Clinical Neurophysiology 121 (2010) 672–685

Contents lists available at ScienceDirect

Clinical Neurophysiology

journal homepage: www.elsevier.com/locate/clinph

Effect of the static magnetic field of the MR-scanner on ERPs: Evaluation of visual, cognitive and motor potentials

S. Assecondi [a,*], K. Vanderperren [b], N. Novitskiy [c], J.R. Ramautar [d,c], W. Fias [e], S. Staelens [a], P. Stiers [f], S. Sunaert [g], S. Van Huffel [b], I. Lemahieu [a]

[a] Ghent University, Department of Electronics and Information Systems, MEDISIP-IBBT-IbiTech, De Pintelaan 185, B-9000 Ghent, Belgium
[b] Katholieke Universiteit Leuven, Department of Electrical Engineering, ESAT-SCD, B-3001 Leuven, Belgium
[c] Katholieke Universiteit Leuven, Department of Psychology, Laboratory of Experimental Psychology, B-3000 Leuven, Belgium
[d] Katholieke Universiteit Leuven, Department of Pediatric Neurology, B-3000 Leuven, Belgium
[e] Ghent University, Department of Experimental Psychology, Ghent, Belgium
[f] Maastricht University, Faculty of Psychology and Neuroscience, Maastricht, The Netherlands
[g] Katholieke Universiteit Leuven, Department of Radiology, Leuven, Belgium

ARTICLE INFO

Article history:
Accepted 20 December 2009
Available online 25 January 2010

Keywords:
ERP
EEG–fMRI
3 T static magnetic field
Ballistocardiographic artifact
Artifact removal

ABSTRACT

Objective: This work investigates the influence of the static magnetic field of the MR-scanner on ERPs extracted from simultaneous EEG–fMRI recordings. The quality of the ERPs after BallistoCardioGraphic (BCG) artifact removal, as well as the reproducibility of the waveforms in different environments is investigated.

Methods: We consider a Detection, a Go–Nogo and a Motor task, eliciting peaks that differ in amplitude, latency and scalp topography, repeated in two situations: outside the scanner room (0 T) and inside the MR-scanner but without gradients (3 T). The BCG artifact is removed by means of three techniques: the Average Artifact Subtraction (AAS) method, the Optimal Basis Set (OBS) method and the Canonical Correlation Analysis (CCA) approach.

Results: The performance of the three methods depends on the amount of averaged trials. Moreover, differences are found on both amplitude and latency of ERP components recorded in two environments (0 T vs 3 T).

Conclusions: We showed that, while ERPs can be extracted from simultaneous EEG–fMRI data at 3 T, the static magnetic field might affect the physiological processes under investigation.

Significance: The reproducibility of the ERPs in different recording environments (0 T vs 3 T) is a relevant issue that deserves further investigation to clarify the equivalence of cognitive processes in both behavioral and imaging studies.

© 2010 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.

1. Introduction

In the past 20 years, the increased combined use of electrophysiologically- and hemodynamically-based techniques significantly enhanced our understanding of the brain (see Shibasaki (2008) for a recent review). The combination of these techniques improves the accuracy and precision of the identification of active brain regions. Among others, the integration of ElectroEncephaloGraphy (EEG) and functional Magnetic Resonance Imaging (fMRI) is gaining more and more popularity because of the non-invasiveness that characterizes the two modalities and because of its low extra cost if an MR-scanner is available. The latter accelerates the widespread purchase of the required equipments in several research centers. With the combination of EEG and fMRI one can simultaneously achieve the high spatial resolution of fMRI (mm) and the high temporal resolution of the EEG (ms). These two features offer together an insight into the brain dynamics not achievable with any other non-invasive technique (Herrmann and Debener, 2008).

The simultaneous recording of EEG and fMRI is not straightforward. In particular, the static and gradient magnetic fields of the MR-scanner obscure the EEG with high amplitude artifacts, the BallistoCardioGraphic (BCG) artifact and the gradient artifact, respectively (Hamandi et al., 2004; Lemieux et al., 2002). Almost all the commercially available MR-compatible EEG systems provide basic solutions for the removal of gradient and ballistocardiographic artifacts, usually based on the subtraction of an average artifact template.

However, the BCG artifact shows an intrinsic inter- and intra-subject variability in time (Vanderperren et al., 2007), being a phe-

* Corresponding author. Tel.: +32 9 332 43 26.
E-mail address: Sara.Assecondi@gmail.com (S. Assecondi).

1388-2457/$36.00 © 2010 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.clinph.2009.12.032

nomenon related to the pulsatory blood flow in scalp arteries, leading to electrode movements, and to the ferromagnetic properties of blood (Allen et al., 1998; Debener et al., 2008). It appears synchronously to the heart beat and it compromises the entire duration of the recordings. Even though several algorithms have been proposed to remove the BCG artifact, based either on template matching (Allen et al., 1998; Goldman et al., 2000; Sijbers et al., 2000; Ellingson et al., 2004), or on Blind Source Separation (BSS) (Bénar et al., 2003; Niazy et al., 2005; Srivastava et al., 2005; Briselli et al., 2006; Nakamura et al., 2006; Mantini et al., 2007; Assecondi et al., 2009; Dyrholm et al., 2009; Leclercq et al., 2009), a common agreement upon the optimal solution has not yet been reached. As Grouiller et al. (2007) pointed out, each method affects the data in a different way, and each of them is suitable for different situations, depending on the strength of the static magnetic field, and therefore the amplitude of the BCG artifact (Debener et al., 2008), and the EEG feature of interest. Especially in the case of ERPs the effect of BCG artifact removal techniques cannot be predicted on the basis of results obtained in routine EEG recordings because the methodologies of analysis of ERPs and EEG are intrinsically different.

On one hand, the amplitude of an ERP is much smaller than the EEG ($\sim$10 μV versus $\sim$100 μV) since the signal of interest is embedded in background stochastic EEG. In order to increase the signal-to-noise ratio, the single trials are commonly averaged within conditions. This additional average step may help in further reducing random artifact residuals. On the other hand, there is no general agreement in the scientific community about the similarity of ERPs recorded inside and outside the MR-scanner, given the very different environment (in terms of light, noise and stimulus presentation) in which the data are recorded, and this matter has not been fully investigated yet.

Only a few studies report a quantitative analysis of ERPs extracted from simultaneous EEG–fMRI recordings. Bregadze and Lavric (2006) examined mid latency ERPs and found that ERPs recorded with concurrent fMRI at 1.5 T are not significantly compromised. Mulert et al. (2004) used an auditory Detection task and found differences in latency for the N1 component inside and outside the MR-scanner. A difference in latency of the early components in a visual task was also found by Becker et al. (2005) and Comi et al. (2005), when comparing recordings with and without concurrent imaging. The first thorough study was performed by Sammer et al. (2005): they considered three tasks and found that the typical characteristics of the ERPs are preserved after artifact removal. However, all the aforementioned studies were performed at 1.5 T. Few studies present ERPs recorded simultaneously with fMRI at 3 T (e.g. Kruggel et al. (2000), Sadeh et al. (2008), Strobel et al. (2008)). In these studies however, the effect of the static magnetic field of the MR-scanner on the recorded ERPs was not taken into account and a direct comparison of the ERP waves recorded inside and outside the MR-scanner was not shown.

In this paper we investigate the performance of three different, fully automated, BCG artifact removal techniques for simultaneously recorded EEG–fMRI, namely the Average Artifact Subtraction (AAS) (Allen et al., 1998) method, the Optimal Basis Set (OBS) (Niazy et al., 2005) method and the Canonical Correlation Analysis (CCA) (Assecondi et al., 2009) approach, already proven to be successful in recovering routine EEG. The common rationale behind these methods is that a reference signal, representing the BCG artifact, is subtracted from the original signal. The methods differ in the way this reference signal is calculated. In AAS one uses as reference an average BCG artifact. In OBS, the reference is computed on the basis of a selected number of principal components extracted by means of singular value decomposition. In CCA the reference is given by orthogonal canonical variates maximally correlated between two consecutive EEG epochs. Other proposed methods to remove the BCG artifact are based on Independent Component Analysis (ICA) (Srivastava et al., 2005; Debener et al., 2007, 2008; Mantini et al., 2007). However, these techniques present difficulties in selecting the artifact-related components. This is due to the fact that ICA identifies underlying components without assigning them a specific order. For a more detailed evaluation of methods based on Independent Component Analysis or OBS, we refer the reader to Vanderperren et al. (2009).

In order to assess the performance of the aforementioned artifact removal techniques, we considered three ERP tasks, eliciting different components with different characteristics. We used a visual task that elicits low-amplitude early-latency components, a cognitive task that elicits medium-amplitude middle-latency components and a Motor task characterized by a very low-amplitude long-latency increase at low frequency.

The data were recorded in three laboratories, using three MR-scanner models and distinct MR-compatible EEG systems from the same manufacturer. The EEG systems differ in the electrodes setup. This allows us to assess the effect of different equipments and experimental setups on the BCG artifact and its removal. Each subject underwent the same task in two different situations: inside the MR-scanner but without MR-gradients (referred to as 3 T) and outside the MR-scanner (referred to as 0 T). The main reason for not applying the gradients was to avoid any effect due to the gradient artifact removal in the interpretation of the results. This study allowed us to investigate the effect of BCG artifact removal techniques on the extraction of ERPs from simultaneously recorded EEG–fMRI data. Moreover, the effect of the type of MR-scanner as well as the issue of the reproducibility of the ERP waveform inside and outside the scanner, are addressed.

2. Method

2.1. Experimental protocol

Fourteen healthy subjects (9 males and 5 females, mean age: 24.5 ± 3.9 years) participated in the study. They performed three different tasks. Seven subjects performed both the Go–Nogo and the Detection task, while the other seven subjects performed only the Motor task. An overview of the datasets is given in Table 1. The study was approved by the local ethic committee of the institution where the data were recorded (University of Maastricht, K.U.Leuven Gasthuisberg Hospital, Ghent University Hospital) and the subjects gave their informed consent. Each task, with the same experimental design optimized for both EEG and fMRI, was repeated in two different situations, namely outside the scanner room (0 T), and inside the magnet bore but without MR-gradients (3 T). The tasks are shown in Supplementary Figure S1 and described here below.

2.1.1. Experiment 1: Detection task

The Detection task (Supplementary Figure S1a) requires a unimanual keypress to the presentation of a visual stimulus. The stimulus consisted of segments of a circular black-and-white checkerboard, presented one at the time, in randomized sequences to one of the four quadrants of the visual field. In addition, a full circular black-and-white checkerboard is presented in the central part of the screen. The experiment consisted of 4 blocks, in which a total of 400 trials (80 for each stimulus type) were randomly presented. In each condition the stimulus was presented for 100 ms preceded by a fixation dot for 500 ms. The Inter-Stimulus Interval (ISI) varied between 900 and 2400 ms. This task was selected as it evokes a large P1 component (around 100 ms after stimulus onset), followed by a N1 component (around 180 ms), with their maximal

Table 1
Overview of the datasets: for each dataset the age (years), gender (m = male, f = female), location in which the data were recorded (M = Maastricht; L = Leuven; G = Ghent), number of electrodes used, task performed and mean reaction time at 3 T and 0 T are reported. Significant differences between reaction times are marked by stars (***$p < 0{,}001$, **$p < 0.01$, *$p < 0.05$).

Dataset	Age	Gender	Location	Number of electrodes	Task	Rt (3 T)	Rt (0 T)	p-Value
1	28	m	M	62	Detection	309	252	***
2	25	f	M	62	Detection	289	297	–
3	21	f	M	62	Detection	410	363	***
4	32	m	M	62	Detection	342	362	**
5	20	f	M	62	Detection	301	278	***
6	25	m	L	62	Detection	361	429	***
7	25	m	L	62	Detection	325	464	***
8	28	m	M	62	Go–Nogo	338	304	***
9	25	f	M	62	Go–Nogo	368	354	*
10	21	f	M	62	Go–Nogo	428	427	–
11	32	m	M	62	Go–Nogo	401	374	***
12	20	f	M	62	Go–Nogo	349	349	–
13	25	m	L	62	Go–Nogo	422	488	***
14	25	m	L	62	Go–Nogo	469	498	***
15	20	f	G	30	Motor	351	329	**
16	22	m	G	30	Motor	350	331	*
17	24	f	G	30	Motor	453	382	***
18	23	m	G	30	Motor	408	419	–
19	21	m	G	30	Motor	412	350	***
20	31	m	G	30	Motor	443	423	*
21	21	m	G	30	Motor	375	354	**

amplitude on parietal and occipital channels (Vanderperren et al., 2007, 2009).

2.1.2. Experiment 2: Go–Nogo task

The Go–Nogo task (Supplementary Figure S1b) requires a unimanual keypress to the presentation of Go stimuli (red squares), while the subject must withhold when Nogo stimuli are presented (red circle). A total of 320 (80%) Go and 80 (20%) Nogo trials were recorded per subject. In each condition the stimulus was presented for 150 ms preceded by a fixation dot for 100 ms. The ISI varied between 850 and 5850 ms. The Go–Nogo task was chosen as the Nogo trials elicit the prominent P3 component around 300 ms preceded by the N2 component (between 200 and 300 ms), characteristic of cognitive tasks, with their maximal amplitude in the central and frontal areas (Pfefferbaum et al., 1985; Kok, 1986).

2.1.3. Experiment 3: Motor task

The Motor task (Supplementary Figure S1c) requires a right-hand keypress to the presentation of a right stimulus (arrow pointing right) and a left-hand keypress to the presentation of a left stimulus (arrow pointing left) while the subject must withhold when a *catch* stimulus is presented. The same number of left and right stimuli was presented. The stimulus was presented for 200 ms preceded by a fixation cross for 500 ms. A total of 16 blocks of 15 trials each were presented with a pause of 15.5 s between blocks. The ISI varied between 1700 and 2200 ms. The Motor task was chosen as it elicits a low-amplitude low-frequency contralateral deactivation over the motor area, synchronous to the response (Eimer, 1998).

2.2. Data acquisition parameters

The experiments were conducted in three locations, namely the University of Maastricht (Maastricht, The Netherlands), K.U.Leuven Gasthuisberg Hospital (Leuven, Belgium) and Ghent University Hospital (Ghent, Belgium). The Motor task was performed in Ghent while the Detection and the Go–Nogo task were performed by five subjects in Maastricht and by two subjects in Leuven. The EEG was acquired with three distinct MR-compatible systems from the same manufacturer (BrainProducts, Munich, Germany). The electrode setup differed for each location, as explained below. In all cases, an additional electrode was placed below the left eye to record ElectroOculoGraphic (EOG) activity. Moreover, ElectroCardioGraphic (ECG) activity was recorded by an electrode whose position differed for each location. This allows us to assess the effect of different equipments and experimental setups on the artifact removal. Impedances were kept as low as possible (below 5 kΩ). The data were sampled at 5000 Hz and hardware filtered between 0.016 and 250 Hz.

2.2.1. Location 1 (University of Maastricht, Maastricht, The Netherlands)

A 64 channels EEG cap (equidistant 64 channels arrangement with 62 unipolar EEG channels) was used. The data were referenced to Cz and grounded to POz. The ECG electrode was placed on the right lower back. A Siemens 3 T Allegra MR-scanner (Siemens, Munich, Germany) was used to record the data in the presence of a static magnetic field.

2.2.2. Location 2 (K.U.Leuven Gasthuisberg Hospital, Leuven, Belgium)

A 64 channels EEG cap (62 unipolar EEG channels, arranged according to the standard 10–20 system) referenced to FCz and grounded to Iz, was used. The ECG electrode was placed on the left upper back. A Philips 3 T Intera MR-scanner (Philips, Amsterdam, The Netherlands) was used to record the data in the presence of a static magnetic field. During the recording, the helium pump of the MR-scanner was switched off.

2.2.3. Location 3 (Ghent University Hospital, Ghent, Belgium)

A 32 channels EEG cap (30 unipolar EEG channels, arranged according to the standard 10–20 system), referenced to FCz and grounded to Iz, was used. The ECG electrode was placed on the left upper torso. A Siemens 3 T Trio MR-scanner (Siemens, Munich, Germany) was used to record the data in the presence of a static magnetic field.

2.3. Signal processing

The data processing was implemented in Matlab R2008a (Mathworks, Natick, Massachusetts) and the EEGLAB toolbox was used (Delorme and Makeig, 2003). As a preprocessing step, the signals were band-pass filtered between 0.16 and 30 Hz (window-based

finite impulse response filter, Hamming window) and downsampled to 256 Hz. Since the BCG artifact is synchronous to the ejection phase of the cardiac cycle, it was identified by detecting the *R*-peaks on the ECG channel. The *R*-peaks were detected by using the FMRIB plug-in (Niazy et al., 2005) for EEGLAB. The BCG artifact was then removed by using either the AAS, OBS or CCA method.

2.4. Average artifact subtraction

The Average Artifact Subtraction method (AAS), as implemented in the FMRIB plug-in for EEGLAB, was used (Niazy et al., 2005). This implementation is based on the algorithm developed by Allen et al. (1998). Each channel is processed independently. Once the *R*-peaks are detected, they are shifted in time by a period approximately equal to the mean delay between the *R*-peak and the BCG artifact, i.e. 0.21 s. The EEG is then epoched around the detection (with a window equal to 1.5 times the mean RR interval) and the average of 21 artifacts (10 before and 10 after the current one, also included) is subtracted from the EEG.

2.5. Optimal basis set

The Optimal Basis Set method (OBS), as implemented in the FMRIB plug-in for EEGLAB (Niazy et al., 2005), was used. Similarly to AAS, OBS subtracts an artifact template from the EEG on a channel-wise basis. The data are segmented around the detected *R*-peaks (with a window equal to 1.5 times the mean RR interval), shifted by an a-priori determined BCG delay, and aligned in a matrix to calculate the principal components of the artifact using Principal Component Analysis (PCA). The first 3 Principal Components (PC) form an Optimal Basis Set (OBS), where the weighting factors of the PCA decomposition are assumed to account for the variability of the artifact. The OBS is then fitted to and subtracted from each EEG epoch containing the BCG artifact.

2.6. Canonical correlation analysis

We (Assecondi et al., 2009) proposed a method, referred to as CCA, based on Canonical Correlation Analysis (CCA) to remove the BCG artifact. The method extracts sources common to two consecutive EEG epochs, to take into account the intra-subject physiological variability of the BCG artifact. This variability is one of the main properties of the BCG artifact, being the artifact related to the ejection phase of the cardiac cycle. In this study we optimized this method in order to avoid the removal of physiological EEG rhythms and in order to deal with small-amplitude ERPs. This is done by extracting sources common to the current EEG epoch and an average BCG artifact.

Canonical Correlation Analysis (CCA) (Hotelling, 1936; De Clercq et al., 2006) is a multivariate technique that finds linear relations common to two sets of variables.

Let us consider two sets of zero-mean random variables $\mathbf{X} = [\mathbf{x}_1(t) \dots \mathbf{x}_M(t)], t = 1 \dots N$ and $\mathbf{Y} = [\mathbf{y}_1(t) \dots \mathbf{y}_M(t)], t = 1 \dots N$, where N is the number of samples and M is the number of stochastic variables. CCA finds a linear combination of the original variables $\mathbf{X}$ and $\mathbf{Y}$ for which the linear correlation is maximized. The new variables can be written as follows:

$$\begin{aligned} \mathbf{u}(t) &= \mathbf{Xa} \\ \mathbf{v}(t) &= \mathbf{Yb} \end{aligned} \tag{1}$$

$\mathbf{u}(t)$ and $\mathbf{v}(t)$ are called *canonical variates* and $\mathbf{a}$ and $\mathbf{b}$ are two vectors containing the *M canonical coefficients or weights*. The canonical weights are estimated by maximizing the correlation between the two new variables with respect to $\mathbf{a}$ and $\mathbf{b}$. The correlation, which is a scaled version of the covariance, can be expressed as follows:

$$\rho_{(\mathbf{u},\mathbf{v})}(\mathbf{a},\mathbf{b}) = \frac{\mathbf{a}^{\mathbf{T}}\mathbf{C}_{\mathbf{XY}}\mathbf{b}}{\sqrt{(\mathbf{a}^{\mathbf{T}}\mathbf{C}_{\mathbf{XX}}\mathbf{a})(\mathbf{b}^{\mathbf{T}}\mathbf{C}_{\mathbf{YY}}\mathbf{b})}} \tag{2}$$

where $\rho_{(\mathbf{u},\mathbf{v})}$ is the correlation between $\mathbf{u}(t)$ and $\mathbf{v}(t)$. The covariance matrices $\mathbf{C}_{\mathbf{XX}}, \mathbf{C}_{\mathbf{YY}}$ and $\mathbf{C}_{\mathbf{XY}}$ are estimated from the data.

$\rho_{(\mathbf{u},\mathbf{v})}$ (Eq. (2)) is called *canonical correlation* R_c and represents the square root of the variance shared by the two canonical variates. Once the first couple of canonical variates is found, its effect is removed from the original data and the procedure is repeated by searching for a new couple of canonical variates, again with maximal canonical correlation but now orthogonal to the previous set (i.e. uncorrelated with the first pair of canonical variates). In this way two orthogonal matrices $\mathbf{U} = [\mathbf{u}_1(t) \dots \mathbf{u}_M(t)], t = 1 \dots N$ and $\mathbf{V} = [\mathbf{v}_1(t) \dots \mathbf{v}_M(t)], t = 1 \dots N$ are found. An efficient and robust algorithm for implementing CCA is described in Golub and Van Loan (1996).

In addition to the canonical correlation R_c, other coefficients can be calculated, called *redundancy*, as follows:

$$\begin{aligned} R_{\mathbf{X}} &= \frac{\rho_{(\mathbf{u},\mathbf{X})^{\mathbf{T}}}\rho_{(\mathbf{u},\mathbf{X})}}{M} R_c^2 \\ R_{\mathbf{Y}} &= \frac{\rho^{\mathbf{T}}_{(\mathbf{v},\mathbf{Y})}\rho_{(\mathbf{v},\mathbf{Y})}}{M} R_c^2 \end{aligned} \tag{3}$$

where $R_{\mathbf{X}}$ and $R_{\mathbf{Y}}$ are the redundancies of $\mathbf{X}$ and $\mathbf{Y}$, respectively while $\rho_{(\mathbf{u},\mathbf{X})}$ and $\rho_{(\mathbf{v},\mathbf{Y})}$ are the correlations between the canonical variate and each random variable $\mathbf{x}_i(t)$ and $\mathbf{y}_i(t)$, respectively. The squared canonical correlation R_c represents the amount of variance common to the canonical variates, while the redundancies $R_{\mathbf{X}}$ and $R_{\mathbf{Y}}$ represent the proportion of variance in the original sets explained by the canonical correlation (Stewart and Love, 1968; Cooley and Lohnes, 1971). Thus the redundancy represents the amount of actual overlap between the original sets for which the first canonical relationship is responsible.

2.6.1. BCG artifact removal by means of CCA

Our proposed method takes the advantages of both AAS and CCA. On one hand, the EEG epoch is compared to an average artifact, computed as in AAS, representative for the deterministic component of the artifact. On the other hand, CCA, as well as the selection of the artifact-related components, is applied on a single epoch-basis, thus allowing the method to take into account the intra-subject variability of the BCG artifact.

The procedure can be divided in two steps: the identification of sources and the selection of the artifact sources. The method is described in Fig. 1 and proceeds as follows:

- *Extraction of the canonical variates*. All the channels are processed simultaneously, in a window ($\mathrm{EEG}_{\mathrm{epoch}}$) of length equal to the current RR interval, i.e. the distance between two consecutive *R*-peaks, to take into account at the same time both the spatial distribution and the temporal waveform of the BCG artifact. An average artifact (BCG_{21}) is computed by averaging 21 consecutive EEG epochs (10 before and 10 after the current one, also included). The averaging allows us to extract the deterministic component of the artifact from the background EEG. CCA is then applied to the current $\mathrm{EEG}_{\mathrm{epoch}}$ and BCG_{21}. The algorithm extracts sources (canonical variates), ordered according to their canonical correlation R_c that are maximally temporal correlated and have highly correlated topographies. Since BCG_{21} is unlikely to contain physiological EEG rhythms (because of the averaging), these rhythms are not extracted as components.
- *Identification of artifact sources*. Once the sources are identified, the artifact sources must be selected. For each canonical variate, the *redundancy* coefficients R_{EEG} and R_{BCG} are calculated, and the variates ordered according to increasing redundancy. Two

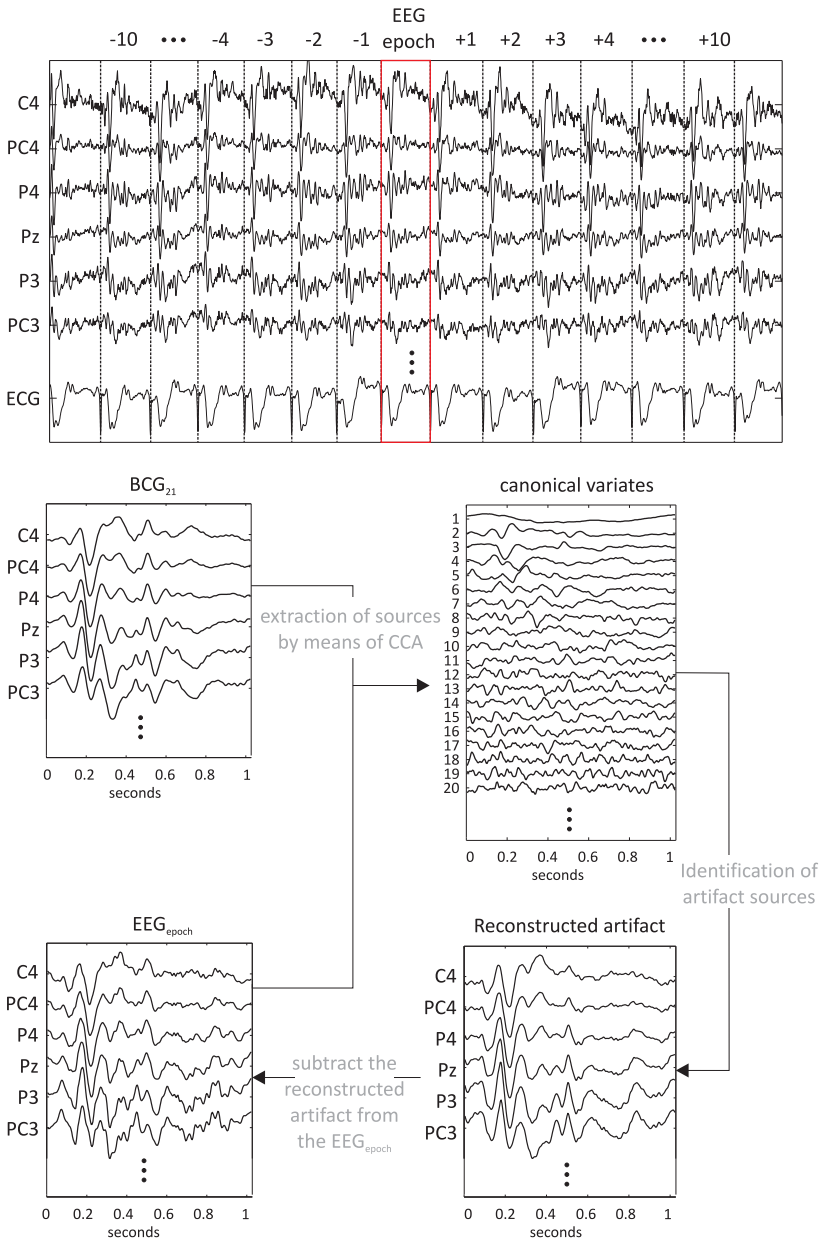


Fig. 1. BCG artifact removal by means of CCA. The EEG signal is epoched around the BCG artifact and an average of 21 epochs (BCG_{21}) around the current EEG is computed. CCA is applied to the current EEG epoch (EEG_{epoch}) and to BCG_{21} to extract the canonical variates. Through several selection criteria, artifact sources are identified and the reconstructed artifact is subtracted from the original signal.

thresholds (tr_{EEG} and tr_{BCG}) for the cumulative redundancy were calculated for the EEG_{epoch} and the BCG_{21}, respectively. tr_{EEG} and tr_{BCG} are the variance explained in the EEG_{epoch} and the BCG_{21}, respectively, by an average BCG artifact calculated from all the data (BCG_{all}). BCG_{all} is a better approximation of the deterministic component than BCG_{21}, because of the higher number of the averaged epochs. The procedure is shown in Fig. 2.

The artifact components are then selected as follows:

1. The variates with a canonical correlation equal or higher than 99%, to account for rounding errors, are labeled as artifacts.
2. The variates, ordered according to increasing R_{EEG}, for which the cumulative redundancy reaches tr_{EEG}, are labeled as artifacts;
3. The variates, ordered according to increasing R_{BCG}, for which the cumulative redundancy reaches tr_{BCG}, are labeled as artifacts;
4. If a component is labeled as artifact by all three previous tests, it is selected as actual artifact source.

2.7. Removal of eye blinks

The removal of eye blinks is an important issue in ERP studies. Since any artifact removal technique may affect the residual data and, therefore, any following validation, we decided to identify trials contaminated by eye blinks and to reject them before proceeding to the validation steps. Single epochs were labeled as rejected if the maximum amplitude of the EOG channel in that epoch exceeded a given dataset-dependent threshold. The threshold was automatically calculated as follows: the EOG channel was squared and the mean amplitude and standard deviation of the amplitude distribution were calculated. A threshold equal to the mean amplitude plus three times the standard deviation was set, as it can be

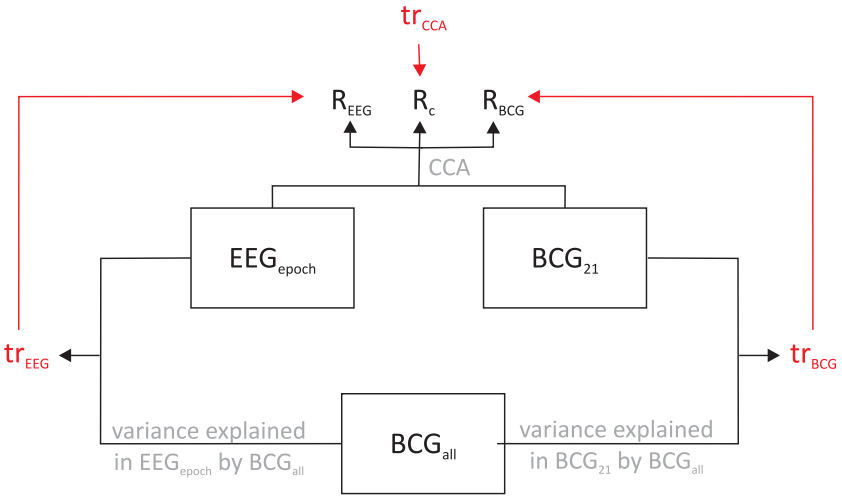


Fig. 2. Identification of artifact sources. Two thresholds, tr_{EEG} and tr_{BCG} are calculated as the variance explained by the BCG artifact (BCG_{all}) in EEG_{epoch} and BCG_{21}, respectively. These thresholds, together with tr_{CCA} (i.e. 99%) are used to threshold R_{EEG}, R_{BCG} and R_C and identify the artifact components.

proven that such an interval contains at least 89% of the data, regardless of the underlying distribution (Delorme, 2006).

2.8. ERP extraction

2.8.1. Detection task

Each dataset was re-referenced to the average signal recorded at electrode locations TP9 and TP10, as the locations closest to the mastoids (not recorded in the Leuven setup). The re-referenced data were segmented for the left (stimulus presented in the upper or lower left field of view) and the right (stimulus presented in the upper or lower right field of view) condition, in a window of 100 ms before and 400 ms after stimulus onset. The trials were baseline corrected (100 ms before stimulus onset) and averaged separately for the two conditions. Then the ERPs for the left and right stimulus were averaged electrode-wise across hemispheres to create a contralateral and ipsilateral condition. For subsequent analysis, the average across electrodes in the occipital region was used, as shown in Supplementary Figure S2 (DiRusso et al., 2001). On this average, the contra- and ipsilateral P1 and N1 components were identified.

2.8.2. Go–Nogo task

Each dataset was re-referenced to the average signal recorded at preauricular electrodes (A11 and A12) and segmented separately for the Go and the Nogo condition, with a window of 100 ms before and 600 ms after stimulus onset and baseline corrected (100 ms before stimulus onset). The trials were averaged separately for the two conditions. Since we are mainly interested in the N2 and P3 peaks, only midline channels are considered for subsequent analysis, as shown in Supplementary Figure S2. The Nogo N2 component was identified on Fz while the Go and Nogo P3 were identified on Pz.

2.8.3. Motor task

Each dataset was re-referenced to the average signal recorded at electrode locations T9 and T10, as the locations closest to the mastoids (not recorded in the Ghent setup). The re-referenced data were segmented for the left (left-hand keypress) and the right (right-hand keypress) condition in a window of 100 ms before and 1000 ms after the stimulus. The trials were baseline corrected (100 ms before stimulus onset) and further segmented around the response (700 ms before and 100 ms after response onset) to obtain response-locked averages. Then the ERPs for the left and right stimulus were averaged electrode-wise across hemispheres to create a contralateral and ipsilateral condition. For subsequent analysis, only the electrodes placed over the motor cortex were considered, as shown in Supplementary Figure S2 and the Lateralized Readiness Potential (LRP) was also calculated as follows:

$$\mathrm{LRP} = \frac{(\mathrm{C3}-\mathrm{C4})_{\mathrm{right}} + (\mathrm{C4}-\mathrm{C3})_{\mathrm{left}}}{2} \tag{4}$$

where C3 and C4 are the responses recorded at the electrodes with labels C3 and C4 while left and right indicate the condition. The LRP results in a negative potential. The most important parameters measured on the LRP are the LRP onset, i.e. the time at which the LRP starts to differ from the baseline, and the LRP peak, i.e. the maximum of the LRP before the response. The onset latency was calculated by means of a regression-based method (Mordkoff and Gianaros, 2000), as the break-point between two intersecting straight lines that are fitted to the LRP waveform. The LRP peak was calculated as the global minimum preceding the response.

2.9. Validation

The validation proposed in this work is fourfold. Firstly, we address the dependence on the electrodes number of the criterion to adaptively select thresholds in CCA. Secondly, we investigate the ability of each method (AAS, OBS and CCA) to extract BCG artifacts that well represent the characteristics of the artifacts in the data. Thirdly, we assess the capability of the three methods to recover a reliable ERP signal. Fourthly, we investigate the effect of the static magnetic field of the MR-scanner on the repeatability of ERPs recorded in different environments.

1. *Adaptive thresholds*. The mean number of removed components and the mean thresholds used by CCA, together with their standard deviation, are calculated for each subject.
2. *Extracted artifact*. The extracted artifact is obtained by subtracting the EEG after AAS-, OBS- and CCA-based artifact removal from the recordings at 3 T. In order to evaluate the quality of the extracted artifact, the data in five different cases (3 T, after AAS, after OBS, after CCA, 0 T), were segmented around the *R*-peak (100 ms before and 800 ms after the *R*-peak) and the epoch mean was subtracted. Then the following quantities were calculated.
   - Mean extracted artifact. In order to obtain a reference BCG artifact, the epochs extracted from the 3 T data were averaged to remove background stochastic EEG activity while

keeping the deterministic component of the artifact. The artifact extracted by the AAS, OBS and CCA method, respectively, was also segmented and averaged to obtain a mean extracted artifact. The mean extracted artifact was then compared to the reference artifact and, for each method, the Pearson's correlation coefficient for each channel, as well as its mean across channels, was calculated. Any difference between the mean extracted artifacts would be related to systematic errors due to the artifact extraction method. For visual inspection, the 0 T data, were also synchronously averaged to the *R*-peaks to show the absence of BCG-related activity outside the scanner. The mean Pearson's correlations were tested for significance with a repeated measurements ANOVA (with a within-subject factor Method (AAS, OBS, CCA) × a between-subject factor Location (Maastricht, Leuven, Ghent)) to assess the effect of different setups on the artifact removal.

- Area Under the Envelope (AUE). For each epoch the maximum and minimum envelope (i.e. the maximum and minimum amplitude across channels for each time point) were calculated, as well as the area encompassed by the two envelopes. Since the BCG artifact has a different polarity on different channels, the AUE represents the amount of residual artifact in the data. The AUE for epoched 0 T data was calculated as a reference. The mean AUE obtained in the 3 T data and in the data after AAS-, OBS- and CCA-based artifact removal were tested for significance with a repeated measurements ANOVA (with a within-subject factor Method (AAS, OBS, CCA) × a between-subject factor Location (Maastricht, Leuven, Ghent)) (Debener et al., 2007).

3. *Effect of BCG artifact removal on ERPs*. Once the artifact was removed, the residual EEG was inspected to assess the reliability of the extracted ERP waveform. The data were segmented according to paragraph Section 2.8 and the quality of the grand-average (across subjects) and the single-subject averages was evaluated. For the Detection and the Go–Nogo tasks, only data recorded in Maastricht were included in the analysis, because of the different electrode setups.
   - Grand-averages were calculated for the three different tasks. On this wave, meaningful peaks, as described in Section 2.8, were measured. For each method, the peaks were considered present if they were statistically different from the baseline (*z*-score higher than 1.960 times the standard deviations, $p < 0.05$). Scalp maps were plotted at the latencies of the measured peaks. For each peak, scalp maps obtained in different situations (0 T, 3 T, AAS, OBS, CCA) were compared by means of a repeated measurements ANOVA (Electrodes × Method). Before applying ANOVA, the electrodes were grouped according to their positions and the mean voltages over each group normalized by subtracting the minimum value and dividing by the difference between the maximum and minimum voltage. A list of the groups and the electrodes assigned to each group is reported in Supplementary Figure S3. A significant interaction between Electrode and Method would mean a significant difference between scalp maps (Picton et al., 1995, 2000).
   - ERP modulation. For the Detection and the Go–Nogo tasks amplitude differences between conditions at the peak's latency were tested for significance, for each peak, by means of a paired two-tailed *t*-test. No correction was used, since each peak and each method were considered independent from the others. The procedure was repeated for each method to assess the ability of the method to reproduce the experimental modulation, as obtained outside the MR-scanner. The latencies of meaningful peaks on the single-subject data were derived as the latency of the global maximum or minimum in a time window of 120 ms around the peak manually marked on the grand-average. The peak amplitude was calculated as the mean amplitude over 12 ms around the peak latency.

4. *Reproducibility of ERPs in different environments*. Differences between amplitude and latency of meaningful peaks at 0 T and in the 3 T, AAS, OBS and CCA cases were tested for significance (paired *t*-test). Moreover, differences between the reaction times outside and inside the MR-scanner were tested for significance in each subject (two-tailed paired *t*-test). Results are labeled as significant ($p < 0.05$), very significant ($p < 0.01$) and highly significant ($p < 0.001$).

3. Results

3.1. Adaptive thresholds

For each artifact occurrence, two thresholds were determined: one for the EEG_{epoch} and one for the BCG_{21}. Fig. 3 shows the histograms for these two thresholds, in red and white, respectively, for one subject. Table 2 reports the mean and standard deviation for the two thresholds for each dataset, as well as the mean number of components removed by CCA. The mean thresholds determined for the EEG_{epoch} fragments are lower than those determined for the BCG_{21} fragments. This agrees with the observation that the EEG epoch contains both brain activity and artifact, while the reference contains mainly artifact. We found a significant main effect ($F_{(2,18)} = 14.58, p < 0.001$) of location on the number of removed components, being lower in Ghent than in Maastricht or Leuven, while no significant difference was found between the Leuven and Maastricht datasets.

3.2. Extracted BCG artifact

An example of the mean extracted artifact is shown in Fig. 4: the average artifact extracted by the AAS, OBS and CCA method is superimposed to the mean artifact obtained by averaging the 3 T data. The superposition of the mean artifact for 0 T and 3 T data is also shown.

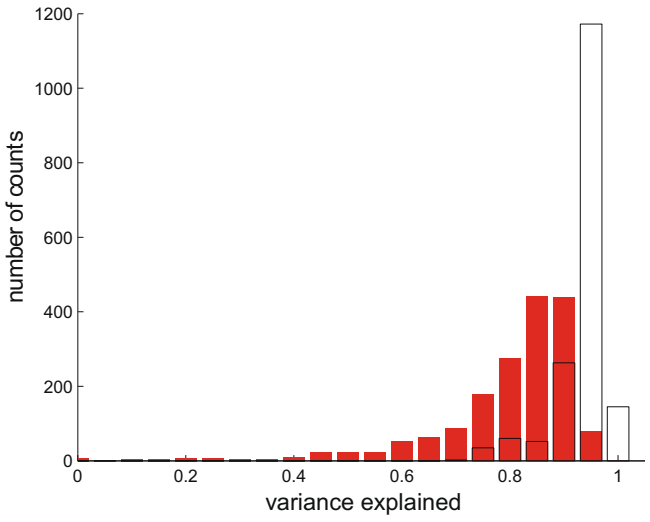


Fig. 3. Histograms of the thresholds for the EEG fragment (red bars) and the reference (white bars) adaptively determined by CCA. (For interpretation of color mentioned in this figure the reader is referred to the web version of the article.)

Table 2
The adaptively determined thresholds are reported with their mean and standard deviation, as well as the mean number of components removed by CCA.

Dataset	Location	Removed components	Threshold EEG_{epoch}	Threshold EEG_{ref}
1	M	6 ± 2	0.80 ± 0.14	0.94 ± 0.05
2	M	5 ± 2	0.75 ± 0.10	0.92 ± 0.02
3	M	6 ± 3	0.81 ± 0.11	0.96 ± 0.04
4	M	6 ± 2	0.74 ± 0.13	0.96 ± 0.01
5	M	7 ± 3	0.77 ± 0.16	0.96 ± 0.02
6	L	7 ± 2	0.81 ± 0.14	0.97 ± 0.01
7	L	7 ± 2	0.84 ± 0.12	0.97 ± 0.01
8	M	6 ± 2	0.78 ± 0.14	0.94 ± 0.05
9	M	6 ± 2	0.75 ± 0.13	0.93 ± 0.02
10	M	6 ± 2	0.82 ± 0.11	0.94 ± 0.04
11	M	6 ± 2	0.74 ± 0.13	0.96 ± 0.02
12	M	7 ± 3	0.82 ± 0.12	0.97 ± 0.02
13	L	6 ± 2	0.74 ± 0.23	0.96 ± 0.01
14	L	7 ± 2	0.87 ± 0.11	0.97 ± 0.01
15	G	5 ± 2	0.73 ± 0.11	0.96 ± 0.01
16	G	5 ± 3	0.74 ± 0.16	0.96 ± 0.03
17	G	5 ± 3	0.56 ± 0.18	0.92 ± 0.05
18	G	5 ± 3	0.68 ± 0.24	0.95 ± 0.02
19	G	5 ± 3	0.74 ± 0.18	0.96 ± 0.01
20	G	5 ± 2	0.79 ± 0.13	0.98 ± 0.01
21	G	6 ± 2	0.89 ± 0.08	0.98 ± 0.01

The average Pearson correlation across channels and across datasets is shown in Fig. 5A: even though all the single-channel correlations are significant ($p < 0.01$), we found a main effect of Location ($F_{(2,36)} = 11.033, p < 0.01$) and Method ($F_{(2,36)} = 43.513, p < 0.001$). A follow-up analysis revealed that the mean correlation achieved with the data recorded in Ghent was lower than the one achieved with data recorded in Maastricht or Leuven ($p < 0.001$) while the correlation achieved with the Maastricht data was higher than with

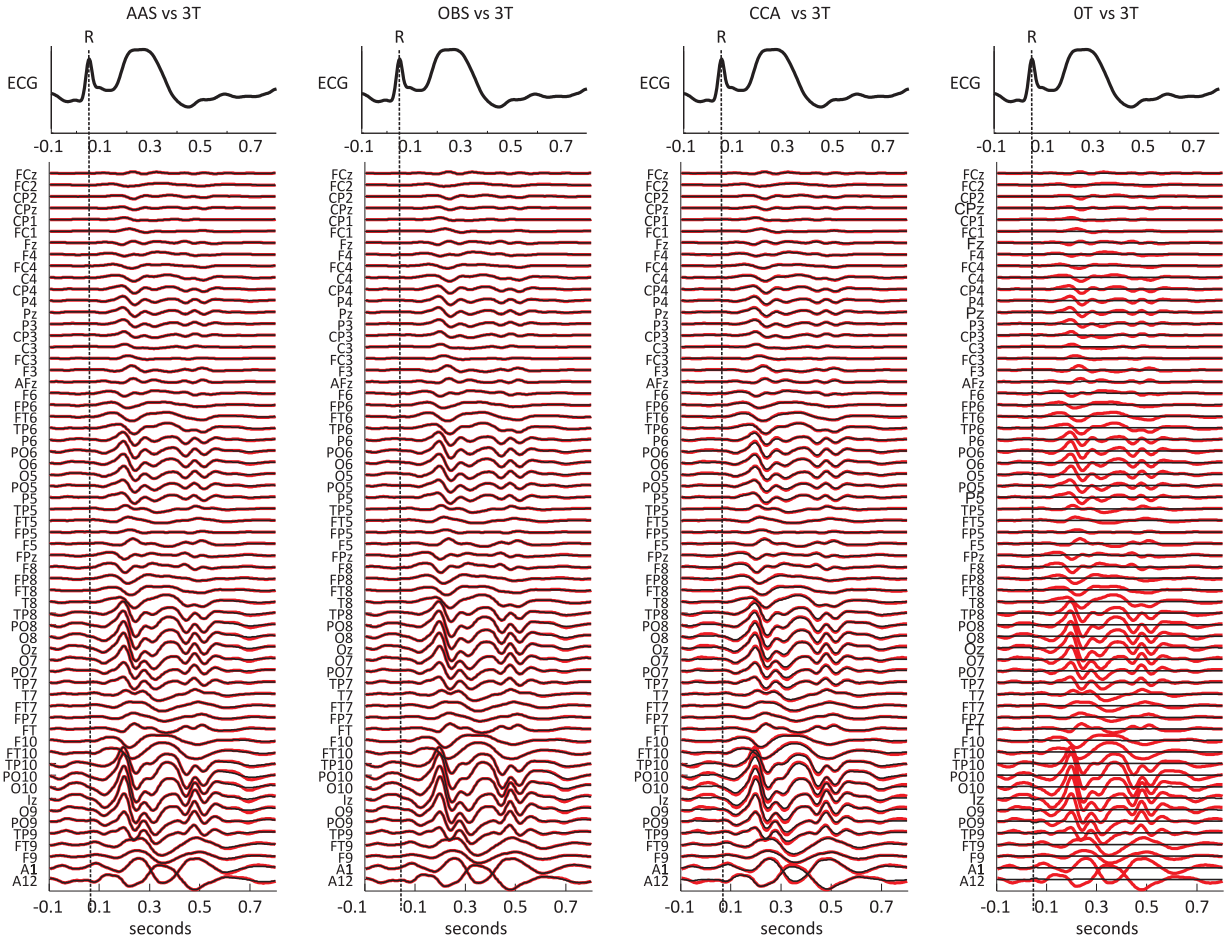


Fig. 4. Comparison of the mean BCG (red thick line) and the mean BCG extracted by AAS, OBS, CCA (black thin line). The 0 T data, epoched around the *R*-peak and averaged, are also shown. In the top of the figure, the mean ECG is reported as well as the marked *R*-peak. (For interpretation of color mentioned in this figure the reader is referred to the web version of the article.)

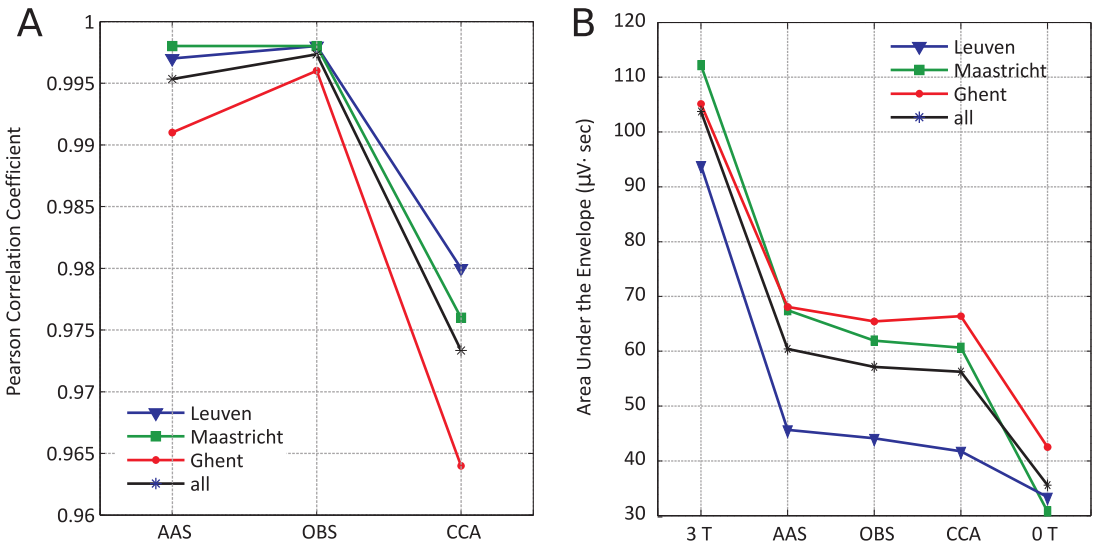


Fig. 5. (A) Pearson correlation coefficient between the mean BCG artifact and the mean artifact extracted by AAS, OBS, CCA, averaged across channels and across subjects. The means of the datasets grouped by location are shown, as well as the global mean. (B) Area under the envelope calculated from the data before (3 T) and after artifact removal by means of AAS, OBS, CCA, as well as from the data recorded outside the MR-scanner. The values are averaged across channels and across subjects. The means of the datasets grouped by location are shown, as well as the global mean.

the Leuven data ($p < 0.001$). Moreover, the correlation achieved by AAS-based artifact removal was significantly lower than that one achieved by OBS-based artifact removal. Correlation achieved by CCA was significantly lower than those achieved by both AAS-based and OBS-based artifact removal.

Fig. 5B shows the mean area under the envelope, averaged across epochs and across subjects, for the five different situations. When tested for significance we found a main effect ($F_{(4,72)} = 75.977, p < 0.001$) of Method. Follow-up *t*-tests showed that the area under the envelope is always significantly higher in 3 T compared to the other cases. In 0 T it is always significantly smaller. No significant differences were found between the area obtained after AAS- OBS- or CCA-based artifact removal.

3.3. Effect of BCG artifact removal on ERPs

3.3.1. Detection task

Fig. 6 shows the grand-averages of the Detection task based on 5 datasets, for the ipsilateral and the contralateral hemisphere for each situation (0 T, 3 T, AAS, OBS, CCA). Two main peaks are iden-

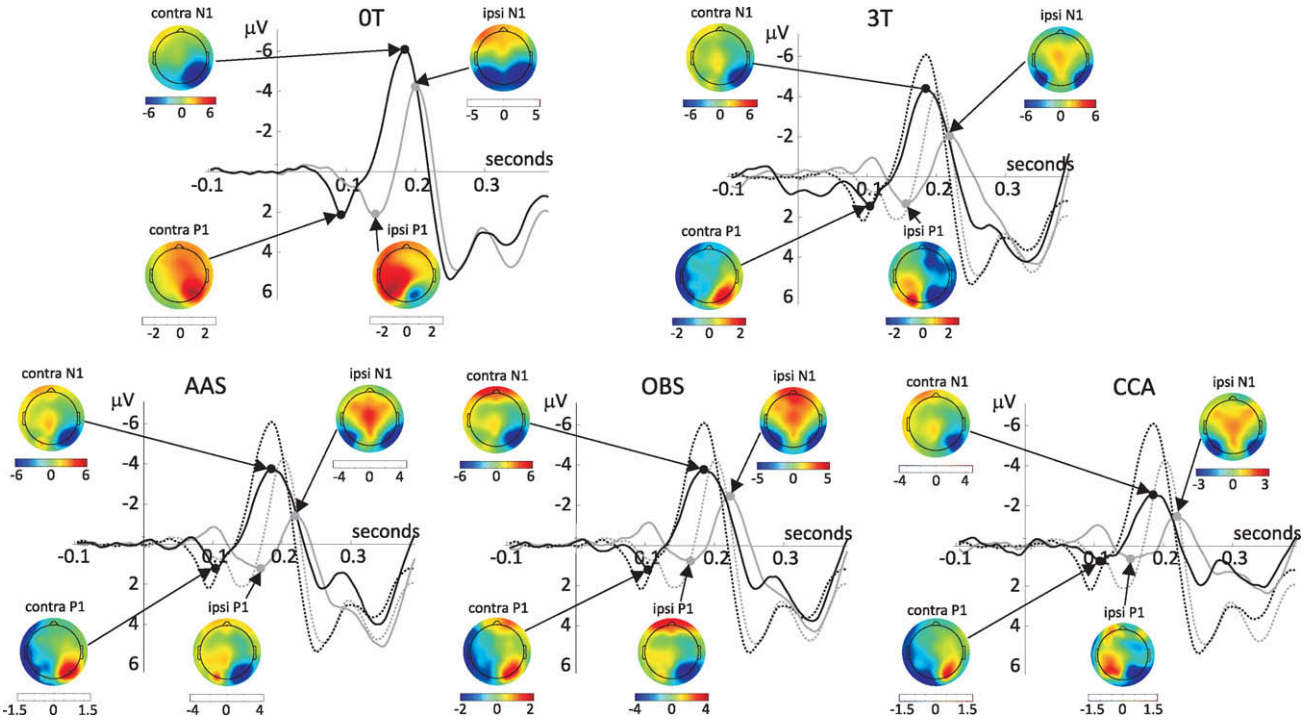


Fig. 6. Grand-averages (of the ERP averaged across occipital channels) obtained in the Detection task for the contralateral (black line) and the ipsilateral (gray line) condition in four cases (0 T, 3 T, AAS, OBS, CCA). In the 3 T, AAS, OBS and CCA cases, the 0 T ERPs are superimposed (dotted lines). The N1 and P1 components are identified, marked by dots, and scalp maps are drawn in correspondence of their latencies.

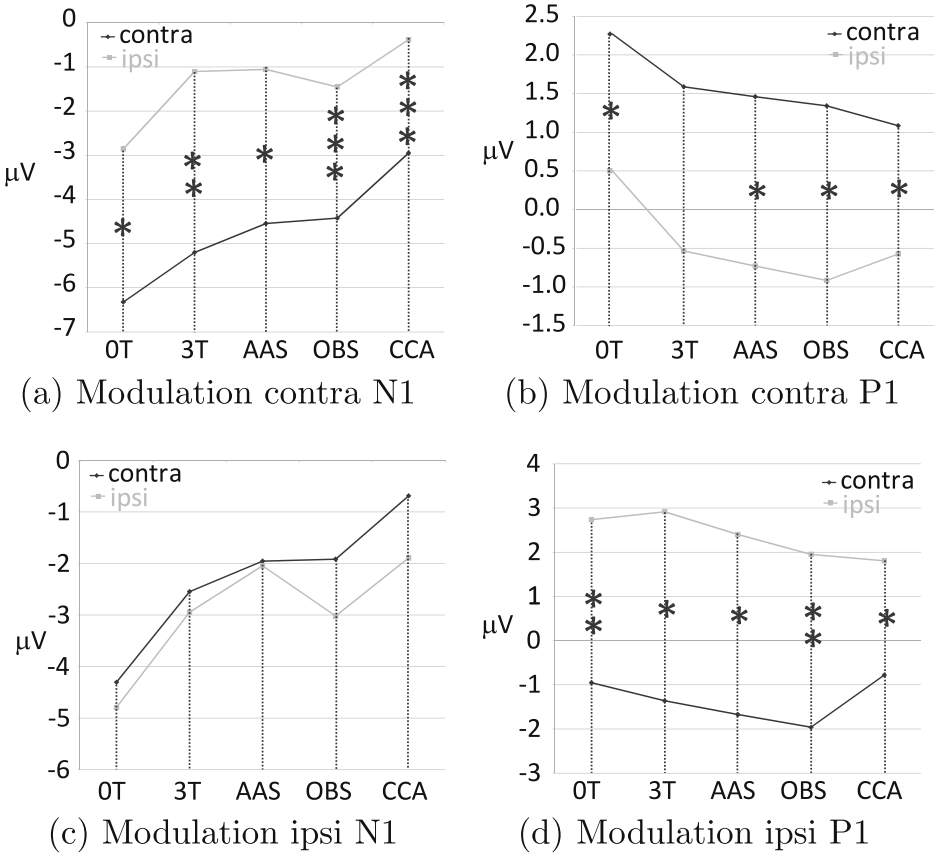


Fig. 7. Modulation effect for N1 and P1 on the contra- and ipsilateral condition. Significant differences are marked by stars (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$).

tified, namely the P1 (around 100 ms) and the N1 (around 180 ms). The peak amplitudes were always significantly different from the baseline (z-score $> 1.960 * \sigma, p < 0.05$). Scalp maps are drawn at the latencies of those peaks. The scalp distribution of the ipsilateral N1 at 0 T was significantly different from the scalp distribution at 3 T ($F_{(9,80)} = 2.135, p < 0.05$) and after AAS-based artifact removal ($F_{(9,80)} = 2.175, p < 0.05$).

The modulation effect is reported in Fig. 7: significant differences are marked by the stars. The modulation effect found in the 0 T data for the contralateral N1 and P1 peaks and the ipsilateral P1, was also found in the data after AAS- OBS- or CCA-based artifact removal. The modulation of the contralateral P1 was not found in the 3 T data.

3.3.2. Go–Nogo task

Fig. 8 shows the grand-averages of the Go–Nogo task based on 5 datasets, for the Go and the Nogo condition for each situation (0 T, 3 T, AAS, OBS, CCA). Two main peaks are identified, namely the N2 (between 200 and 300 ms) and the P3 (around 300 ms). The peak amplitudes were always significantly different from the baseline (z-score $> 1.960 * \sigma, p < 0.05$). Scalp maps are drawn at the latencies of those peaks. The scalp distribution of the Go P3 component at 0 T was significantly different from the scalp distribution at 3 T ($F_{(9,80)} = 4.885, p < 0.001$) and after AAS-based ($F_{(9,80)} = 4.216$, $p < 0.001$), OBS-based ($F_{(9,80)} = 3.034, p < 0.01$) and CCA-based ($F_{(9,80)} = 2.728, p < 0.01$) artifact removal.

The modulation effect is reported in Fig. 9: significant differences are marked by the stars. A modulation effect for the Nogo N2 peak was only found in the 0 T data, while a modulation of the Nogo P3 was only found in the data after CCA-based artifact removal.

3.3.3. Motor task

Fig. 10 shows the grand-averages of the ipsilateral and contralateral condition for the Motor task, when 7 datasets are included in the analysis. The ipsilateral and contralateral channels, as well as the extracted LRP are shown. Two main points are identified, namely the LRP onset and the LRP peak. The LRP amplitude at the peak was always significantly higher than the baseline (z-score $> 1.960 * \sigma, p < 0.05$) while the LRP amplitude at the onset did not differ from the baseline. Scalp maps are drawn at the latencies of the LRP onset and peak. The scalp distribution at the onset latency at 0 T was significantly different from the scalp distribution after CCA-based artifact removal ($F_{(9,120)} = 3.403, p < 0.01$). It is worth mentioning that an actual peak (defined as a global maximum) was not found in the LRP after AAS- and OBS-based artifact removal, as it is visible in Fig. 10. Therefore, the peak was considered at the latency of the response.

3.4. Reproducibility of ERPs in different environments

In the Detection task we found a significant latency shift for the ipsilateral N1 component in the data after OBS-based artifact removal ($t_{(4)} = 3.33, p < 0.05$), when compared to the 0 T data. A significant decrease in amplitude was found for the contralateral N1 ($t_{(4)} = 2.81, p < 0.05$), the ipsilateral P1 ($t_{(4)} = 4.10, p < 0.05$) and the ipsilateral N1 ($t_{(4)} = 3.78, p < 0.05$) component after CCA-based artifact removal and for the ipsilateral N1 component after AAS-based artifact removal ($t_{(4)} = 3.97, p < 0.05$).

In the Go–Nogo task we found a significant latency shift for the Go P3 peak in the 3 T data ($t_{(4)} = 3.55, p < 0.05$) and in the data after AAS- ($t_{(4)} = 3.56, p < 0.05$), OBS- ($t_{(4)} = 3.55, p < 0.05$) and CCA-based ($t_{(4)} = 3.52, p < 0.05$) artifact removal, when compared

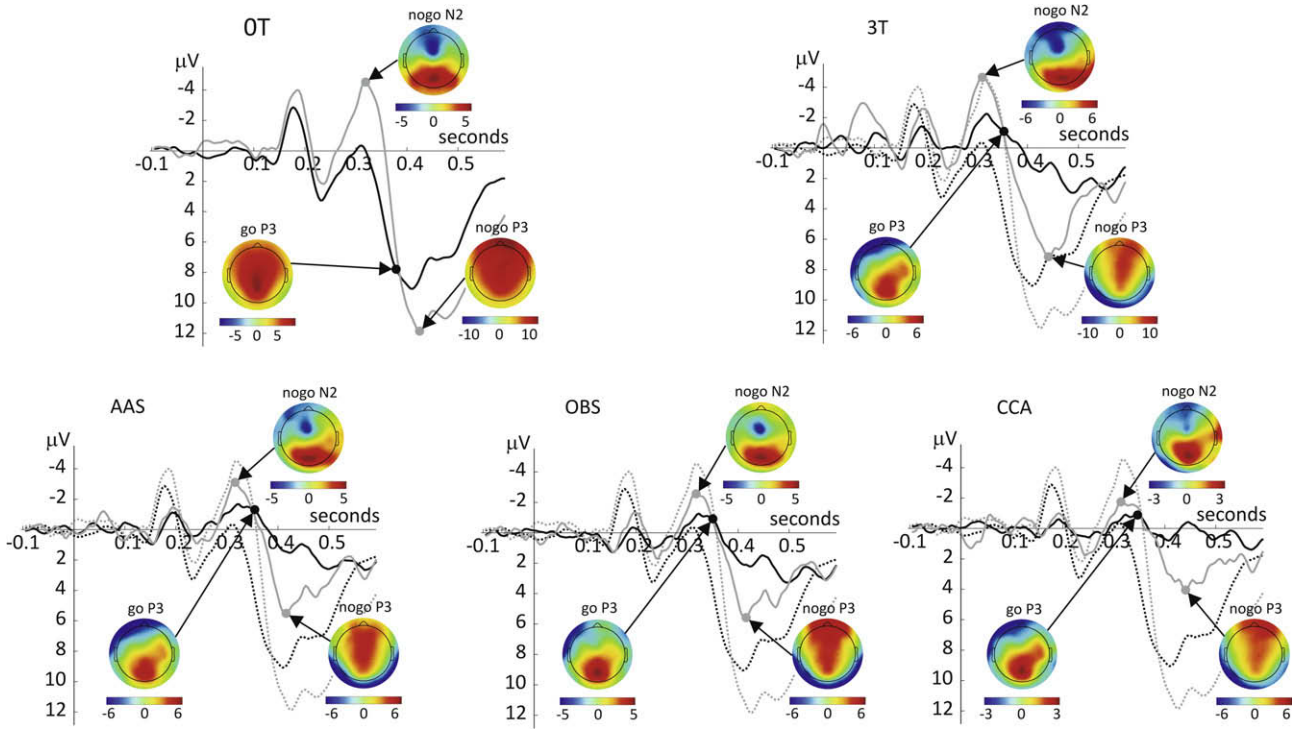


Fig. 8. Grand-averages (Fz channel) obtained in the Go–Nogo task for the Go (black line) and the Nogo (gray line) condition in four cases (0 T, 3 T, AAS, OBS, CCA). In the 3 T, AAS, OBS and CCA cases, the 0 T ERPs are superimposed (dotted lines). The N2 and P3 components are identified, marked by dots, and scalp maps are drawn in correspondence of their latencies.

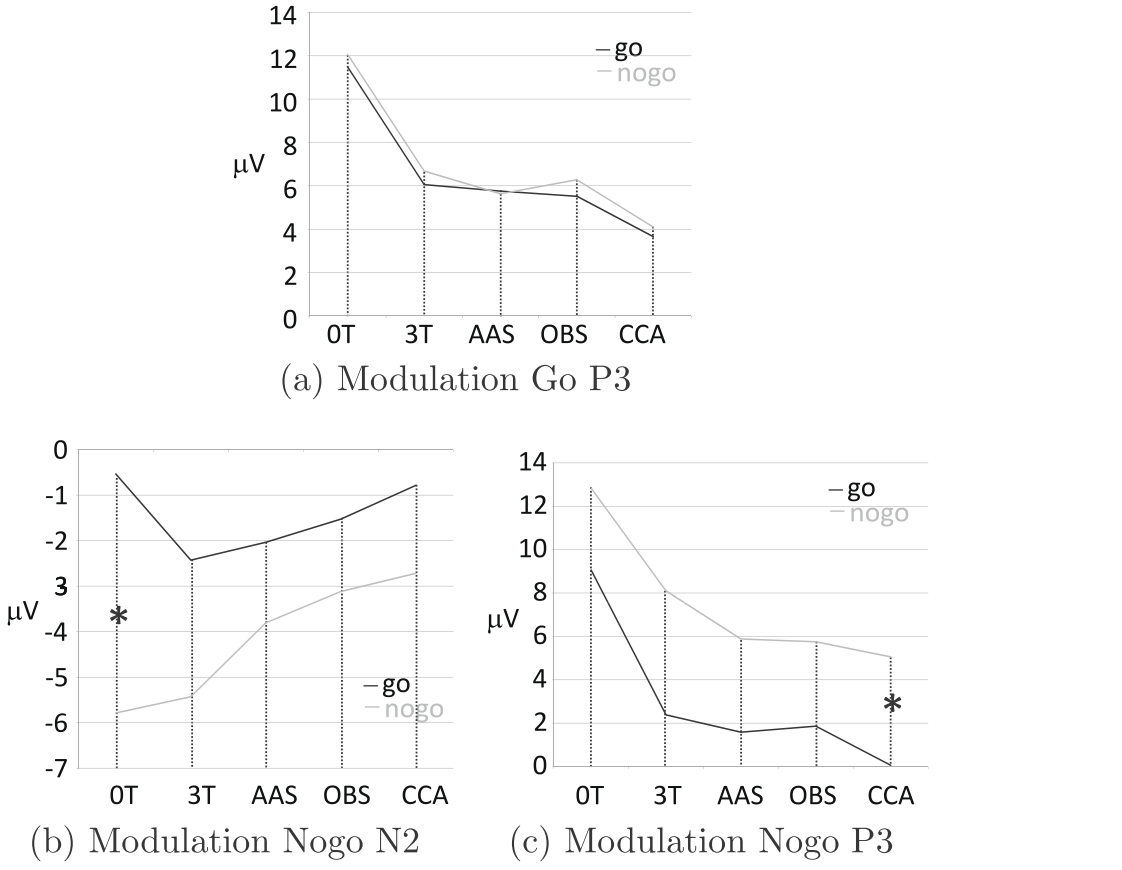


Fig. 9. Modulation effect for N2 and P3 on the Go and Nogo condition. Significant differences are marked by stars (***$p < 0.001$, **$p < 0.01$, *$p < 0.05$).

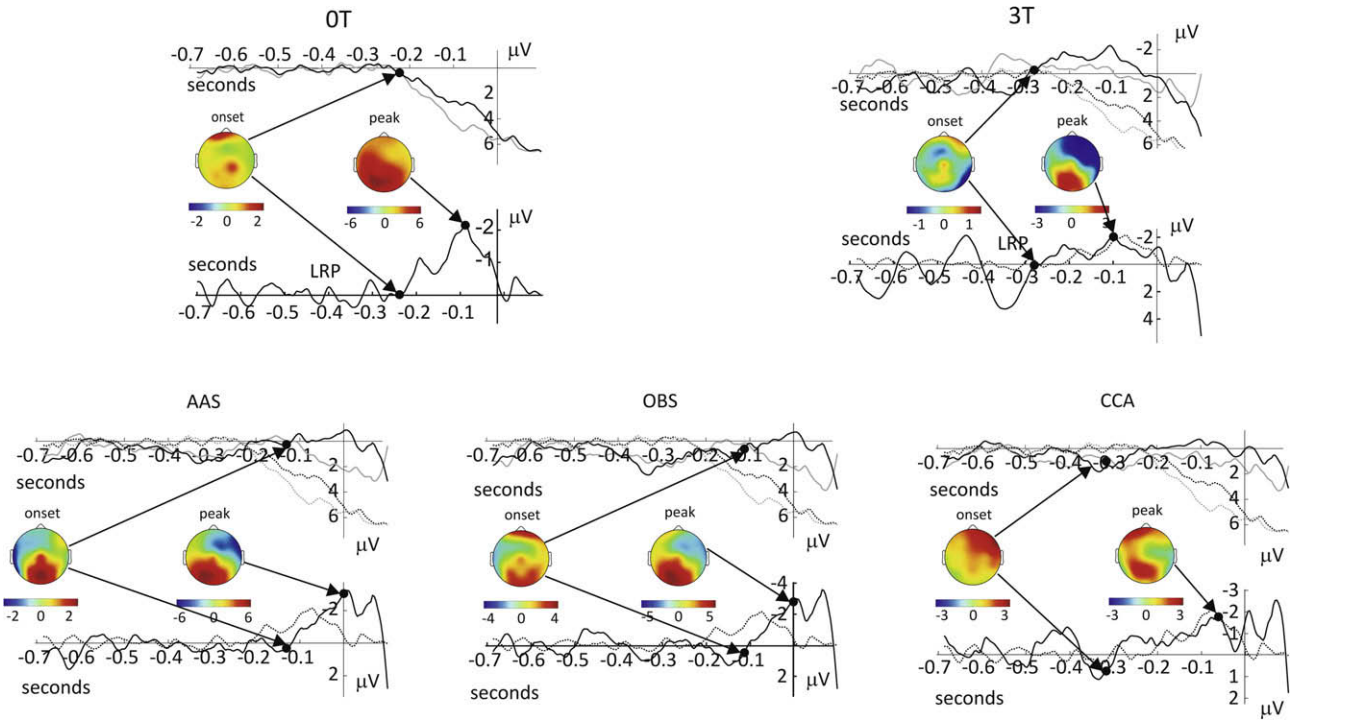


Fig. 10. Grand-averages obtained in the Motor task for the contralateral (black line) and the ipsilateral (gray line) condition in four cases (0 T, 3 T, AAS, OBS, CCA). For each case, the LRP is also shown. In the 3 T, AAS, OBS and CCA cases, the 0 T LRP is superimposed (dotted lines). The onset and peak of the LRP are identified, marked by dots, and scalp maps are drawn in correspondence of their latency.

to the 0 T data. For the same peak a significant decrease in amplitude was also found in the 3 T data ($t_{(4)} = 2.97, p < 0.05$) and in the data after AAS- ($t_{(4)} = 3.51, p < 0.05$), OBS- ($t_{(4)} = 3.44, p < 0.05$) and CCA-based ($t_{(4)} = 4.56, p < 0.05$) artifact removal, when compared to the 0 T data. A significant latency shift was found for the Nogo P3 peak only in the 3 T data ($t_{(4)} = 3.58, p < 0.05$).

In the Motor task we did not find any significant latency or amplitude difference for the LRP onset or peak.

The differences in reaction times in the two environments (0 T vs 3 T) are reported in Table 1. In the Detection task three out of five subjects (recorded in Maastricht) had reaction times significantly longer inside the MR-scanner than outside, while in two subjects (Leuven data) the reaction times were significantly longer outside the scanner. Similar results were found in the Go–Nogo task. In the Motor task six out of seven subjects had reaction times outside the MR-scanner significantly shorter than inside.

4. Discussion

In this work we compared the AAS-, OBS- and CCA-based BCG artifact removal techniques, for recovering ERPs from simultaneous EEG–fMRI data. The AAS- and OBS-based techniques were chosen as they were already successfully used in routine EEG and ERPs and because they are fully automated. The CCA-based technique proposed here is also fully automated. This guarantees a fair evaluation of the compared algorithms.

Three different tasks, namely a visual Detection task, a Go–Nogo task and a Motor task, were used to assess the quality of ERPs recorded in a magnetic environment, when the BCG artifact was removed. The choice of the different ERPs allowed us to test the artifact removal procedures on peaks elicited by different mechanisms. The Detection task elicits early latency components with a clear occipital topography (N1, P1) and a small latency variability which are, therefore, more stable across subjects. The Go–Nogo task elicits two cognitive components (N2, P3) fronto-centrally with a higher latency variability across subjects. The Motor task elicits a low-amplitude low-frequency component over the motor cortex. These components are considered as representative of the possible ERP features, that one would like to extract from EEG–fMRI recordings. The same tasks were performed outside the MR-scanner (0 T) and inside the MR-scanner but without MR-gradients (3 T).

ERPs during actual fMRI were not considered, to avoid the additional necessary step of gradient artifact removal. The removal of gradient artifacts would affect the data to some extent, and would complicate the interpretation of the results.

4.1. Extracted BCG artifact

The mean extracted artifact gives information about the average behavior of the methods. The artifact can be seen as the sum of two contributions, a deterministic one, common to all artifact occurrences, and a random one, variable from artifact to artifact.

We showed that the mean extracted artifacts are significantly correlated ($p < 0.001$) with the average artifact, regardless of the method used for the artifact removal. This allows us to state that none of the methods considered produced systematic errors. Moreover, there was a small (0.01) but noticeable decrease in the mean correlation obtained with CCA, when compared to AAS or OBS. This difference is due to the higher variability of the artifact extracted by CCA. However, the fact that the correlations are significant ensures that the CCA-extracted artifact still well represents the average BCG.

The average artifact captures the deterministic component of the BCG while it ignores the random variations on the single trial level, that are canceled out by any average measure. For this reason it is important to evaluate not only the mean extracted artifact but to consider the single trial residuals as well. The single trial residuals are evaluated in terms of the area under the envelope: this measure is sensitive to the epoch-related variability of the residu-

als, that is not taken into account by the average EEG after artifact removal. We found a significant effect of the method on the reduction of the area under the envelope, with the value after artifact removal being lower than at 3 T and higher than at 0 T. The three artifact removal techniques did not yield significant differences.

Our datasets contained a different number of electrodes. This motivated the development of a dataset-dependent threshold for CCA. Our criterion is able to adapt to the data, as confirmed by the differences between the histograms reported in Fig. 3, as well as the significant main effect of location on the number of removed components. The variance explained in the EEG_{epoch} is lower than the mean variance explained in the BCG_{21}: this is due to the fact that BCG_{21} contains mainly the artifact while EEG_{epoch} contains also background EEG that must not be removed.

We found an effect of location on the performance of the methods. This is due to the fact that the different laboratories use different MR-scanners and different setups. The BCG artifact is due to the static magnetic field of the MR-scanner: differences in the manufacturing of the scanner may influence the properties of the static magnetic field and, therefore, the BCG artifact.

The use of a reference signal allows us to combine the advantages of both the AAS and CCA methods: on one hand, AAS extracts the deterministic component of the artifact, on the other hand, CCA is more sensitive in capturing the intra-subject variability and the variable component of the artifact by processing the data on an epoch-based basis and by adapting the thresholds to the subject and to the epoch.

4.2. Recovered ERP

The quality of ERPs extracted from simultaneous EEG–fMRI recordings was assessed by investigating the possibility of reproducing the same topographies as obtained in 0 T data for a given peak. Moreover, the capability of obtaining the same experimental modulation as in the 0 T data was considered.

We found that the removal of the BCG artifact is mandatory if one wants to recover ERPs from simultaneous EEG–fMRI. The visual inspection of the grand-average waveforms shows that the BCG artifact heavily contaminates the data, making the waveforms and the topographies not always reproducible. This is especially the case when a limited number of trials are averaged, such as in the Nogo condition of the Go–Nogo and in Motor task. Meaningful ERPs can almost always be recovered from simultaneous EEG–fMRI recordings in the Detection and the Go–Nogo tasks, regardless of the technique used to remove the BCG artifact. However, only after CCA-based artifact removal we could identify the LRP peak in the time window preceding the response. The LRP is a small-amplitude low-frequency potential, suggesting that CCA is more effective when small components must be extracted from a smaller amount of trials.

4.3. Reproducibility of ERPs in different environments

There is no general agreement that the ERPs recorded outside the scanner are perfectly reproducible inside a magnetic field. To the authors' knowledge, only few studies addressed the same problem in a 3 T MR-scanner (e.g. Kruggel et al. (2000), Sadeh et al. (2008), Strobel et al. (2008)). The recording conditions during an EEG–fMRI experiment are very different from a normal ERP recording experiment. The noise of the MR scanner, the different ambient light, the position of the subject inside the magnet and the magnetic fields make the MR environment intrinsically hostile to ERP recordings that require, by definition, a quiet laboratory with controlled environmental conditions in terms of light, noise and position of the subject with respect to the presentation screen. Moreover, it has already been shown that the strong static magnetic field has an influence on the data (Debener et al., 2008), in terms of amplitude of the BCG artifact. This effect becomes more pronounced with increasing static magnetic field.

In the Go–Nogo task we found significant differences in latency and amplitude for the Go P3, regardless of the method used for artifact removal. This might suggest that some brain path may be affected by the different environment. In the Detection and the Motor task, a latency shift was visible on the grand-average but not significant: this might be due to the limited number of subjects included in the study. Similar findings have been reported previously at 1.5 T (Mulert et al., 2004). In order to answer whether it is an effect of the static magnetic field or the different position of the subject, the same experiment should be performed in a dummy scanner. Koch et al. (2003) tested the differences in reaction times in three situations, namely in an ERP laboratory (behavioral condition), in a simulated MRI (participant in a lying position on a stretcher) and in conventional fMRI (operating scanner): they found that the situation in which the experiment was performed had a main effect on the reaction times. Specifically, reaction times in the behavioral group were shorter than in the conventional MRI and the simulated MRI was in between. We obtain similar results in the reaction times, when tested separately for each subject. Whether this effect is related to the different position of the subject with respect to traditional behavioral studies or to the static magnetic field is still unclear. This is a relevant issue that deserves further investigation in order to clarify the equivalence of cognitive processes in both behavioral and imaging studies and to correctly relate new multi-modal results with former literature studies.

5. Conclusion

In this work we investigated the effect of the static magnetic field of the MR-scanner on the quality of ERPs extracted from simultaneously recorded EEG–fMRI.

We compared three fully automated methods to remove the BCG artifact, already proven successful when working with routine EEG. We showed that, when the number of averaged trials is high, AAS, OBS and CCA methods performed equally good. When less trials are averaged, CCA-based artifact removal obtained better results, especially when dealing with small low-frequency components such as the LRP. The proposed CCA-based artifact removal takes into account the intra-subject variability of the BCG artifact through an epoch-based approach and adaptively determined thresholds.

The reproducibility of the ERPs in different recording environments (0 T vs 3 T) was also considered. However, further investigations are needed in order to assess to what extent changes in the recording environment affect the extracted ERP.

Acknowledgement

K. Vanderperren is supported by an IWT PhD grant. Her research, together with that of S. Van Huffel, is supported by GOA-AMBioRICS, GOA MaNet, IUAP P6/04, FWO project G. 0360.05 and Neuromath (COST-BM0601). N. Novitskiy is a post-doctoral fellow of the Fund for Scientific Research, Flanders (FWO). All aforementioned authors, together with S. Sunaert, J. Ramautar and P. Stiers, are supported by the K.U.Leuven Research Fund (K.U.Leuven Onderzoeksfonds) IDO 05/010 EEG–fMRI. The authors would like to thank Sven Gijsen, MR technician at the Maastricht Brain Imaging Centre (MBIC), Ronald Peeters, Physicist at the Department of Radiology University Hospitals Leuven, and Pieter Vandemaele, support engineer of the Ghent Institute of functional and Metabolic Imaging (GIfMI), Ghent University Hospital, for their help in setting

up the equipment and the experiments at Maastricht, Leuven and Ghent location, respectively.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.clinph.2009.12.032.

References

Allen PJ, Polizzi G, Krakow K, Fish DR, Lemieux L. Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. Neuroimage 1998;8:229–39.

Assecondi S, Hallez H, Staelens S, Bianchi AM, Huiskamp GM, Lemahieu I. Removal of the ballistocardiographic artifact from EEG–fMRI data: a canonical correlation approach. Phys Med Biol 2009;54:1–17.

Becker R, Ritter P, Moosmann M, Villringer A. Visual evoked potentials recovered from fMRI scan periods. Human Brain Mapping 2005;26:221–30.

Bénar CG, Aghakhani Y, Wang Y, Izenber A, Al-Asmi A, Dubeau F, et al. Quality of EEG in simultaneous EEG–fMRI for epilepsy. Clin Neurophysiol 2003;114:569–80.

Bregadze N, Lavric A. ERP differences with vs. without concurrent fMRI. Int J Psychophysiol 2006;62(1):54–9.

Briselli E, Garreffa G, Bianchi L, Bianciardi M, Macaluso E, Abbafati M, et al. An independent component analysis-based approach on ballistocardiogram artifact removing. Magn Reson Imaging 2006;24:393–400.

Comi E, Annovazzi P, Silva AM, Cursi M, Blasi V, Cadioli M, et al. Visual evoked potentials may be recorded simultaneously with fMRI scanning: a validation study. Human Brain Mapping 2005;24:291–8.

Cooley W, Lohnes P. Multivariate data analysis. John Wiley & Sons, Inc.; 1971.

De Clercq W, Vergult A, Vanrumste B, Van Paesschen W, Van Huffel S. Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. IEEE Trans Biomed Eng 2006;53:2583–7.

Debener S, Mullinger KJ, Niazy RK, Bowtell RW. Properties of the ballistocardiogram artefact as revealed by EEG recordings at 1.5, 3 and 7 T static magnetic field strength. Int J Psychophysiol 2008;67:189–99.

Debener S, Strobel A, Sorger B, Peters J, Kranczioch C, Engel A, et al. Improved quality of auditory event-related potentials recorded simultaneously with 3 T fMRI: removal of the ballistocardiogram artefact. Neuroimage 2007;34:587–97.

Delorme A. Encyclopedia of medical device and instrumentation 2006;vol. 6. Wiley Interscience; 2006 [chapter Statistical methods, p. 240–64].

Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial dynamics including independent component analysis. J Neurosci Methods 2003;134:9–21.

Di Russo F, Martínez A, Sereno MI, Pitzalis S, Hillyard SA. Cortical sources of the early components of the visual evoked potential. Hum Brain Mapp 2001;15:95–111.

Dyrholm M, Goldman R, Sajda P, Brown TR. Removal of BCG artifacts using a non-kirchhoffian overcomplete representation. IEEE Trans Biomed Eng 2009;56(2):200–4.

Eimer M. The lateralized readiness potential as an on-line measure of central response activation processes. Behav Res Meth Instrum Comput 1998;30(1):146–56.

Ellingson ML, Liebenthal E, Spanaki MV, Prieto TE, Binder JR, Ropella KM. Ballistocardiogram artifact reduction in the simultaneous acquisition of auditory ERPS and fMRI. Neuroimage 2004;22(4):1534–42.

Goldman RI, Stern JM, Engel Jr J, Cohen MS. Acquiring simultaneous EEG and functional MRI. Clin Neurophysiol 2000;111:1974–80.

Golub G, Van Loan C. Matrix computations. third ed. Baltimore: The Johns Hopkins University Press; 1996.

Grouiller F, Vercueil L, Krainik A, Segebarth C, Kahane P, David O. A comparative study of different artefact removal algorithms for EEG signals acquired during functional MRI. Neuroimage 2007;38(1):124–37.

Hamandi K, Salek-Haddadi A, Fish DR, Lemieux L. EEG/functional MRI in epilepsy: the queen square experience. J Clin Neurophysiol 2004;21:241–8.

Herrmann CS, Debener S. Simultaneous recording of EEG and BOLD responses: a historical perspective. Int J Psychophysiol 2008;67:161–8.

Hotelling H. Relations between two sets of variates. Biometrika 1936;28(3/4):321–77.

Koch I, Ruge H, Brass M, Rubin O, Meiran N, Prinza W. Equivalence of cognitive processes in brain imaging and behavioral studies: evidence from task switching. Neuroimage 2003;20:572–7.

Kok A. Effects of degradation of visual stimuli on components of the event-related potential (ERP) in go/nogo reaction tasks. Biol Psychol 1986;23:21–38.

Kruggel F, Wiggins CJ, Herrmann CS, von Cramon DY. Recording of the event-related potentials during functional MRI at 3.0 tesla field strength. Magn Reson Med 2000;44:277–82.

Leclercq Y, Balteau E, Dang-Vu T, Schabus M, Luxen A, Maquet P, et al. Rejection of pulse related artefact (PRA) from continuous electroencephalographic (EEG) time series recorded during functional magnetic resonance imaging (fMRI) using constraint independent component analysis (cICA). Neuroimage 2009;44(3):679–91.

Lemieux L, Salek-Haddadi A, Hoffmann A, Gotman J, Fish DR. EEG-correlated functional MRI: recent methodologic progress and current issues. Epilepsia 2002;43(Suppl. 1):64–8.

Mantini D, Perrucci MG, Cugini S, Ferretti A, Romani GL, Del Gratta C. Complete artifact removal for EEG recorded during continuous fMRI using independent component analysis. Neuroimage 2007;34(2):598–607.

Mordkoff JT, Gianaros PJ. Detecting the onset of the lateralized readiness potential: a comparison of available methods and procedures. Psychophysiology 2000;37:347–60.

Mulert C, Jäger L, Schmitt R, Bussfeld P, Pogarell O, Möller HJ, et al. Integration of fMRI and simultaneous EEG: towards a comprehensive understanding of localization and time-course of brain activity in target detection. Neuroimage 2004;22(1):83–94.

Nakamura W, Anami K, Mori T, Saitoh O, Cichocki A, Amari S. Removal of ballistocardiogram artifacts from simultaneously recorded EEG and fMRI data using independent component analysis. IEEE Trans Biomed Eng 2006;53(7):1294–308.

Niazy RK, Beckmann CF, Iannetti GD, Brady JM, Smith SM. Removal of fMRI environment artifacts from EEG data using optimal basis sets. Neuroimage 2005;28:720–37.

Pfefferbaum A, Ford JM, Weller BJ, Kopell BS. ERPs to response production and inhibition. Electroencephalogr Clin Neurophysiol 1985;60:423–34.

Picton TW, Bentin S, Berg P, Donchin E, Hillyard SA, Johnson R, et al. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 2000;37(2):127–52.

Picton TW, Lins OG, Scherg M. The recording and analysis of event-related potentials. In: Handbook of neuropsychology. Elsevier Science B.V; 1995.

Sadeh B, Zhdanov A, Podlipsky I, Hendler T, Yovel G. The validity of the face-selective ERP N170 component during simultaneous recording with functional MRI. Neuroimage 2008;42:778–86.

Sammer G, Blecker C, Gebhardt H, Kirsch P, Stark R, Vaitl D. Acquisition of typical EEG waveforms during fMRI: SSVEP, LRP, and frontal theta. Neuroimage 2005;24(4):1012–24.

Shibasaki H. Human brain mapping: hemodynamic response and electrophysiology. Clin Neurophysiol 2008;119(4):731–43.

Sijbers J, Van Audekerke J, Verhoye M, Van der Linden A, Van Dyck D. Reduction of ECG and gradient related artifacts in simultaneously recorded human EEG/MRI data. Magn Reson Imaging 2000;18(7):881–6.

Srivastava G, Crottaz-Herbette S, Lau KM, Glover GH, Menon V. ICA-based procedures for removing ballistocardiogram artifacts from EEG data acquired in the MRI scanner. Neuroimage 2005;24:50–60.

Stewart D, Love W. A general canonical correlation index. Psychol Bull 1968;70(3):160–3.

Strobel A, Debener S, Sorger B, Peters JC, Kranczioch C, Hoechstetter K, et al. Novelty and target processing during an auditory novelty oddball: a simultaneous event-related potential and functional magnetic resonance imaging study. Neuroimage 2008;40:869–83.

Vanderperren K, De Vos M, Ramautar JR, Novitskiy N, Vanrumste B, Stiers P, Van den Bergh B, Wagemans J, Lagae L, Sunaert S, Van Huffel S. Removal of BCG artifacts from EEG recordings inside the MR scanner: a comparison of methodological and validation-related aspects, in revision; 2009.

Vanderperren K, Ramautar J, Novitski N, De Vos M, Mennes M, Vanrumste B, et al. Ballistocardiogram artifacts in simultaneous EEG–fMRI acquisitions. International Journal of Bioelectromagnetism 2007;9(3):146–50.