latency, the significant main effect of group was driven by significantly faster peak latencies in young individuals compared to both the HP group ( $t(30) = 2.764$ ,  $pFDR = 0.017$ ) and the LP group ( $t(27) = 3.522$ ,  $pFDR = 0.003$ ). There was no significant difference in P3b latency for LP compared to HP ( $t(29) = 0.914$ ,  $pFDR = 0.366$ ). Young individuals showed greater P3b peak amplitude compared to HP ( $t(30) = 2.745$ ,  $pFDR = 0.018$ ) and LP individuals ( $t(27) = 4.992$ ,  $pFDR < 0.001$ ), however, P3b peak amplitude was also larger in HP compared to LP ( $t(29) = 2.445$ ,  $pFDR = 0.019$ ).

In summary, whilst latency and amplitude in both the P3a and P3b distinguished young and old individuals, only P3b peak amplitude distinguished between HP and LP individuals. The same pattern of results was clear when ERPs were compared on a time-point basis (see Fig. 5).

### 3.2.4. Individual differences in P3a and P3b derived using FSS

In an effort to move beyond group-level approaches we applied machine learning techniques (SVM) in order to establish whether P300 ERP responses (peak amplitude and latency) were capable of accurately categorizing Young, HP, and LP individuals (see Table 3). Highest overall accuracy to distinguish between the 3 groups was found using FSS P3b data (average 81.25% accuracy across all comparisons using FSS P3b ERP

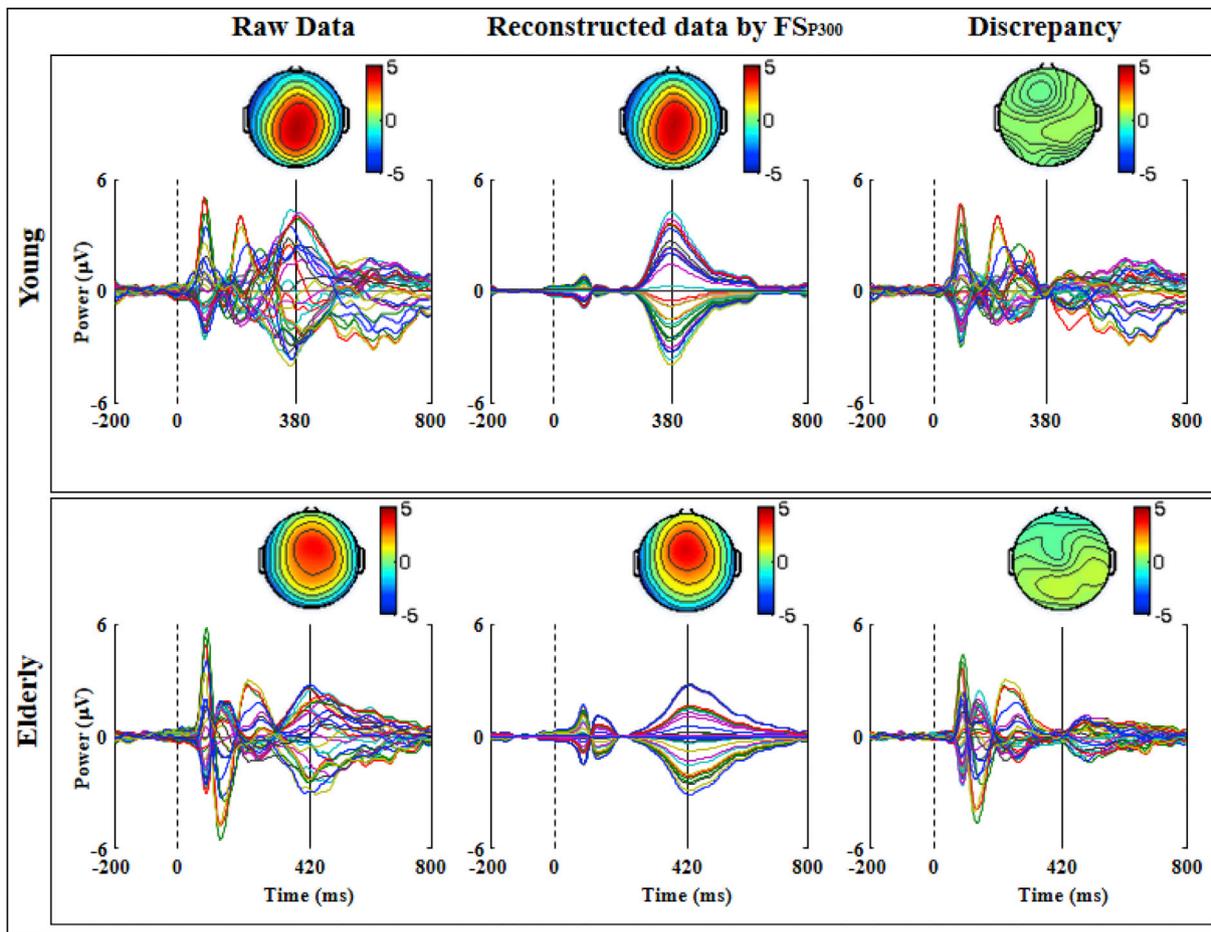


Fig. 2. ERP FS<sub>P300</sub> discrepancy quality check.

ERPs and topographic maps for Young (top row) and Old (bottom row) participants. Top and Bottom left, grand average ERP on the raw data. Top and Bottom middle ERP of reconstructed FS<sub>P300</sub> source. Top and bottom Right, ERP discrepancy i.e. Raw Data minus ERP of reconstructed data by FS<sub>P300</sub>. Topographic maps at 380 ms (Young Group) and 420 ms (Old Group) are also shown for Raw Data, Reconstructed data using FS<sub>P300</sub>, and Discrepancy.

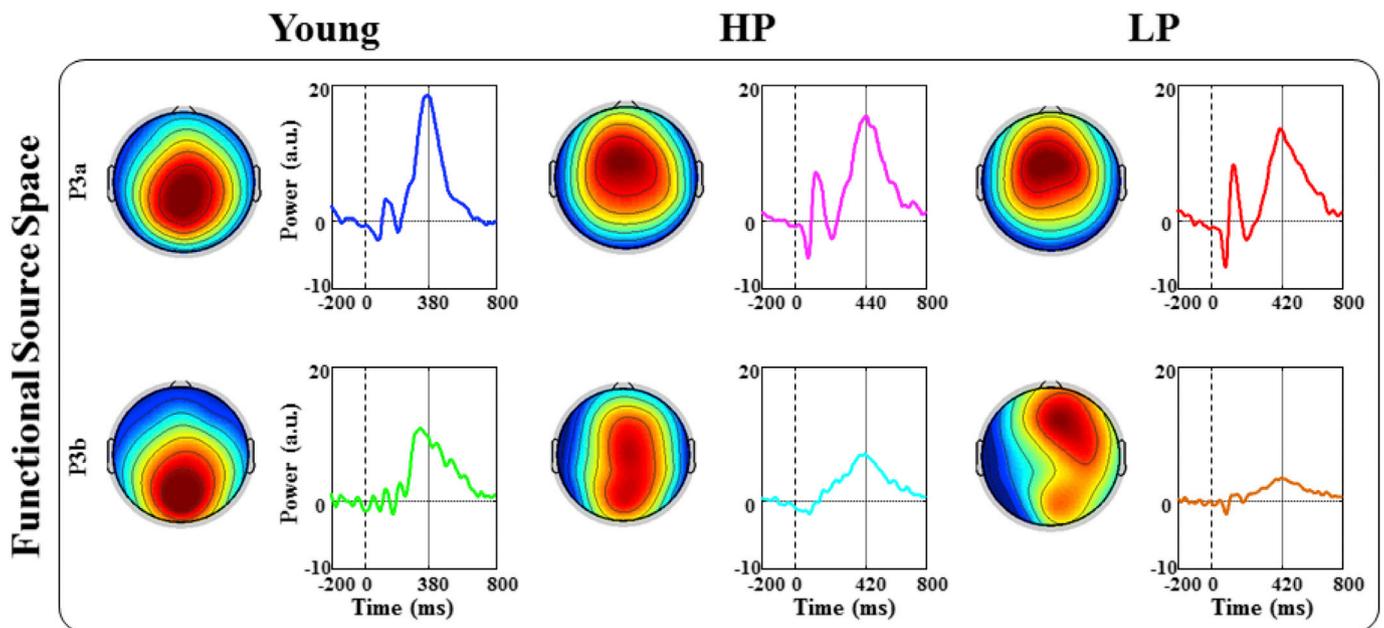


Fig. 3. Topographic maps and ERPs of P3a and P3b ERP responses in young and old groups extracted using FSS.

ERPs and topographic maps for the three groups (Young vs. HP vs. LP) on FS<sub>P300</sub>. Blue, magenta and red lines indicate FS<sub>P3a</sub> and green, cyan and brown lines indicate FS<sub>P3b</sub> for Young, HP and LP groups respectively.

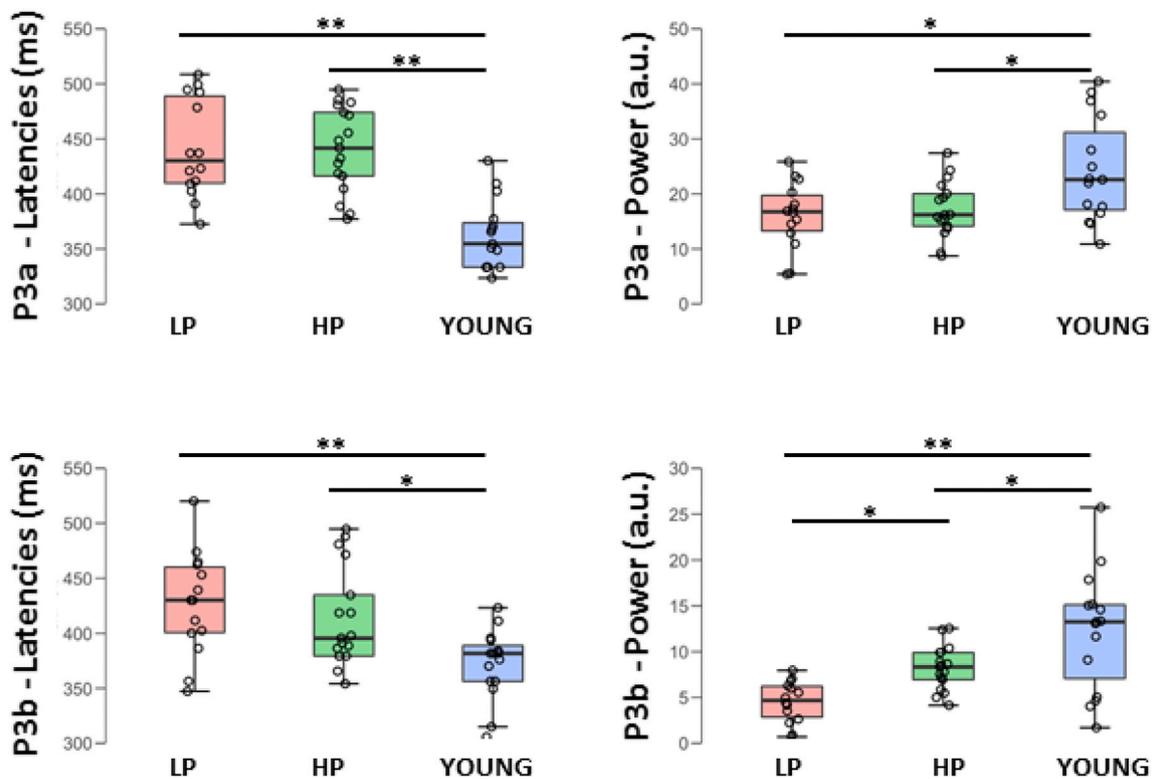


Fig. 4. Group differences in P3a and P3b peak amplitude and latency extracted using FSS.

Box plots illustrating the first quartile, median, and third quartile and 95% confidence limits. Circles show the individual data points. Black horizontal lines highlight significant differences between the groups (\* $p < 0.05$  FDR corrected, \*\* $p < 0.005$  FDR corrected).

information). In particular, this feature enabled us to classify HP from LP participants with the highest accuracy (78.1% accuracy, 76.47% sensitivity, 80% specificity,  $p\text{FDR} < 0.001$ ). It was also possible to distinguish between all comparisons using P3b ERP data from electrode Pz, however this was not as accurate or specific as using FSS derived ERPs. We note that when comparing HP and LP individuals, sensitivity (true positives) was the same for both FSS ERPs and Pz ERPs (76.47%), however FSS ERPs showed higher specificity (true negatives) compared to classification using Pz ERP data (80% compared to 66.67%).

### 3.2.5. Is FSS more sensitive than traditional methods?

In order to establish whether FSS is more sensitive to individual differences in aging compared to traditional channel selection we compared P3b amplitude in all three groups at Fz, Cz, and Pz. There was no significant effect of group at Fz ( $F(2,43) = 2.387$ ,  $p = 0.104$ , partial eta squared = 0.1), however there were significant group differences in P3b amplitude at Cz ( $F(2,43) = 4.706$ ,  $p = 0.014$ , partial eta squared = 0.18) and Pz ( $F(2,43) = 9.327$ ,  $p < 0.001$ , partial eta squared = 0.303). Crucially, significant group differences at electrodes Cz and Pz were driven by larger P3b peak amplitude in the Young group compared to both the HP (Pz:  $t(30) = 3.387$ ,  $p\text{FDR} = 0.003$ ) and LP groups (Cz:  $t(27) = 3.017$ ,  $p\text{FDR} = 0.013$ ; Pz:  $t(27) = 4.026$ ,  $p\text{FDR} < 0.001$ ). There were no significant differences between HP and LP groups in P3b amplitudes from channel space (Fz:  $t(29) = 1.587$ ,  $p\text{FDR} = 0.24$ ; Cz:  $t(29) = 1.134$ ,  $p\text{FDR} = 0.263$ ; Pz:  $t(29) = 0.821$ ,  $p\text{FDR} = 0.416$ ). There were also no significant differences between HP and LP individuals in P3b peak latency and P3a peak latency and amplitude.

To further compare FSS and channel-based methods we used Bayesian independent samples t-tests. There are a number of advantages for using Bayesian methods over traditional frequentist approaches. First, Bayesian approaches allow one to make inferences about null hypothesis as well as the alternative hypothesis. Second, it is possible to compare

Bayes factors across analyses and, based on the magnitude of the Bayes Factor, make statements about whether one result is stronger than another. Using the FSS approach, there was a significant group difference in P3b amplitude (HP > LP:  $t(29) = 4.379$ ,  $p < 0.001$ , Cohen's  $d = 1.58$ ,  $BF_{10} = 156.252$ ). P3b peak amplitude extracted from electrodes Fz and Pz was not significantly different between HP and LP (Fz:  $t(29) = 1.537$ ,  $p = 0.135$ , Cohen's  $d = 0.555$ ,  $BF_{10} = 0.826$ ; Pz:  $t(29) = 1.23$ ,  $p = 0.228$ , Cohen's  $d = 0.444$ ,  $BF_{10} = 0.604$ ). P3b peak amplitude was significantly different between HP and LP groups ( $t(30) = 2.149$ ,  $p = 0.04$ , Cohen's  $d = 0.776$ ,  $BF_{10} = 1.848$ ), however, the magnitudes of the Bayes factors for group differences using FSS and Cz were very different (156.252 for FSS compared to 1.848 for Cz). This suggests that the alternative hypothesis (HP <> LP) was 1.8 times more likely than the null hypothesis using data from electrode Cz, whereas same the alternative hypothesis was 156 times more likely than the null hypothesis using FSS. This means the alternative hypothesis was 87 times more likely to be supported using FSS compared to the best electrode (Cz).

Finally, we used a regression analysis to investigate which parameters best explained individual variability in our metric of cognitive decline (Logical Memory-NART). Independent variables included: Age, MMSE, NART (Z transformed), FSS P3b amplitude, Fz P3b amplitude, Cz P3b amplitude, and Pz P3b amplitude. A significant regression equation was found ( $F(7,23) = 3.215$ ,  $p = 0.016$ ) with an  $R^2$  of 0.495. Only the FSS P3b amplitude significantly explained cognitive decline scores across individuals ( $\beta = 0.162$ ,  $SE = 0.048$ ,  $t = 3.367$ ,  $p = 0.003$ ). A Bayesian regression using the same variables found that the best model to explain Logical Memory-NART scores was to use only the FSS P3b amplitude ( $BF_{10} = 205.485$  compared to the null model). A Bayesian model average of the candidate models for each of the independent variables showed that only the inclusion of FSS P3b amplitude had a strong influence on the model ( $BF_{\text{inclusion}} = 59.741$ , all other variables had a  $BF < 1$ ). A separate Bayesian regression was run using independent variables of:

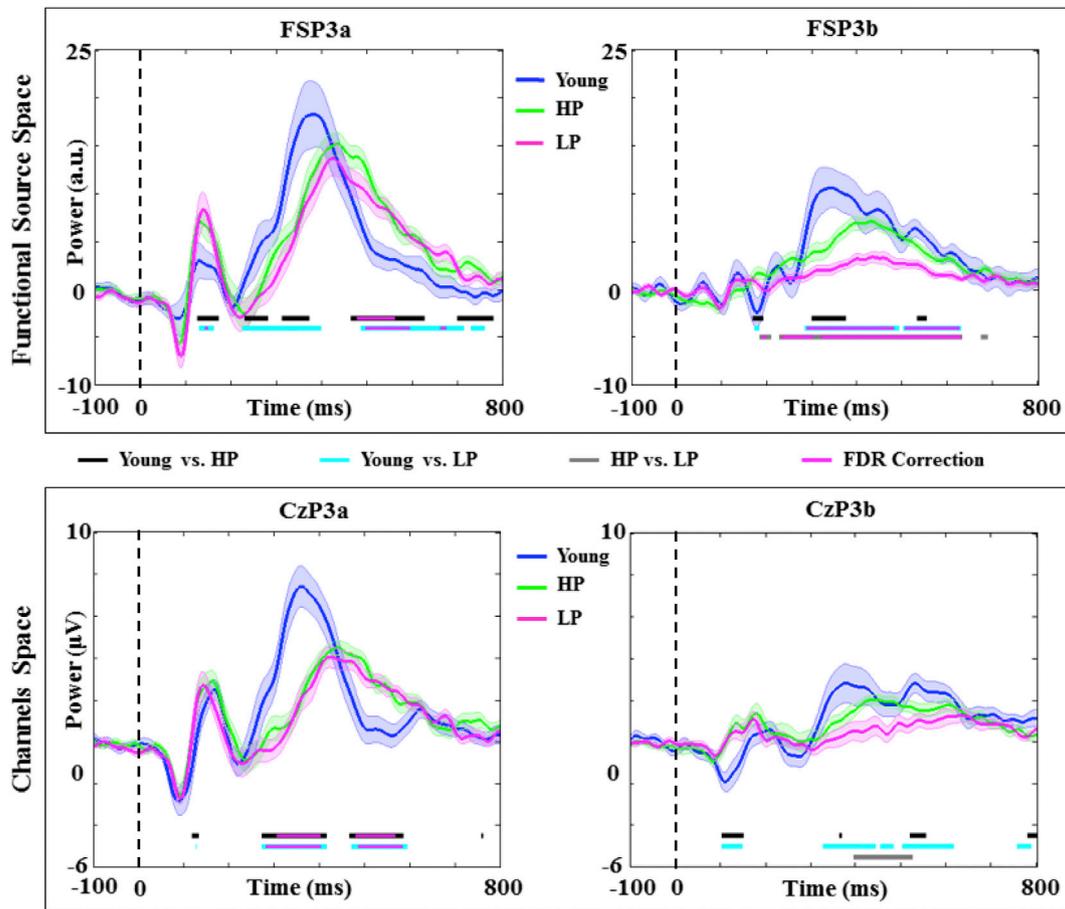


Fig. 5. P3a and P3b functional behaviour between young and old participants.

Point-by-point analysis of P3a and P3b ERPs extracted using FSS (top panels) and the best performing single EEG channel (CZ; bottom panels). In all plots Young (blue Line), HP (Green Line), and LP (Magenta Line) are shown with the shaded area of the same colour highlighting standard error. Horizontal black, cyan and grey thick lines indicate a significant group difference between Y vs. HP, Y vs. LP and HP vs. LP respectively (permutation  $t$ -test at  $p = 0.05$ ; horizontal pink line indicates  $pFDR < 0.05$ ).

Age, MMSE, NART (Z transformed), FSS P3a amplitude, FSS P3a latency, FSS P3b amplitude, FSS P3b latency. This analysis was run to establish which of the neural signals extracted by FSS best captured the variance in Logical Memory-NART scores. Once again, the winning model was a model that only included FSS P3b amplitude ( $BF_{10} = 205.485$  compared to the null model). Similarly, Bayesian model averaging of the candidate models found that only inclusion of the FSS P3b amplitude variable influenced the model ( $BF_{inclusion} = 74.176$ , all other variables had  $BF < 1$ ).

#### 4. Discussion

Despite decades of research investigating the P300 ERP and aging, it has been unclear whether changes in the P3a and P3b ERPs only reflect the process of aging or whether these could be used as markers of cognitive impairment. We posited that signal extraction techniques commonly used in EEG analyses (i.e. using single electrodes or a small subset of electrodes) could have reduced signal-to-noise by excluding potential neural generators. The FSS approach used in this study utilised a weighted combination of all available electrodes in an effort to improve signal-to-noise ratio and avoid missing potential neural generators that contribute to the P300 ERP. First, we demonstrated using FSS that whilst both the P3a and P3b change with age, only changes in P3b amplitude discriminated between individual differences in cognitive decline (episodic memory baselined by IQ). Crucially, the relationship between P3b amplitude and cognitive decline could not be explained by age.

Second, we confirmed that EEG signals extracted using FSS were 87 times more sensitive (based on Bayesian analyses) to these individual differences compared to EEG signals derived from the best single electrode (Cz). These analyses not only highlight the importance of using more advanced EEG signal processing methods such as FSS, but further highlight the utility of the P3b ERP as an easy-to-acquire marker of early, age-independent, loss of memory performance.

##### 4.1. The advantages of Functional Source Separation (FSS)

Although non-invasive electrophysiological techniques, such as EEG, provide the opportunity to directly measure the activity of large-scale neuronal populations, different challenges remain in characterizing this activity. In particular, electrical potentials generated by neuronal activity are not only detected close to neuronal sources but also at distant sites due to electric field propagation. Therefore, each channel positioned across the whole head derives its signal from more than one source (Porcaro and Tecchio, 2014; Rusiniak et al., 2013; Siegel et al., 2012). Since the P300 undeniably arises from a widely distributed network (Fjell et al., 2007; Linden, 2005; Polich, 2007, 2003), selecting a channel or averaging a group of channels based on the topographic representation might be misleading if one aims to describe a distributed brain network. In this respect, we believe that methods capable of extracting the neural source under investigation (such as FSS) are suitable to avoid selection of channels and to overcome possibly misleading results. The FSS approach has been successfully applied to extract EEG specific features using

Table 3

Machine learning accuracy for categorizing participants based on ERP data (peak amplitude and latency). P-values in bold survived FDR correction for multiple comparisons (pFDR<0.05).

	Young vs HP		Young vs LP		HP vs LP		Average Accuracy	
	Accuracy	P value	Accuracy	P value	Accuracy	P Value	Young vs Old	All comparisons
FSS P3a	81.25	<b>&lt;0.001</b>	75.00	<b>0.006</b>	28.13	0.914	78.13	61.46
Fz P3a	56.25	0.276	71.88	<b>0.007</b>	53.13	0.402	64.06	60.42
Cz P3a	81.25	<b>&lt;0.001</b>	81.25	<b>&lt;0.001</b>	56.25	0.258	81.25	72.92
Pz P3a	84.38	<b>&lt;0.001</b>	84.38	<b>&lt;0.001</b>	50.00	0.518	84.38	72.92
FSS P3b	75.00	<b>0.002</b>	90.63	<b>&lt;0.001</b>	78.13	<b>&lt;0.001</b>	82.81	81.25
Fz P3b	78.13	<b>0.002</b>	75.00	<b>0.003</b>	56.25	0.26	76.56	69.79
Cz P3b	65.63	0.039	59.38	0.161	56.25	0.256	62.50	60.42
Pz P3b	71.88	<b>0.004</b>	75.00	<b>0.002</b>	71.88	<b>0.004</b>	73.44	72.92

	Young vs HP		Young vs LP		HP vs LP		Average Young vs Old		Average All comparisons	
	Sensitivity (TP)	Specificity (TN)	Sensitivity (TP)	Specificity (TN)	Sensitivity (TP)	Specificity (TN)	Sensitivity (TP)	Specificity (TN)	Sensitivity (TP)	Specificity (TN)
FSS P3a	<b>80.00</b>	<b>82.35</b>	<b>73.33</b>	<b>76.47</b>	35.29	20.00	76.67	79.41	62.88	59.61
Fz P3a	60.00	52.94	<b>73.33</b>	<b>70.59</b>	47.06	60.00	66.67	61.76	60.13	61.18
Cz P3a	<b>73.33</b>	<b>88.24</b>	<b>80.00</b>	<b>82.35</b>	58.82	53.33	76.67	85.29	70.72	74.64
Pz P3a	73.33	94.12	<b>80.00</b>	<b>88.24</b>	41.18	60.00	76.67	91.18	64.84	80.78
FSS P3b	<b>73.33</b>	<b>76.47</b>	<b>86.67</b>	<b>94.12</b>	<b>76.47</b>	<b>80.00</b>	80.00	85.29	78.82	83.53
Fz P3b	<b>80.00</b>	<b>76.47</b>	<b>80.00</b>	<b>70.59</b>	35.29	80.00	80.00	73.53	65.10	75.69
Cz P3b	66.67	64.71	60.00	58.82	64.71	46.67	63.33	61.76	63.79	56.73
Pz P3b	<b>66.67</b>	<b>76.47</b>	<b>60.00</b>	<b>88.24</b>	<b>76.47</b>	<b>66.67</b>	63.33	82.35	67.71	77.12

different number of channels for primary motor 28 MEG Sensors, (Porcaro et al., 2008); EEG Electrodes, (Di Pino et al., 2012); primary sensory 28 MEG Sensors, (Porcaro et al., 2008); 23 EEG Electrodes (Porcaro et al., 2009a, 2009b); 39 EEG Electrodes (Porcaro et al., 2017) and primary visual 64 EEG Electrodes (Ostwald et al., 2011; Porcaro et al., 2011, 2010) areas but this is the first demonstration that FSS can be used to extract non-primary and wide spread brain activations. Based on the spatial (source localization and topography) and functional behaviour of the sources extracted by FSS, it seems that the impact of the number of the electrodes is rather small. Moreover, functional behaviour and signal-to-noise ratio (SNR) seems to be improved when FSS is applied, independently of the number of channels used (Porcaro et al., 2009b, 2017, 2018, 2009a, 2011; 2010) making this a useful tool for clinical datasets collected using a low number of electrodes. Since the P300 has been intensively investigated, it provides a perfect framework to test the quality of functional sources extracted by FSS (i.e. comparing the results obtained by FSS with those obtained in previous studies and well established in the literature). Here, we used specific temporal fingerprint information, i.e. we maximized the P300 response in the chosen window 320–480 ms for both young old participants to extract the functional sources (FSs) of interest. The features extracted by FSS were well matched with those reported in the literature using pre-selected channels such as slower peak latency of the P300 response and a more frontal topography map in the old group than the young one (Balsters et al., 2013; O'Connell et al., 2012a). Moreover, our discrepancy measure (i.e. the P300 ERP of Raw Data minus the P300 ERP obtained reconstructing the data using the source FS<sub>P300</sub>) showed that the residual activity of the P300 peak is near to zero for both P3a and P3b as expected if all the electrical activity is in the FS<sub>P300</sub>. This check assured us that we were able to extract all the electrical activity underlying the P300 response. Finally, results obtained by FSS seemed to be more sensitive for discriminating between HP and LP elderly than those obtained using the Cz channel

which has been conventionally chosen to study the P3a and P3b components of the P300 (Johnson, 1993; Linden, 2005; Magnano et al., 2006; Polich, 2007; Rossini et al., 2007).

Independent Components Analysis (ICA) is the most common approach for extracting neural sources underpinning the P300 ERP response (Makeig et al., 2004). Recently, van Dinteren et al. (van Dinteren et al., 2017) used group ICA to investigate the P300 aging effect, however, a major drawback of using ICA compared to FSS is that multiple independent components might describe the P300 ERP. Porcaro et al. (2018) compared FSS and ICA approaches on data from primary motor and primary sensory areas. The results showed that ICA tended to describe the functional behaviour in multiple components whereas a single functional source was extracted by FSS allowing for simpler analyses. In the case of van Dinteren et al. (van Dinteren et al., 2017), they report four components that describe the P300 ERP, some of which show differences with age. In addition to being difficult to interpret, the ICA-based approach to extracting neural sources raises a number of methodological issues. First, the selection of 'relevant' components is usually biased by the choice of the user (depending on user experience). This selection process can also be very demanding and time-consuming on the user if they are working with high density EEG data (i.e. 128 channels or more). Finally, the combination of multiple components requires the user to reconstruct the data by removing all non P300 components from the data and returning to channel space. This then requires channel selection for analysis, re-introducing previously mentioned concerns about identifying sources. Compared to this, FSS exploits a well-known a-priori 'functional' property (i.e. relevant 'fingerprint' information regarding frequency range or time course characteristics) to identify the source of interest. This means FSS gives the user one source that maximises the pre-defined functional property of the data, thus addressing all the concerns raised above. The potential drawback is that FSS cannot be used to extract unexpected or unknown brain activities,

however previous studies have found that this semi-blind approach often outperforms fully blind source separation approaches like ICA (Barbati et al., 2008; Porcaro et al., 2011, 2010).

#### 4.2. P300, aging, and cognitive decline

Consistent with previous studies we showed that both peak latency and peak amplitude of the P3a and P3b ERPs changed with age (Fjell et al., 2007; O'Connell et al., 2012a; Polich, 1997; Walhovd and Fjell, 2001). Specifically, the peak latency of both the P3a and P3b responses increased in the elderly group, whereas the peak amplitude of both the P3a and P3b decreased with age. However, only the P3b amplitude distinguished between the two elderly subgroups (HP vs LP). This effect was also apparent when ERPs were compared at each sampling point and when using machine learning to detect group classification. Fig. 5 highlights that there was no form of delayed P3b response in the LP group compared to the HP, rather there was an absence of the P3b ERP in the LP group. This is consistent with Dockree et al. (2015) who also subdivided older individuals into HP and LP groups using the same criteria as this study. Dockree et al. (2015) found that both young and HP showed greater EEG activity in centroparietal regions when recognising previously learnt words, whereas ERP responses in LP individuals were not different for learnt and novel words. Similar to this study (see Fig. 3), ERP topographies presented in Dockree et al. (2015) show additional frontal peaks in HP and LP individuals compared to purely centroparietal activity in the Young group. Whilst some have suggested that the increased frontal activity in aging reflects compensatory activity (Davis and Friedrich, 2010; van Dinteren et al., 2014), others have shown that increased frontalisation of the P300 response is associated with poorer performance in neuropsychological tests (Fabiani et al., 1998; West et al., 2010). It has previously been suggested that the prefrontal cortex contributes to the initial target detection trials, however, after a few trials a model of the task is created and prefrontal control is no longer required (Fabiani and Friedman, 1995; Richardson et al., 2011; West et al., 2010). In line with this, it has been suggested that increased frontalization of P3b responses in elderly individuals may reflect an inability to establish a strong mental representation of the task stimuli, and as such there is a continued reliance on both frontal and parietal regions in order to maintain task performance. One limitation of these previous studies has been the inability to localize this general frontalization effect to discrete neural sources. In order to address this problem, O'Connell et al. (Redmond G. O'Connell et al., 2012) used simultaneous EEG-fMRI recordings to investigate the neural generators underpinning this phenomenon. Crucially, O'Connell et al. (O'Connell et al., 2012a) demonstrated that this marked frontalization effect with age was driven by distinct neuro-anatomical changes for the P3a and P3b responses. Whilst the increase in frontal P3a responses with age was produced by increased activity in the left inferior frontal gyrus and middle cingulate cortex, the increase in frontal P3b responses was driven by increased activity in the right middle frontal gyrus and putamen. This once again highlights that the P3a and P3b responses are underpinned by unique patterns of brain activity and are likely contributing to different cognitive processes. In addition, these findings suggest that the general process of frontalization that occurs during aging can be driven by distinct neural generators. Given that the LP group showed reduced amplitude for P3b only, it is possible that the cognitive decline may reflect aberrant cortico-striatal connectivity. This is in keeping with previous work investigating cortico-striatal connectivity in healthy elderly individuals by Ystad et al. (Ystad et al., 2011, 2010). Whilst Ystad et al. (2011) found that elderly individuals who performed better on executive function tasks had stronger connections between the putamen and dorso-medial parietal cortex (putatively dorsal attention network), Ystad et al. (2010) showed a negative relationship between the episodic memory and the number of cortico-striatal connections. This suggests that whilst increased connectivity between the putamen and parietal lobe is beneficial in aging, diffuse and more distributed connectivity between the putamen and multiple other regions

is detrimental. This could suggest that increased frontal-striatal connectivity is actually decreasing task performance, and perhaps taking resources away from beneficial parieto-striatal connectivity. However, EEG is not suited to address cortico-subcortical hypotheses and further research in this area would require fMRI or simultaneous EEG-fMRI.

With the greying population there is a growing need for sensitive markers of cognitive decline in ageing. Here, we have defined cognitive decline in ageing using a combination of neuropsychological tests (age-adjusted scores from the Logical Memory subtest of the Wechsler Memory Scale III baselined against NART IQ). The results of this study build on a growing literature highlighting neural differences between HP and LP elderly individuals as defined here (Dockree et al., 2015; Hogan et al., 2011, 2012; O'Hara et al., 2013). Furthermore, LP individuals as defined above, have shown differential receptor expression and cytokine profiles indicative of inflammation that may be involved in the prodromal processes leading to the development of neurodegenerative disease (Downer et al., 2013). However, to our knowledge these findings have only been demonstrated at the group level, and it is not clear whether the any of these markers (including the P300 ERP) index individual differences in cognitive decline. In an attempt to address this issue, we used a machine learning approach to establish whether the P3b signals extracted using FSS could be a useful tool to categorize elderly individuals as HP or LP. This would be particularly advantageous clinically given the speed with which P300 ERP data could be collected (~10mins), which is much faster than the time needed to collect the neuropsychological battery (~2hrs). The results of our SVM analysis were promising, demonstrating that P3b signals extracted using FSS could categorize individuals as HP or LP with 78% accuracy (76% sensitivity; 80% specificity). However, it is important to treat cross-sectional results with caution (Fisher et al., 2018). Nyberg et al. (2010) found that whilst cross-sectional evidence showed an increase in prefrontal activity with ageing, a longitudinal examination showed the opposite trend. The results presented here provide an important contribution to the ageing literature, however, it is crucial to replicate these results with an independent sample, and to conduct longitudinal designs to establish whether individual changes in the P300 ERP index individual changes in cognitive decline.

It will also be important to move towards more mechanistic explanations of perturbed neural activity (Hanks and Summerfield, 2017; Montague et al., 2012). Converging evidence from mathematical modelling and primate electrophysiology have suggested that the P300 ERP (specifically the P3b) has all the characteristics of a 'decision variable' signal (Kelly and O'Connell, 2015; O'Connell et al., 2012b; Twomey et al., 2015). A decision variable signal represents the accumulation of sensory information over time until there is sufficient information (also referred to as evidence) to pass through a boundary criterion. The absence of the P3b response in the LP group could suggest that elderly individuals at risk of cognitive decline have a reduced ability to accumulate sensory evidence over time. This is in keeping with the findings of Dockree et al. (2015), who found LP individuals showed significantly slower evidence accumulation during an episodic memory task compared to HP individuals. Unfortunately, traditional EEG paradigms such as the three-stimulus oddball and the memory encoding paradigm used by Dockree et al. (2015) are not optimal for investigating evidence accumulation mechanisms. Further research is necessary to determine whether the absence of this fundamental neural mechanism (evidence accumulation) could become a crucial marker of cognitive decline and dementia.

## 5. Conclusion

In conclusion, using a simple well-established paradigm in combination with advanced signal processing techniques (FSS), we were able to reliably extract both subcomponents of the P300 ERP. Our analyses demonstrated: 1) P3b amplitude could be a useful and easy-to-acquire metric of age independent memory loss – and possibly prodromal disease – in elderly individuals, and 2) more advanced signal processing

methods such as FSS are necessary in order to detect subtle variations in EEG signals that are likely obscured by much larger effects such as age-related differences.

## Acknowledgement

The work was partly supported by the Research Foundation Flanders (grants G0F76.16N, EOS.30446199 and G0936.16N), KU Leuven Special Research Fund (grant C16/15/070), Senior Fellowship KU Leuven (ZKD1331 - SF/16/011), SNSF (IZK0Z3\_163614 and 320030\_175616) and the ETH Visiting Professors Program.

## References

- Balsters, J.H., O'Connell, R.G., Galli, A., Nolan, H., Greco, E., Kilcullen, S.M., Bokde, A.L.W., Lai, R., Upton, N., Robertson, I.H., 2013. Changes in resting connectivity with age: a simultaneous electroencephalogram and functional magnetic resonance imaging investigation. *Neurobiol. Aging* 34, 2194–2207. <https://doi.org/10.1016/j.neurobiolaging.2013.03.004>.
- Barbati, G., Porcaro, C., Zappasodi, F., Rossini, P.M., Tecchio, F., 2004. Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals. *Clin. Neurophysiol.* 115 (5), 1220–1232. <https://doi.org/10.1016/j.clinph.2003.12.015>.
- Barbati, G., Sigismondi, R., Zappasodi, F., Porcaro, C., Graziadio, S., Valente, G., Balsi, M., Rossini, P.M., Tecchio, F., 2006. Functional source separation from magnetoencephalographic signals. *Hum. Brain Mapp.* 27, 925–934. <https://doi.org/10.1002/hbm.20232>.
- Barbati, G., Porcaro, C., Hadjipapas, A., Adjajian, P., Pizzella, V., Romani, G.L., Seri, S., Tecchio, F., Barnes, G.R., 2008. Functional source separation applied to induced visual gamma activity. *Hum. Brain Mapp.* 29, 131–141. <https://doi.org/10.1002/hbm.20375>.
- Davis, N.C., Friedrich, D., 2010. Age stereotypes in middle-aged through old-old adults. *Int. J. Aging Hum. Dev.* <https://doi.org/10.2190/AG.70.3.b>.
- Davis, S.W., Dennis, N.A., Daselaar, S.M., Fleck, M.S., Cabeza, R., 2008. Qué PASA? the posterior-anterior shift in aging. *Cerebr. Cortex* 18, 1201–1209. <https://doi.org/10.1093/cercor/bhm155>.
- Di Pino, G., Porcaro, C., Tombini, M., Assenza, G., Pellegrino, G., Tecchio, F., Rossini, P.M., 2012. A neurally-interfaced hand prosthesis tuned inter-hemispheric communication. *Restor. Neurol. Neurosci.* 30 (5), 407–418. <https://doi.org/10.3233/RNN-2012-120224>.
- Dockree, P.M., Brennan, S., O'Sullivan, M., Robertson, I.H., O'Connell, R.G., 2015. Characterising neural signatures of successful aging: electrophysiological correlates of preserved episodic memory in older age. *Brain Cognit.* 97, 40–50. <https://doi.org/10.1016/j.bandc.2015.04.002>.
- Donchin, E., Coles, M.G.H., 1988. Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* 11, 357. <https://doi.org/10.1017/S0140525X00058027>.
- Downer, E.J., Jones, R.S., McDonald, C.L., Greco, E., Brennan, S., Connor, T.J., Robertson, I.H., Lynch, M.A., 2013. Identifying early inflammatory changes in monocyte-derived macrophages from a population with IQ-discrepant episodic memory. *PLoS One* 8. <https://doi.org/10.1371/journal.pone.0063194>.
- Fabiani, M., Friedman, D., 1995. Changes in brain activity patterns in aging: the novelty oddball. *Psychophysiology* 32, 579–594. <https://doi.org/10.1111/j.1469-8986.1995.tb01234.x>.
- Fabiani, M., Friedman, D., Cheng, J.C., 1998. Individual differences in P3 scalp distribution in older adults, and their relationship to frontal lobe function. *Psychophysiology* 35, 698–708. <https://doi.org/10.1111/1469-8986.3560698>.
- Fisher, A.J., Medaglia, J.D., Jeronimus, B.F., 2018. Lack of group-to-individual generalizability is a threat to human subjects research. *Proc. Natl. Acad. Sci. Unit. States Am.* 201711978 <https://doi.org/10.1073/pnas.1711978115>.
- Fjell, A.M., Walhovd, K.B., Fischl, B., Reinvang, I., 2007. Cognitive function, P3a/P3b brain potentials, and cortical thickness in aging. *Hum. Brain Mapp.* 28, 1098–1116. <https://doi.org/10.1002/hbm.20335>.
- Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6).
- Goh, J.O., Park, D.C., 2009. Neuroplasticity and cognitive aging: the scaffolding theory of aging and cognition. *Restor. Neurol. Neurosci.* <https://doi.org/10.3233/RNN-2009-0493>.
- Hanks, T.D., Summerfield, C., 2017. Perceptual decision making in rodents, monkeys, and humans. *Neuron*. <https://doi.org/10.1016/j.neuron.2016.12.003>.
- Hillyard, S.A., Squires, K.C., Bauer, J.W., Lindsay, P.H., 1971. Evoked potential correlates of auditory signal detection. *Science* (80-. ) 172, 1357–1360. <https://doi.org/10.1126/science.172.3990.1357>.
- Hogan, M., Collins, P., Keane, M., Kilmartin, L., Kaiser, J., Kenney, J., Lai, R., Upton, N., 2011. Electroencephalographic coherence, aging, and memory: distinct responses to background context and stimulus repetition in younger, older, and older declined groups. *Exp. Brain Res.* <https://doi.org/10.1007/s00221-011-2726-8>.
- Hogan, M.J., Kilmartin, L., Keane, M., Collins, P., Staff, R.T., Kaiser, J., Lai, R., Upton, N., 2012. Electrophysiological entropy in younger adults, older controls and older cognitively declined adults. *Brain Res.* <https://doi.org/10.1016/j.brainres.2012.01.027>.
- Johnson, R., 1993. On the neural generators of the P300 component of the event-related potential. *Psychophysiology* 30, 90–97. <https://doi.org/10.1111/j.1469-8986.1993.tb03208.x>.
- Kappenman, E.S., Luck, S.J., 2010. The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology* 47 (5), 888–904. <http://doi.org/10.1111/j.1469-8986.2010.01009.x>.
- Kelly, S.P., O'Connell, R.G., 2015. The neural processes underlying perceptual decision making in humans: recent progress and future directions. *J. Physiol. Paris*. <https://doi.org/10.1016/j.jphysparis.2014.08.003>.
- Khan, O.I., Farooq, F., Akram, F., Choi, M.T., Han, S.M., Kim, T.S., 2012. Robust extraction of P300 using constrained ICA for BCI applications. *Med. Biol. Eng. Comput.* 50, 231–241. <https://doi.org/10.1007/s11517-012-0861-4>.
- Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P., 1983. Optimization by simulated annealing. *Science* (80-. ) 220, 671–680. <https://doi.org/10.1126/science.220.4598.671>.
- Linden, D.E.J., 2005. The P300: where in the brain is it produced and what does it tell us? *Neuroscience* 11, 563–576. <https://doi.org/10.1177/1073858405280524>.
- Murphy, P.R., Robertson, I.H., Balsters, J., O'Connell, R.G., 2011. Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology* 48 (11), 1532–1543. <http://doi.org/10.1111/j.1469-8986.2011.01226.x>.
- Lu, W., Rajapakse, J.C., 2005. Approach and applications of constrained ICA. *IEEE Trans. Neural Network.* 16, 203–212. <https://doi.org/10.1109/TNN.2004.836795>.
- Magnano, I., Aiello, I., Piras, M.R., 2006. Cognitive impairment and neurophysiological correlates in MS. *J. Neurol. Sci.* 245, 117–122. <https://doi.org/10.1016/j.jns.2005.08.027>.
- Makeig, S., Delorme, A., Westerfield, M., Jung, T.P., Townsend, J., Courchesne, E., Sejnowski, T.J., 2004. Electroencephalographic brain dynamics following manually responded visual targets. *PLoS Biol.* 2. <https://doi.org/10.1371/journal.pbio.0020176>.
- Mars, R.B., Debener, S., Gladwin, T.E., Harrison, L.M., Haggard, P., Rothwell, J.C., Bestmann, S., 2008. Trial-by-Trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *J. Neurosci.* 28, 12539–12545. <https://doi.org/10.1523/JNEUROSCI.2925-08.2008>.
- Montague, P.R., Dolan, R.J., Friston, K.J., 2012. Computational psychiatry (vol 16, pg 72, 2012). *Trends Cogn.* <https://doi.org/10.1016/j.tics.2011.11.018>.
- Nelson, H.E., 1982. The National Adult Reading Test (NART): Test Manual. Wind. UK NFER-Nelson 124, pp. 0–25. [https://doi.org/Thesis\\_references-Converted#319](https://doi.org/Thesis_references-Converted#319).
- Nieuwenhuis, S., De Geus, E.J., Aston-Jones, G., 2011. The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology*. <https://doi.org/10.1111/j.1469-8986.2010.01057.x>.
- Nyberg, L., Salami, A., Andersson, M., Eriksson, J., Kalpouzos, G., Kauppi, K., 2010. Longitudinal Evidence for Diminished Frontal Cortex Function in Aging. *Pnas*. <https://doi.org/10.1073/pnas.1012651108/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073.pnas.1012651108>.
- Ojala, M., Garriga, G.C., 2009. Permutation tests for studying classifier performance. In: *Proceedings - IEEE International Conference on Data Mining. ICDM.* <https://doi.org/10.1109/ICDM.2009.108>.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4).
- Ostwald, D., Porcaro, C., Bagshaw, A.P., 2011. Voxel-wise information theoretic EEG-fMRI feature integration. *Neuroimage* 55. <https://doi.org/10.1016/j.neuroimage.2010.12.029>.
- O'Connell, R.G., Balsters, J.H., Kilcullen, S.M., Campbell, W., Bokde, A.W., Lai, R., Upton, N., Robertson, I.H., 2012a. A simultaneous ERP/fMRI investigation of the P300 aging effect. *Neurobiol. Aging* 33, 2448–2461. <https://doi.org/10.1016/j.neurobiolaging.2011.12.021>.
- O'Connell, R.G., Dockree, P.M., Kelly, S.P., 2012b. A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat. Neurosci.* 15, 1729–1735. <https://doi.org/10.1038/nn.3248>.
- O'Hora, D., Schinkel, S., Hogan, M.J., Kilmartin, L., Keane, M., Lai, R., Upton, N., 2013. Age-related task sensitivity of frontal EEG entropy during encoding predicts retrieval. *Brain Topogr.* <https://doi.org/10.1007/s10548-013-0278-x>.
- Polich, J., 1997. On the relationship between EEG and P300: individual differences, aging, and ultradian rhythms. *Int. J. Psychophysiol.* 299–317. [https://doi.org/10.1016/S0167-8760\(97\)00772-1](https://doi.org/10.1016/S0167-8760(97)00772-1).
- Polich, J., 2003. Detection of Change: Event-related Potential and fMRI Findings, Overview of P3a and P3b. <https://doi.org/10.1007/978-1-4615-0294-4>.
- Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* <https://doi.org/10.1016/j.clinph.2007.04.019>.
- Polich, J., 2012. Neuropsychology of P300. In: *The Oxford Handbook of Event-related Potential Components.* <https://doi.org/10.1093/oxfordhb/9780195374148.013.0089>.
- Polich, J., Corey-Bloom, J., 2005. Alzheimer's disease and P300: review and evaluation of task and modality. *Curr. Alzheimer Res.* 2, 515–525. <https://doi.org/10.2174/156720505774932214>.
- Polich, J., Criado, J.R., 2006. Polich J & Criado JR 2006.pdf. *Int. J. Psychophysiol.* 60, 172–185. <https://doi.org/10.1016/j.jpsycho.2005.12.012>.
- Porcaro, Camillo, Tecchio, Franca, 2014. Semi-blind functional source separation algorithm from non-invasive electrophysiology to neuroimaging. In: Naik, G.R., Wang, W., B.S.S (Eds.), *Blind Source Separation: Signals and Communication Technology.* Springer-Verlag Berlin Heidelberg 2014, Springer-Verlag, Berlin, Heidelberg, pp. 521–551. [https://doi.org/10.1007/978-3-642-55016-4\\_19](https://doi.org/10.1007/978-3-642-55016-4_19).
- Porcaro, C., Barbati, G., Zappasodi, F., Rossini, P.M., Tecchio, F., 2008. Hand sensory-motor cortical network assessed by functional source separation. *Hum. Brain Mapp.* 29, 70–81. <https://doi.org/10.1002/hbm.20367>.

- Porcaro, C., Coppola, G., Lorenzo, G.D., Zappasodi, F., Siracusano, A., Pierelli, F., Rossini, P.M., Tecchio, F., Seri, S., 2009a. Hand somatosensory subcortical and cortical sources assessed by functional source separation: an EEG study. *Hum. Brain Mapp.* 30. <https://doi.org/10.1002/hbm.20533>.
- Porcaro, C., Zappasodi, F., Rossini, P.M., Tecchio, F., 2009b. Choice of multivariate autoregressive model order affecting real network functional connectivity estimate. *Clin. Neurophysiol.* 120 (2), 436–448. <https://doi.org/10.1016/j.clinph.2008.11.011>.
- Porcaro, C., Ostwald, D., Bagshaw, A.P., 2010. Functional source separation improves the quality of single trial visual evoked potentials recorded during concurrent EEG-fMRI. *Neuroimage* 50, 112–123. <https://doi.org/10.1016/j.neuroimage.2009.12.002>.
- Porcaro, C., Ostwald, D., Hadjipapas, A., Barnes, G.R., Bagshaw, A.P., 2011. The relationship between the visual evoked potential and the gamma band investigated by blind and semi-blind methods. *Neuroimage* 56, 1059–1071. <https://doi.org/10.1016/j.neuroimage.2011.03.008>.
- Porcaro, C., Medaglia, M.T., Krott, A., 2015. Removing speech artifacts from electroencephalographic recordings during overt picture naming. *Neuroimage* 105, 171–180. <https://doi.org/10.1016/j.neuroimage.2014.10.049>.
- Porcaro, C., Di Lorenzo, G., Seri, S., Pierelli, F., Tecchio, F., Coppola, G., 2017. Impaired brainstem and thalamic high-frequency oscillatory EEG activity in migraine between attacks. *Cephalalgia* 37 (10), 915–926. <https://doi.org/10.1177/0333102416657146>.
- Porcaro, C., Cottone, C., Cancelli, A., Salustri, C., Tecchio, F., 2018 Apr. Functional semi-blind source separation identifies primary motor area without active motor execution. *Int. J. Neural Syst.* 28 (3), 1750047. <https://doi.org/10.1142/S0129065717500472>.
- Reuter-Lorenz, P.A., Cappell, K.A., 2008. Neurocognitive aging and the compensation hypothesis. *Curr. Dir. Psychol. Sci.* <https://doi.org/10.1111/j.1467-8721.2008.00570.x>.
- Richardson, C., Bucks, R.S., Hogan, A.M., 2011. Effects of aging on habituation to novelty: an ERP study. *Int. J. Psychophysiol.* 79, 97–105. <https://doi.org/10.1016/j.ijpsycho.2010.09.007>.
- Rossini, P.M., Rossi, S., Babiloni, C., Polich, J., 2007. Clinical neurophysiology of aging brain: from normal aging to neurodegeneration. *Prog. Neurobiol.* <https://doi.org/10.1016/j.pneurobio.2007.07.010>.
- Rusiniak, M., Lewandowska, M., Wolak, T., Pluta, A., Milner, R., Ganc, M., Włodarczyk, A., Senderski, A., Śliwa, L., Skarzynski, H., 2013. A modified oddball paradigm for investigation of neural correlates of attention: a simultaneous ERP-fMRI study. *Magn. Reson. Mater. Phys. Biol. Med.* 26, 511–526. <https://doi.org/10.1007/s10334-013-0374-7>.
- Siegel, M., Donner, T.H., Engel, A.K., 2012. Spectral fingerprints of large-scale neuronal interactions. *Nat. Rev. Neurosci.* <https://doi.org/10.1038/nrn3137>.
- Squires, K.C., Hillyard, S.A., Lindsay, P.H., 1973. Vertex potentials evoked during auditory signal detection: relation to decision criteria \*. *Percept. Psychophysiol.* 14, 265–272. <https://doi.org/10.3758/BF03212388>.
- Squires, N.K., Squires, K.C., Hillyard, S.A., 1975. Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalogr. Clin. Neurophysiol.* 38, 387–401. [https://doi.org/10.1016/0013-4694\(75\)90263-1](https://doi.org/10.1016/0013-4694(75)90263-1).
- Sutton, S., Braren, M., Zubin, J., John, E.R., 1965. Evoked-potential correlates of stimulus uncertainty. *Science* (80- ) 150, 1187–1188. <https://doi.org/10.1126/science.150.3700.1187>.
- Tecchio, F., Porcaro, C., Barbati, G., Zappasodi, F., 2007. Functional source separation and hand cortical representation for a brain-computer interface feature extraction. *J. Physiol.* 580, 703–721. <https://doi.org/10.1113/jphysiol.2007.129163>.
- The JASP Team, 2017. JASP (Version 0.8.1.1) [WWW Document]. [Computer software]. <https://jasp-stats.org/#>.
- Twomey, D.M., Murphy, P.R., Kelly, S.P., O'Connell, R.G., 2015. The classic P300 encodes a build-to-threshold decision variable. *Eur. J. Neurosci.* 42, 1636–1643. <https://doi.org/10.1111/ejn.12936>.
- van Deursen, J.A., Vuurman, E.F.P.M., Smits, L.L., Verhey, F.R.J., Riedel, W.J., 2009. Response speed, contingent negative variation and P300 in Alzheimer's disease and MCI. *Brain Cognit.* 69, 592–599. <https://doi.org/10.1016/j.bandc.2008.12.007>.
- van Dinteren, R., Arns, M., Jongasma, M.L.A., Kessels, R.P.C., 2014. Combined frontal and parietal P300 amplitudes indicate compensated cognitive processing across the lifespan. *Front. Aging Neurosci.* 6. <https://doi.org/10.3389/fnagi.2014.00294>.
- van Dinteren, R., Huster, R.J., Jongasma, M.L.A., Kessels, R.P.C., Arns, M., 2017. Differences in cortical sources of the event-related P3 potential between young and old participants indicate frontal compensation. *Brain Topogr.* 1–12. <https://doi.org/10.1007/s10548-016-0542-y>.
- Walhovd, K.B., Fjell, a M., 2001. Two- and three-stimuli auditory oddball ERP tasks and neuropsychological measures in aging. *Neuroreport* 12, 3149–3153. <https://doi.org/10.1097/00001756-200110080-00033>.
- Wang, S., James, C.J., 2007. Extracting rhythmic brain activity for brain-computer interfacing through constrained independent component analysis. *Comput. Intell. Neurosci.* 2007. <https://doi.org/10.1155/2007/41468>.
- Wechsler, D., 1997. WAIS-III Administration and Scoring Manual. The Psychological Corporation, San Antonio, TX. <https://doi.org/10.1177/1073191102009001003>.
- West, R., Schwarb, H., Johnson, B.N., 2010a. The influence of age and individual differences in executive function on stimulus processing in the oddball task. *Cortex* 46, 550–563. <https://doi.org/10.1016/j.cortex.2009.08.001>.
- Ystad, M., Eichele, T., Lundervold, A.J., Lundervold, A., 2010. Subcortical functional connectivity and verbal episodic memory in healthy elderly-A resting state fMRI study. *Neuroimage* 52, 379–388. <https://doi.org/10.1016/j.neuroimage.2010.03.062>.
- Ystad, M., Hodneland, E., Adolfsdottir, S., Haász, J., Lundervold, A.J., Eichele, T., Lundervold, A., 2011. Cortico-striatal connectivity and cognition in normal aging: a combined DTI and resting state fMRI study. *Neuroimage* 55, 24–31. <https://doi.org/10.1016/j.neuroimage.2010.11.016>.
- Zigmond, A.S., Snaith, R.P., 1983. The hospital anxiety and depression scale. *Acta Psychiatr. Scand.* 67, 361–370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>.