



Investigating the variability of cardiac pulse artifacts across heartbeats in simultaneous EEG-fMRI recordings: A 7T study

João Jorge^{a,*}, Charlotte Bouloc^a, Lucie Bréchet^a, Christoph M. Michel^{b,c}, Rolf Gruetter^{a,d,e}

^a Laboratory for Functional and Metabolic Imaging, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

^b Functional Brain Mapping Lab, University of Geneva, Geneva, Switzerland

^c Biomedical Imaging Research Center (CIBM), Lausanne and Geneva, Switzerland

^d Department of Radiology, University of Lausanne, Lausanne, Switzerland

^e Department of Radiology, University of Geneva, Geneva, Switzerland

ARTICLE INFO

Keywords:

Simultaneous EEG-fMRI
Pulse artifact
Ballistocardiogram artifact
Average artifact subtraction
Ultra-high field

ABSTRACT

Electroencephalography (EEG) recordings performed in magnetic resonance imaging (MRI) scanners are affected by complex artifacts caused by heart function, often termed pulse artifacts (PAs). PAs can strongly compromise EEG data quality, and remain an open problem for EEG-fMRI. This study investigated the properties and mechanisms of PA variability across heartbeats, which has remained largely unaddressed to date, and evaluated its impact on PA correction approaches. Simultaneous EEG-fMRI was performed at 7T on healthy participants at rest or under visual stimulation, with concurrent recordings of breathing and cardiac activity. PA variability was found to contribute to EEG variance with more than $500 \mu\text{V}^2$ at 7T, which extrapolates to $92 \mu\text{V}^2$ at 3T. Clustering analyses revealed that PA variability not only is linked to variations in head position/orientation, as previously hypothesized, but also, and more importantly, to the respiratory cycle and to heart rate fluctuations. The latter mechanisms are associated to short-timescale variability (even across consecutive heartbeats), and their importance varied across EEG channels. In light of this PA variability, three PA correction techniques were compared: average artifact subtraction (AAS), optimal basis sets (OBS), and an approach based on K-means clustering. All methods allowed the recovery of visual evoked potentials from the EEG data; nonetheless, OBS and K-means tended to outperform AAS, likely due to the inability of the latter in modeling short-timescale variability. Altogether, these results offer novel insights into the dynamics and underlying mechanisms of the pulse artifact, with important consequences for its correction, relevant to most EEG-fMRI applications.

1. Introduction

Scalp electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) can be combined to achieve rich descriptions of the electrical and hemodynamic processes underlying brain function, non-invasively and with high spatiotemporal resolution (Mulert and Lemieux, 2010; Ullsperger and Debener, 2010; Jorge et al., 2014). Unfortunately, however, when acquired together, EEG and fMRI can interfere with each other through electromagnetic interactions that can strongly compromise data quality (Mullinger and Bowtell, 2011). Namely, the ongoing function of the heart in the static magnetic field of MRI scanners (B_0) produces a major contribution to EEG degradation, hereafter termed pulse artifact (PA) – also often named “ballistocardiogram artifact”. In the spectral domain, PA contributions can span a

frequency range up to 30 Hz, overlapping with an important part of the EEG spectrum. Even at moderate field strengths of 1.5T, PAs can often reach amplitudes of $50 \mu\text{V}$ (Debener et al., 2008; Neuner et al., 2013), easily surpassing most contributions from neuronal activity. At higher fields this becomes even more problematic, as PA amplitude increases approximately linearly with B_0 strength (Debener et al., 2008). For these reasons, PAs are a crucial nuisance for simultaneous EEG-fMRI, and have been given considerable attention (Ullsperger and Debener, 2010).

Several mechanisms have been shown to contribute to the generation of PAs, including bulk head motion in B_0 due to blood ejection from the heart, electric fields generated by ion separation in the moving blood (Hall effect), and local electrode movements due to scalp expansion (Yan et al., 2010; Mullinger et al., 2013). Unfortunately, being a consequence of heart function and vascular physiology in the MR environment, these

* Corresponding author. Laboratory for Functional and Metabolic Imaging, LIFMET-CIBM, Station 6, École Polytechnique Fédérale de Lausanne, CH-1015, Lausanne, Switzerland.

E-mail address: joao.jorge@epfl.ch (J. Jorge).

<https://doi.org/10.1016/j.neuroimage.2019.02.021>

Received 15 August 2018; Received in revised form 4 January 2019; Accepted 7 February 2019

Available online 8 February 2019

1053-8119/© 2019 Elsevier Inc. All rights reserved.

contributions cannot be avoided at their origin, and post-acquisition PA correction methods have become ubiquitous steps in EEG data processing.

The cyclic character of cardiac function motivated early on the development of average artifact subtraction (AAS), a technique which segments the EEG signal into epochs corresponding to the cardiac cycles, estimates the artifact in each epoch as an average over several nearby epochs, and subtracts it from the signal (Allen et al., 1998). Despite its simplicity, AAS has proved effective, and remains arguably the most widely used correction approach in EEG-fMRI, along with variants building upon the same principle (Goldman et al., 2000; Sijbers et al., 2000; Ellingson et al., 2004). AAS typically employs local windowed-averaging, where each epoch is corrected by an average over a number of its closest epochs in time. This is intended to account for

changes in the artifact shape and amplitude over time, presumably caused by drifts in head position and orientation (Allen et al., 1998; Moosmann et al., 2009). However, as reported by various groups (Allen et al., 1998; Debener et al., 2007; Grouiller et al., 2016) and shown in this work as well (Fig. 1), important PA residuals are often found after AAS correction, which appear to be caused by rapid variations in the artifact across epochs, and even across consecutive epochs. This type of beat-to-beat variability cannot be addressed by AAS, even with very small averaging windows (Fig. 1).

While considerable attention has been devoted to understanding the mechanisms that generate PAs, the sources of their inter-epoch variability remain, to our knowledge, largely unaddressed. This is, however, a crucial aspect for EEG analysis, given the impact that it can have on the performance of correction techniques such as the widely-used AAS

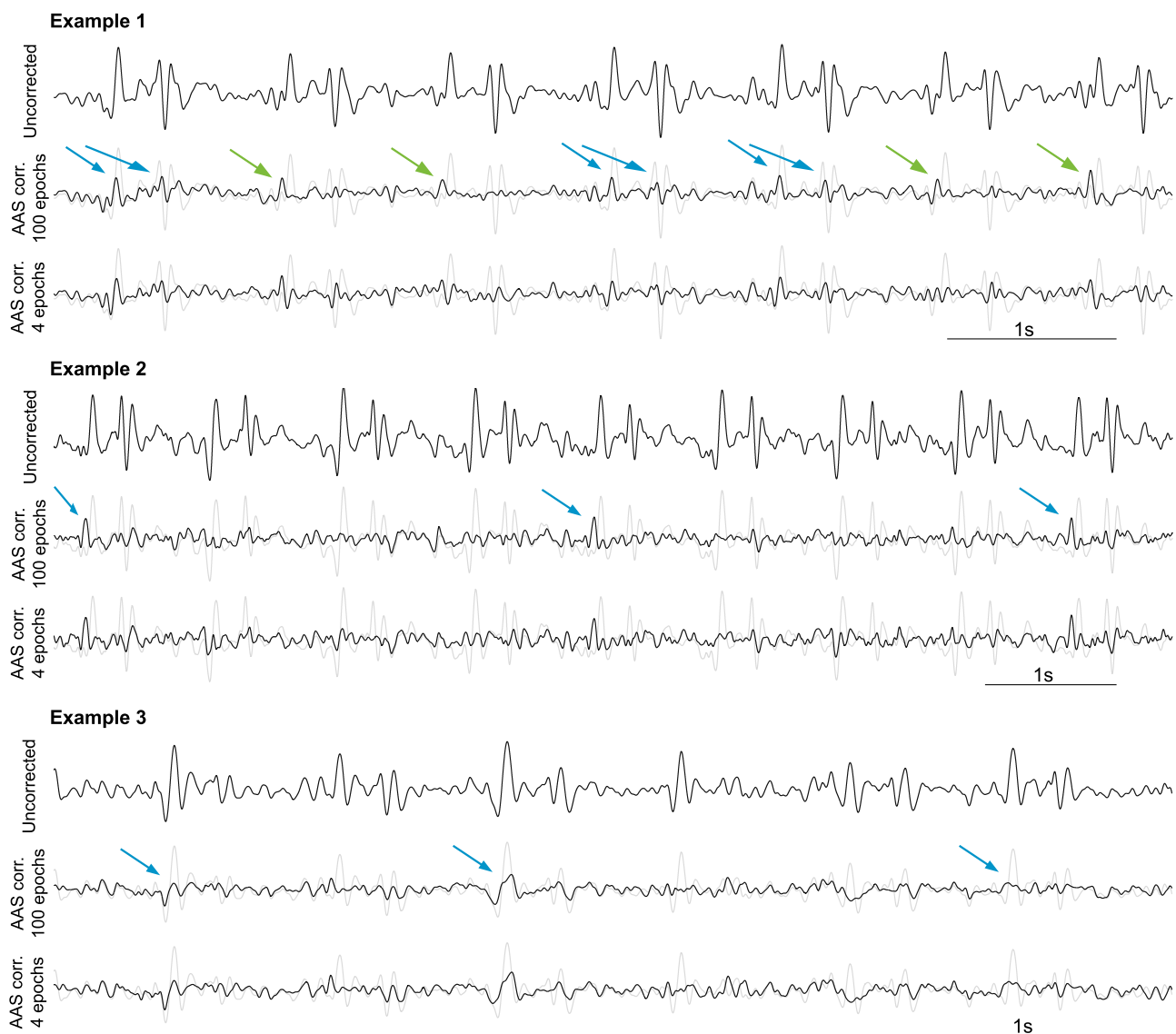


Fig. 1. Examples of short-timescale pulse artifact variability observed in EEG traces recorded in-scanner from three different healthy subjects. All examples correspond to channel Oz, referenced to the average reference. Top row: uncorrected signal; middle row: PA correction with AAS, using a large averaging window (100 epochs); bottom row: PA correction with AAS, short window (4 epochs); the uncorrected signal is also shown in gray, for comparison. Example 1: the PA appears to alternate between two distinct forms roughly every two epochs, resulting in clear residual peaks after AAS correction, appearing either coincident with the main artifact peaks (blue), or immediately before the first positive peak (green). Example 2: roughly every 4 epochs, a different PA shape is observed, lacking an initial undershoot, which results in residual peaks coincident with the (absent) undershoot, after correction. Example 3: the main PA peak is noticeably larger every 1–2 epochs, resulting in a coinciding deflection after AAS correction. Because they occur from one heartbeat to the next, these variations cannot be fully captured by AAS, even with a 4-epoch averaging window.

approach. A deeper understanding of PA variability could prove extremely valuable to better discern the extent of artifact residuals following EEG data correction, and to guide new improvements in correction techniques.

Considering human anatomy and physiology, and their interactions with magnetic fields, a number of sources could be hypothesized to contribute to PA variability. As often discussed in the literature, changes in head position and orientation will change the projection of the EEG wire loops along B_0 , and can also affect the dynamics of pulsatile head motion, scalp expansion and vessel orientation in B_0 , leading to changes in PA amplitude and shape (Yan et al., 2010). Far less often taken into consideration, respiration may also contribute to PA variability (Allen et al., 1998). From a physiology standpoint, it is known that, even at rest, systolic blood pressure exhibits a cyclic variation that is coupled to the breathing cycle, with inspiratory periods being associated to a lower pressure compared to expiratory periods – a mechanism known as “respiratory waves in arterial pressure” (Guyton and Hall, 2006). Such pressure variations, affecting the blood ejected by the heart, could likely have an impact in all three putative PA sources, inducing beat-to-beat variability on the relatively short timescale of the breathing cycle. Additionally, it is known that respiration affects the local magnetic susceptibility, leading to cyclic fluctuations in the B_0 field distribution, and can also induce small cyclic movements of the head (van Gelderen et al., 2007). While the former effect is likely too small to play a relevant role (B_0 variations of only a few parts-per-million (Jorge et al., 2018)), the induced cyclic changes in head position/orientation could potentially have a measurable influence on the ongoing PAs (Gretsch et al., 2018).

Another element that may play a relevant role in PA variability is heart rate itself. Fluctuations in heart rate can cause measurable fMRI signal changes in the brain, especially near regions with cerebrospinal fluid or blood vessels (Chang et al., 2009), suggesting an influence of heart rate on vascular properties. Moreover, in EEG-fMRI, PAs can extend for at least 600 ms following each QRS event (Debener et al., 2008; Mullinger et al., 2013), and potentially even affect subsequent artifact epochs (de Munck et al., 2013; Vincent et al., 2007), suggesting that the observed PA shape may change with variations in heart rate. Another (non-causal) link may arise from the fact that sympathetic and parasympathetic mechanisms that can regulate heart rate also act upon the cardiac contractile strength (and therefore the systolic pressure) simultaneously (Guyton and Hall, 2006).

Therefore, the aim of the present study was to investigate the properties and mechanisms of pulse artifact variability in concurrent EEG-fMRI, and to evaluate its impact on various PA correction approaches. Considering the mechanisms discussed above, we hypothesized that respiration, heart rate, and head motion may be linked to PA variability. Simultaneous EEG-fMRI was performed on healthy human participants at rest or undergoing visual stimulation, with concurrent recordings of breathing and cardiac activity. The acquisitions were performed at 7T, where the PA has a larger signal-to-noise ratio than in more conventional studies. The contributions of PA variability to the recorded EEG signals were first assessed by comparison to additional datasets that do not contain PAs, including: off-scanner human EEG, expected to contain only real EEG activity and EEG-specific artifacts (e.g. eye-blinks, muscle activity), and phantom EEG-fMRI, expected to contain only MR environment-related artifacts (e.g. He coldheads, which can be important at higher fields (Mullinger et al., 2008), and gradient artifact residuals). In a subsequent part, the human EEG-fMRI resting-state data were analyzed with clustering techniques, to determine potential relationships between the variability across different PA epochs and the ongoing respiratory amplitude, cardiac period, head position and orientation. Potential confounds, such as gradient artifact residuals or ongoing EEG activity, were also considered. PA correction was then studied on the visual stimulation data with three distinct approaches, each with specific strengths and limitations in accounting for PA variability: AAS, optimal basis sets (OBS), and a K-means clustering-based method.

2. Materials and methods

This study was approved by the local ethics committee (CER-VD), and involved the participation of 12 healthy volunteers (23 ± 3 years old, 6 male/6 female), who provided written informed consent. All collected data are available upon request, in coded form to preserve anonymity.

2.1. Data acquisition

2.1.1. Human EEG-fMRI

Simultaneous EEG-fMRI acquisitions were performed on an actively-shielded Magnetom 7T/68-cm head scanner (Siemens, Erlangen, Germany), equipped with an 8-channel head RF array (Rapid Biomedical, Rimpf, Germany). Whole-brain functional data were acquired using a 2D simultaneous multi-slice (SMS) gradient-echo EPI sequence, with $TR = 1$ s, 2.2-mm isotropic resolution, 69 sagittal slices with $3 \times$ SMS acceleration, and $2 \times$ in-plane GRAPPA acceleration.

EEG data were recorded using a previously-described compact setup (Jorge et al., 2015b), comprising two 32-channel BrainAmp MR Plus amplifiers (Brain Products, Munich, Germany) and a BrainCap MR model (EasyCap, Herrsching, Germany), connected by 12-cm bundled cables. The 64 cap electrodes were arranged in an extended 10–20 system (referenced to channel FCz), with one electrode dedicated to ECG recording and placed on the back of the subject. Abralyte gel (EasyCap) was used to reduce electrode impedances. After bandpass filtering (0.016–250 Hz) and digitization (0.5- μ V resolution, 5-kHz sampling), the EEG signals were optically transmitted to the control room. Scanner triggers marking the onset of each fMRI volume were also recorded.

Respiratory traces were recorded in synchrony with the EEG-fMRI acquisitions, at 50 Hz, using a respiratory belt from the physiological monitoring unit of the MRI system (Siemens).

Each subject underwent two functional acquisitions while lying in the MRI scanner: an 8-min eyes-open resting-state run, and a 5-min visual stimulation run. Visual stimuli were delivered with an MR-compatible LCD screen (Cambridge Research Systems, Rochester, UK) at 120 frames/s, viewed through a mirror placed on the RF coil. For the resting-state run, the subjects were asked to lie still and focus on a small red cross presented on the screen, to minimize head and eye movements. The visual stimulation run comprised 10 blocks of 10s stimulation followed by 20s fixation periods; the stimuli consisted of grey-scale checkerboards (50% contrast) inverting at 7.5 Hz, with the total luminance kept equal to fixation periods. The timings of each checkerboard reversal were recorded in synchrony with the EEG. A central red cross was also shown, with slight changes in color occurring at random instants; the subjects were instructed to focus on the cross throughout the run and report color changes via a button press, to ensure their attention.

2.1.2. Off-scanner human EEG

These resting-state EEG recordings were performed outside the scanner room, in a sitting position fixating on a cross, for 4 of the 12 subjects.

2.1.3. Phantom EEG-fMRI

This consisted of an 8-min EEG-fMRI acquisition using a head-shaped agar phantom covered with a layer of electrolyte gel, with similar EEG and fMRI sequence parameters.

2.2. Data analysis

All data processing was performed offline in Matlab (MathWorks, Natick MA, USA), using routines developed in-house. Initially, both resting-state and visual stimulation datasets were prepared via a similar pre-processing stage. Subsequently, the former dataset underwent PA variability analysis, while the latter was utilized for PA correction analysis.

2.2.1. Pre-processing

EEG gradient artifacts were corrected slice-by-slice via AAS (Allen et al., 2000) as described in previous work (Jorge et al., 2015a), using a window size of 100 slice epochs, with jittered steps of 8–10 slices separating the selected epochs. After gradient artifact correction, bad channels were identified by visual inspection (1–4 per subject) and replaced by interpolation using 3–6 neighbors. The data then underwent temporal bandpass filtering at 0.5–35 Hz, to remove slow drifts and higher-frequency spurious contributions (including gradient artifact residuals, muscle activity), while preserving the spectral domain of the PA. Subsequently, cardiac triggers were estimated from the ECG channel (Niazy et al., 2005), and fine-tuned with a correlation-maximization approach, using a combination of EEG channels where PAs were most prominent (see Supp. Fig. 1 for further details). For each run, a fixed offset was then added to all triggers, to center the epoch window so as to fully contain the dominant oscillations of the PA (Supp. Fig. 1). The same EEG channels were then used to mark PA outliers, i.e. epochs that were notably distinct from the overall sample in each dataset, likely coinciding with spontaneous motion occurrences. For this purpose, the Euclidean distance between every pair of EEG PA epochs was calculated (i.e. considering each epoch timecourse as a vector, and computing the Euclidean norm of the difference between vectors). Epochs with abnormally large average distances to the others were marked as outliers (Supp. Fig. 1). Relatively strict thresholds were applied, resulting in 7–14% marked outliers per run, to ensure minimal confounds for the subsequent variability analysis. No channel re-referencing was performed at this point.

2.2.2. Pulse artifact variability assessment

For an initial evaluation of the magnitude and distribution of PA variability, after pre-processing, three EEG data groups were considered, differing in their artifact contributions: (i) the 12-subject in-scanner resting-state data (SubIn), expected to contain PA contributions, real

EEG activity, EEG-specific artifacts and MR environment-related artifacts; (ii) the 4-subject off-scanner data (SubOut), containing only EEG activity and EEG-specific artifacts; (iii) the phantom in-scanner data (PhanIn), containing only MR environment-related artifacts. For each subject and EEG channel, the mean PA power was estimated by taking the average PA across all epochs (excluding outliers), and then computing its power across the PA window. PA variability was quantified by computing the variance across epochs for each timepoint of the PA window, and then averaging across the window. For the PhanIn case, which lacked cardiac activity, an analogous approach was simulated by copying the trigger timings from one of the human subjects.

2.2.3. Pulse artifact variability mechanisms – variable-based clustering

This analysis stage focused on investigating possible mechanisms linked to PA variability, following the hypotheses stated in section 1. First, each PA epoch was attributed a set of variables describing its associated: time of occurrence, respiratory amplitude, cardiac period, head position and orientation, and potential confounds (Fig. 2). These putative PA-linked variables are hereafter termed “PA variables”. The time of PA occurrence was simply the PA trigger time; this was not hypothesized to be directly linked to PA variability, but was included in the analysis as it effectively corresponds to the criterion used by AAS for PA template estimation. The respiratory amplitude associated to each PA epoch was extracted from the respiratory signal; this signal was normalized to an amplitude range between -1 (expiratory peak) and $+1$ (inspiratory peak), as proposed by Glover et al. (2000), and then linearly interpolated to the PA trigger timings. The cardiac period was obtained directly from the time differences between PA triggers. The head position and orientation were obtained from the fMRI data; rigid-body motion estimation (Jenkinson et al., 2002) was performed relative to the middle volume of each series, and subsequently re-referenced to the first volume. Linear interpolation was again applied to estimate values for the 6 motion parameters at the PA timings. The confounds included in the analysis

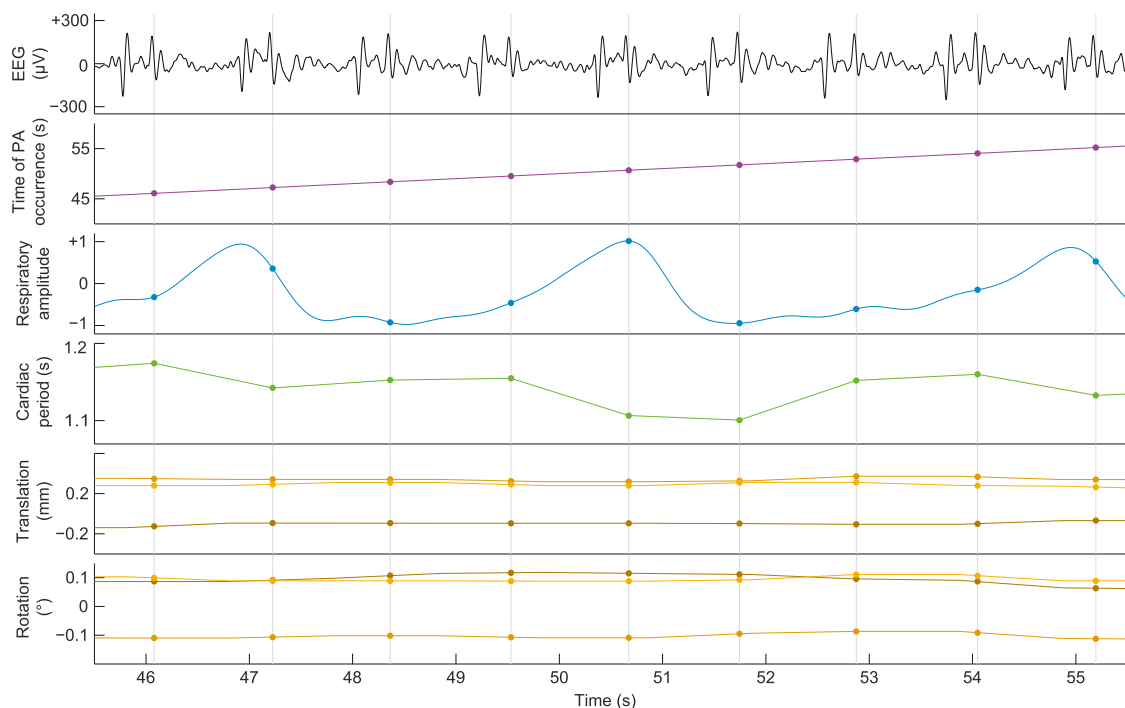


Fig. 2. Illustration of the PA labeling process employed in this work to investigate mechanisms linked to PA variability. Each segmented PA epoch (1st row) was marked with a temporal trigger at the center of the epoch window (gray vertical lines), and then attributed a set of variables describing its: time of PA occurrence, in seconds, relative to the start of the recording (2nd row); respiratory amplitude, normalized to a range of -1 (peak expiration) to $+1$ (peak inspiration) (3rd row); cardiac period, in seconds (4th row); head position/orientation, comprising three translations (5th row) and three rotations (6th row), relative to the first fMRI volume. Additional variables were obtained for a number of potential confounds, including the timing relative to the fMRI acquisition, both at volume and slice scale, as well as the ongoing EEG alpha activity of each subject.

comprised the phase of occurrence within the fMRI volume acquisition sequence (set to 0 at the first volume trigger, and then increasing linearly to multiples of 2π at each following trigger), and analogously for the slice acquisition sequence; these indices were intended to control for potential variability contributions from gradient artifact residuals, and not the PA itself. Other confounds were added to control for the influence of ongoing EEG activity, namely alpha activity; alpha wave timecourses were extracted from each EEG run by ICA, selected as components with a clear isolated peak in the alpha band, and occipito/parietal scalp distributions; up to 3 reliable alpha components were selected from each run, depending on the subject; from these signals, alpha power and phase timecourses were obtained by windowed Fourier analysis, and used as indices of alpha activity for each PA epoch.

Having associated the PA epochs to the above-described variables, each of those variables was then considered independently from the others, and underwent K-means clustering (MacKay, 2003) (see Supp. Fig. 2 for a schematic illustration). The number of clusters K was varied from 1 to 10; cluster centroids were initialized at random samples, and the clustering was repeated for 7 trials with different initializations, after which the solution with the lowest value of the K-means cost function was chosen. The squared Euclidean distance was adopted as distance metric for clustering the time of PA occurrence, respiratory amplitude, cardiac period, head motion (treated as a 6-dimensional vector) or alpha power; to cluster the fMRI volume phase, slice phase, or the alpha phase, the cosine distance was employed instead (defined as $1 - \cos(\text{phase difference})$). An additional series of 1000 clustering runs per subject was performed using a random variable, and another clustering run was performed directly with the EEG PA signals. Subsequently, for each clustering solution of each variable and K value, the EEG PA epochs of each channel (excluding outliers) were grouped following that same clustering scheme, and a measure of intra-cluster variability was obtained as the mean of Euclidean distances between each PA epoch and the mean PA (i.e. centroid) of their cluster (Supp. Fig. 2). Intuitively, it is expected that if a particular variable is linked to PA variability, then the grouping of PA epochs based on a clustering of that variable should yield significantly more homogeneous PA clusters (lower intra-cluster variability), than when they are grouped at random. The use of this clustering approach avoids the need to make a priori assumptions as to which particular features of PA variability (e.g. peak amplitudes, shape) might be linked to any putative variables, and is well suited to treat multidimensional variables like head motion. The testing of different values of K allowed to evaluate how finely the tested variables could potentially discriminate PA variability.

2.2.4. Pulse artifact variability mechanisms – PA-based clustering

This last stage of the variability analysis employed a somewhat reciprocal approach to the previous stage: for a given EEG channel, the PA epochs (excluding outliers) were clustered with K-means directly based on their amplitudes (squared Euclidean distance metric), with K set to 1, 2, 4 or 8. The clustering results were then used to group the PA variables accordingly (see Supp. Fig. 3 for a schematic illustration). This analysis sought to obtain a more detailed picture of the links between PA variability, including PA shape and amplitude, and each of the variables, as well as inter-dependencies between variables.

2.2.5. Pulse artifact correction

Here, three distinct PA correction approaches were compared, each with specific characteristics and adaptiveness regarding PA variability: AAS, OBS, and a technique based on K-means clustering (hereafter referred to as “K-means”). The AAS approach was similar to that of Allen et al. (1998), implemented as a subtraction of the mean PA over the N_a closest neighbors ($N_a/2$ from preceding epochs and $N_a/2$ from the following, excluding outliers). Particular to this work, the PA window size was made adaptive, and set as the minimum cardiac period of the N_a epochs used in each particular template. The tests spanned a range of N_a values from 2 to 100 epochs. OBS was based on the work of Niazy et al.

(2005): for a given EEG channel, all PA epochs were segmented (excluding outliers, and using the mean cardiac period as window size), and underwent principal component analysis across the epochs. Each epoch (including outliers) was then subtracted of a linear combination of the N_o most important principal components. N_o was varied from 1 (simply the grand-average across all epochs) to 50. The K-means approach was motivated by the results of the variability analyses, as well as previous works on PA correction (Goncalves et al., 2007). Here, for a given EEG channel, the PA epochs were segmented (as in OBS), and clustered with K-means (squared Euclidean distance, K clusters); the cluster-average PAs (i.e. centroids) were then subtracted from each of the epochs belonging to the respective cluster; outliers were assigned to their closest cluster and corrected as well. K was varied from 1 (grand-average across epochs) to 50. The case of $K = 1$ was also given attention for corresponding to a completely non-adaptive subtraction approach.

The performance of the correction approaches was assessed on the visual stimulation dataset. The employed stimulus frequency (7.5 Hz) is known to elicit so-called steady-state visual evoked potentials (ssVEPs) – stationary oscillatory responses with a fundamental frequency matching the stimulus frequency, which can be used as a reference to assess functional sensitivity. For each subject, the EEG data were first re-referenced to the average reference, and the channel displaying the strongest ssVEPs was selected for analysis (Oz in most cases). This channel was then PA-corrected with each of the different approaches and parameter values.

Correction performance was assessed based on: (i) the impact on EEG signal root-mean-square (RMS), serving as a general measure of signal variability reduction; and (ii) the quality of the retrieved ssVEPs, namely their average amplitude, serving as a measure of true EEG signal preservation, and their inter-trial variability, as a more specific measure of artifact reduction (or conversely, of accuracy in ssVEP estimation).

To pursue more detailed insights into the impact of PA correction, the RMS estimates were performed not across the full timecourse, but on two distinct temporal subsets: (i) the segments at the periphery of each PA epoch, comprising the outermost 2% of each PA window, and (ii) segments at the center of each PA epoch, comprising the innermost 98%. Given that the largest PA fluctuations occur at the center segments, their RMS is expected to be more strongly reduced by correction than in periphery segments (where artifact fluctuations are much weaker, or even absent) – and ideally, the two cases should achieve a similar RMS. Additionally, the average RMS in the pre-processed, 4-subject off-scanner resting-state data was also estimated, to serve as a PA-free reference.

To assess response amplitude and variability, the ssVEP trials (i.e. each checkerboard reversal) were segmented, and the mean ssVEP amplitude was estimated by computing the mean ssVEP across trials and then obtaining its RMS across the trial window. The ssVEP inter-trial variability was quantified as the standard deviation across trials for every timepoint of the window, subsequently averaged across timepoints. An effective correction approach is expected to preserve the RMS of the mean ssVEP while decreasing its inter-trial variability, which should be dominated by artifact contributions. For both the mean ssVEP and the inter-trial variability, an analogous estimate was made by shifting the ssVEP triggers by 15s to the fixation periods, to serve as a baseline reference; the triggers were also copied to the off-scanner data, to obtain a baseline PA-free (and also ssVEP-free) reference of the same measures.

3. Results

Across the 12 participants of this study, an average of 461 ± 1 PA epochs (range: 409–529) were recorded per subject during the 8-min EEG-fMRI resting-state run, with an average cardiac period of 1.06 ± 0.03 s (0.92–1.18 s).

3.1. Pulse artifact variability assessment

This initial assessment relied on a comparison of the mean PA power

and PA variability between different recording conditions, including human and phantom data, in-scanner and off-scanner. Overall, the subject in-scanner data (SubIn) showed by far the strongest mean PA power, surpassing the subject off-scanner data (SubOut) by roughly four orders of magnitude, and the phantom in-scanner data (PhanIn) by five (Fig. 3). The fairly low power values in SubOut and PhanIn (below $1 \mu\text{V}^2$) are consistent with the absence of measurable PAs in these recordings. The variance across PA epochs was strongest for SubIn, although the difference relative to SubOut was less accentuated (one order of magnitude). Across the scalp, in the SubIn case, the mean PA power was strongest in occipital and lateral regions; the PA variance distribution was similar, with a slight relative increase for frontal electrodes. In SubOut data, both the PA power and variance estimates were strongest at frontal electrodes (likely related to eye-blink contributions), followed by occipital regions. In the PhanIn case, both the PA power and variance lacked a clear distribution, apart from a “checkerboard” pattern that matched the division of channels between the two EEG amplifiers. The channel-averaged estimates for PA power and variance across PA epochs essentially reflected the above-described variations across conditions (Fig. 3, right). Assuming that the signal contributions from PAs, real EEG activity, EEG-specific artifacts and MR-related artifacts are uncorrelated to each other, the variance across PA epochs seen in the SubIn case should then match the sum of the variances from SubOut and PhanIn, plus any inter-epoch variance from the actual PAs. Given the measured variances, this PA-specific variability is indeed highly likely to be present ($p < 0.05$, two-sample t -test), and substantial, on the order of $529 \mu\text{V}^2$.

3.2. Pulse artifact variability mechanisms

A preliminary correlation analysis revealed a number of interdependencies between some of the PA variables (Supp. Fig. 4). Considering their (absolute) correlations, on average across subjects, the 6 rigid-body head motion indices were highly correlated to each other

($r = 0.41\text{--}0.84$), and to the time of PA occurrence ($r = 0.56\text{--}0.79$). Neither the respiratory amplitude nor the cardiac period were substantially correlated to the other variables ($r \leq 0.23$), or to each other ($r = 0.16$). Nonetheless, the correlation between the two was negative in 8 of the 12 subjects (significantly in 6, $p < 0.05$) – reflecting a (healthy) physiological coupling known as respiratory sinus arrhythmia (Hirsch and Bishop, 1981).

The presence of links between PA variability and the PA variables was investigated with clustering analyses. The first analysis consisted of clustering each of the PA variables independently, and grouping the respective PA epochs in the same scheme, to then determine their intra-cluster variability (section 2.2.3, Supp. Fig. 2). For all tested variables, the resulting intra-cluster PA variability was generally found to decrease with the number of clusters, K (Fig. 4, top). This decrease was strongest when using the respiratory amplitude or cardiac period as clustering variables, followed by head motion and the time of PA occurrence. For these four variables, the resulting intra-cluster variability was well below the 95%-range of outcomes found with random clustering, in every subject. In contrast, the variability obtained with the confound variables was generally within the 95%-range of random clustering, in all subjects. On average across channels and subjects, the intra-cluster variability reduction relative to random clustering became stronger with increasing K , and stabilized after $K \approx 10$, except for the direct PA signal-based clustering (Fig. 4, bottom-right). At $K = 10$, the respiratory amplitude and cardiac period produced the largest variability reductions, followed by head motion and time of PA occurrence (Table 1); all these reductions were statistically significant ($p < 0.01$). The confound variables showed only marginal, non-significant reductions. Finally, the PA signal-based clustering achieved markedly stronger reductions than any of the tested variables.

As observed for individual channel curves (Fig. 4), the intra-cluster PA variability reductions achieved by each variable varied across the scalp (Fig. 5). The respiratory amplitude and cardiac period both had

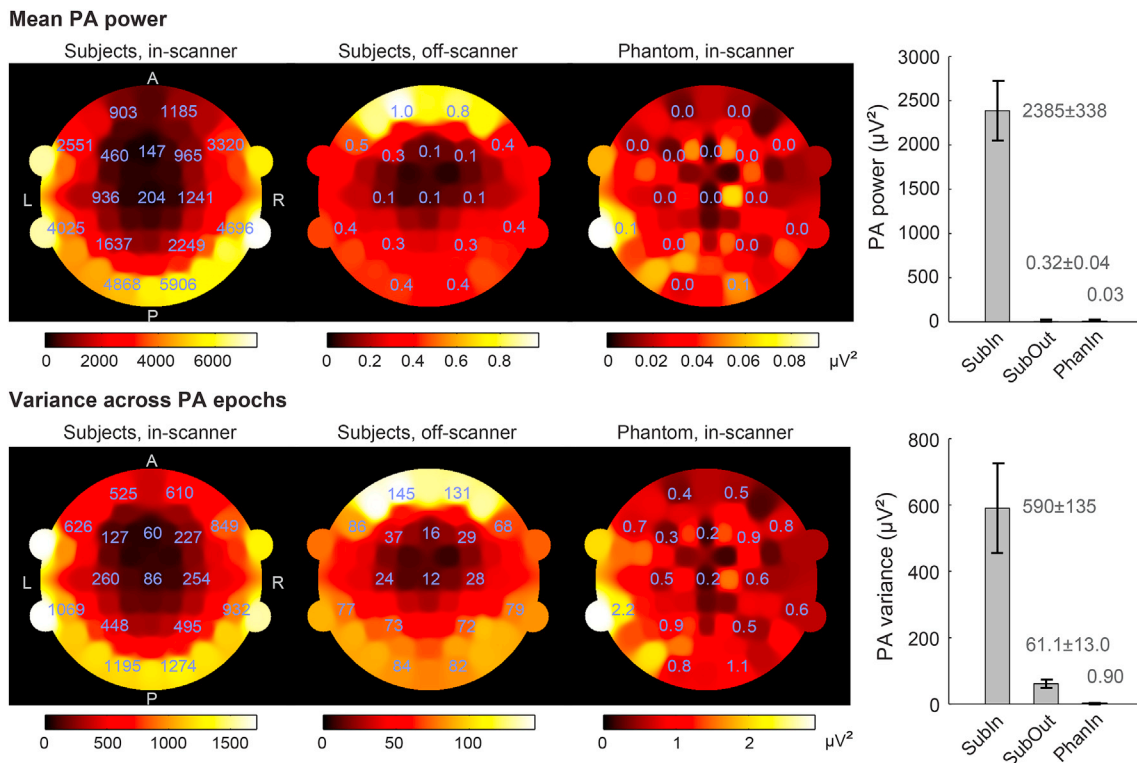


Fig. 3. Pulse artifact variability in three distinct conditions: 12 human subjects in-scanner (SubIn in the bar graphs), 4 human subjects off-scanner (SubOut), and phantom in-scanner (PhanIn). Left: group-average scalp distributions of the power of the mean PA and the variance across PA epochs. Each color map was scaled independently to account for the markedly different amplitude ranges; numerical values are also shown for some of the channels, superimposed on the maps, for a more straight-forward comparison. Right: channel and subject-averaged values for the same measures; the error bars represent the standard error across subjects.

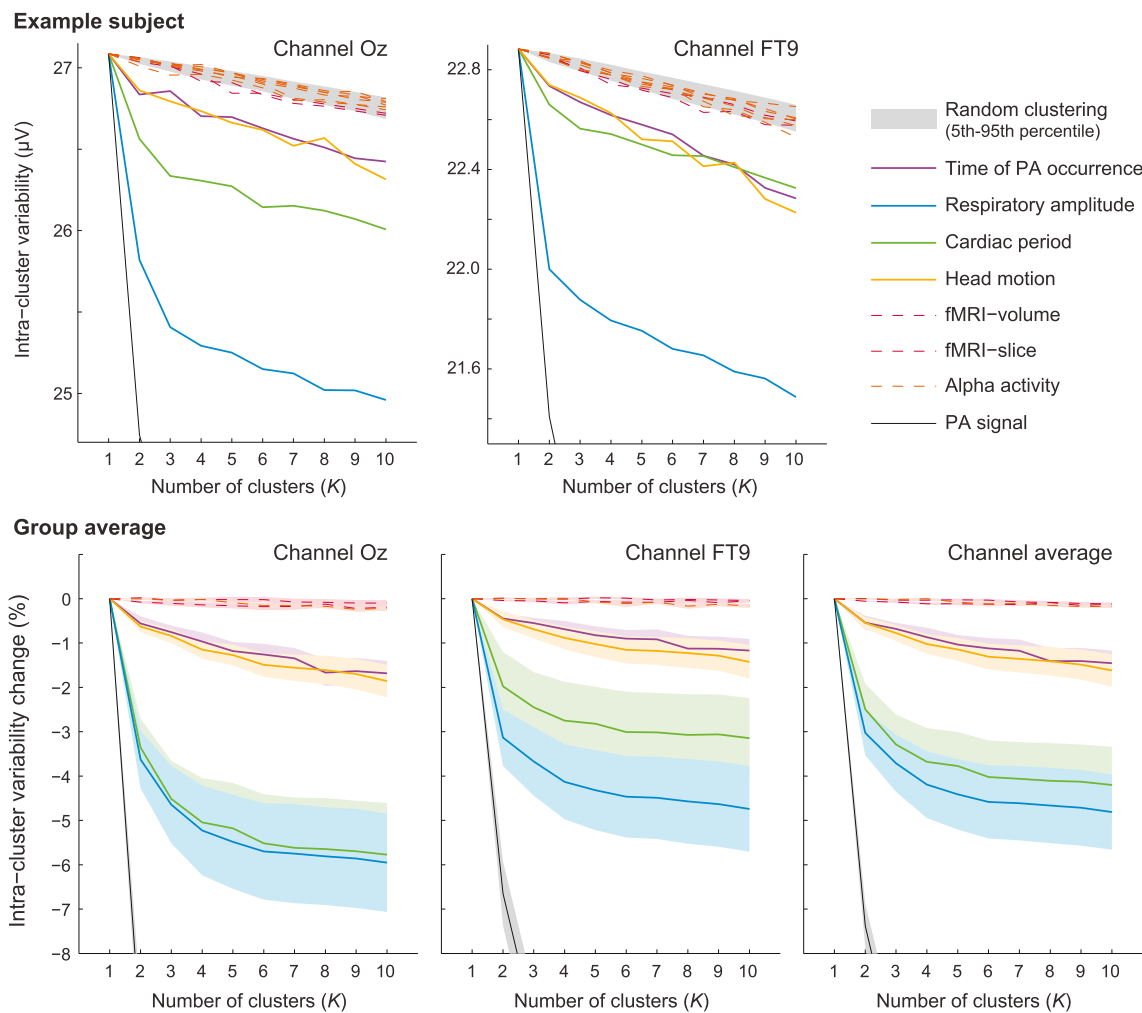


Fig. 4. Intra-cluster variability of pulse artifacts when clustered according to each of their attributed variables (time of PA occurrence, respiratory amplitude, cardiac period, head position/orientation, or potential confounds). The variability curves obtained when clustering the PA epochs directly based on their EEG signals are also shown in black for comparison (due to their much lower variability, parts of these curves are excluded from view). Results are shown for channels Oz and FT9, as well as for all channels averaged (bottom-right). In the example subject (top panels), the grey area represents the 95%-range of outcomes obtained with random clustering for that subject (5th to 95th percentile of the 1000 trials). In the subject group average (bottom panels), the curves were normalized as the change in variability relative to the mean outcome of random clustering, prior to averaging across channels and subjects; the colored margins their standard error.

Table 1
Intra-cluster PA variability decrease with variable-based clustering ($K = 10$).

PA variable	Intra-cluster PA variability decrease (%)
Respiratory amplitude	4.8 ± 0.8
Cardiac period	4.2 ± 0.9
Head motion	1.6 ± 0.4
Time of PA occurrence	1.5 ± 0.3
fMRI volume-phase	0.12 ± 0.03
fMRI slice-phase	0.12 ± 0.02
Alpha activity	0.18 ± 0.04
PA signal ^a	18.9 ± 0.9

Notes. The values shown correspond to averages across channels and subjects, and the error margins to the standard error across subjects.

^a Clustering performed directly based on the EEG PA signals, instead of one of the variables.

their strongest reductions in posterior areas; additionally for the respiratory amplitude, but not the cardiac period, these reductions also extended to more lateral electrodes. The distributions for the time of PA occurrence and head motion were fairly similar to each other, with the peak reductions occurring in central electrodes, followed by posterior

regions. The distributions for confounds were generally less well defined, with a tendency to peak at central electrodes closest to the reference (FCz). The reductions achieved with PA signal clustering were also roughly homogeneous, with a slight increase for regions radially farther from the reference, particularly in occipital and frontal electrodes (Fig. 5).

In the following stage of the analysis, the PA epochs were clustered directly based on their EEG signal, for a given channel, and this clustering scheme was then used to group their attributed variables (section 2.2.4, Supp. Fig. 3). The results were analyzed on a subject-by-subject basis. Considering the previously obtained scalp distributions (Fig. 5), the tests were focused on channel Oz (more occipital) and FT9 (more lateral). For Oz, clustering the PA epochs with $K = 2$ yielded clusters that were generally very well differentiated in terms of respiratory amplitude, or cardiac period, or both, in all subjects (e.g. Fig. 6, top). The same was found with $K = 4$; notably, some clusters that showed similar distributions in respiratory amplitude were well differentiated in cardiac period (e.g. Fig. 6 bottom, clusters 1,4), and others vice-versa (clusters 2,4), while other clusters differed in both variables simultaneously; this was observed both in subjects with a clear anti-correlation between respiratory amplitude and cardiac period (Supp. Fig. 5) as well as without (Supp.

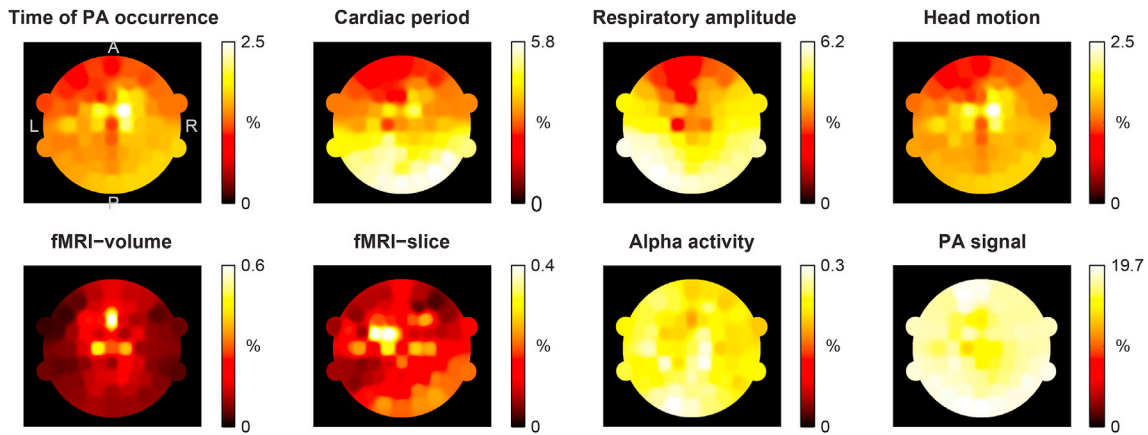


Fig. 5. Group-averaged scalp distributions of the intra-cluster variability of pulse artifacts when clustered according to each of their attributed variables (time of PA occurrence, respiratory amplitude, cardiac period, head position/orientation, or potential confounds), for $K = 10$. The results are presented as the decrease in variability relative to random clustering, averaged across subjects for each EEG channel. The variability reduction obtained when clustering the PA epochs directly based on their EEG signals is also shown for comparison. Each color map is scaled independently.

Fig. 6); 4 subjects also showed clusters that could only be differentiated from the others when considering their head position and/or orientation distributions.

With $K = 8$, the presence of motion-differentiated clusters was now apparent in 11 subjects (e.g. **Fig. 7** top, clusters 3,7); some clusters were also found, in 4 subjects, that did not appear well differentiated from the others in any variable (e.g. **Fig. 7** bottom, clusters 4,8). The results obtained at channel FT9 were generally similar to those of Oz, with subtle differences: respiratory amplitude was more prevalent in cluster differentiation than cardiac period; motion-differentiated clusters were slightly more frequent than with Oz (typically 1 more subject, for the same K). The confound variables were also inspected throughout the analysis, but could not explain any cluster assignments in any of the tests performed.

An effort to understand the links between the dominant PA variables and the shape of the cluster-average PAs (centroids) was made by inspecting specific cases of clusters differing in respiratory amplitude but not in cardiac period, and vice-versa (e.g. **Fig. 6** bottom). The PA centroids differed most frequently in the amplitude of specific peaks, and/or in temporal scale, i.e. a dilation or contraction of the PA shape (**Supp. Fig. 7**). However, across the full dataset, no systematic associations could be solidly established between each of these two specific effects and each of the two dominant variables.

3.3. Pulse artifact correction

The three PA correction approaches under comparison exhibited important performance differences across the range of tested parameters. On average across subjects, before PA correction, the (in-scanner) EEG signals had higher RMS values for PA center segments than for PA periphery segments, and both were considerably stronger than in off-scanner data (**Table 2**). The non-adaptive PA correction (i.e. K-means with $K = 1$) mainly reduced the RMS of center segments, bringing it closer to that of periphery segments – although still higher than off-scanner. AAS yielded similar values to the non-adaptive approach when using N_a larger than ~ 8 epochs, and was relatively insensitive to N_a ; below 8, both segments showed sharp increases in RMS (**Fig. 8a**). With OBS, the RMS of both segment types was steadily reduced when increasing the number of components N_o . Despite starting above the PA periphery segments, the center RMS decreased faster, becoming smaller than the periphery for $N_o > 4$. K-means also showed steady decreases in RMS when increasing the number of clusters K , yet less accentuated than for OBS; the PA center RMS remained stronger than the periphery, but the two became progressively closer with increasing K .

Regarding visual responses, before PA correction, the trial-average

ssVEP RMS showed fairly similar values for either the stimulation or fixation periods, which were considerably higher than off-scanner (**Table 2**). Non-adaptive correction reduced both values, yet the resulting ssVEP RMS for the stimulation periods was substantially stronger than for fixation, consistent with the expected activity. The trends found for each correction technique were generally consistent with those of signal RMS (**Fig. 8b**). The curves for fixation were always lower than those of stimulation, but also remained substantially higher than in off-scanner fixation (except with OBS at very high N_o). Of note, the K-means curve in stimulation was fairly stable for $K < 8$, only decreasing thereafter.

The ssVEP inter-trial variability also showed similar values for either the stimulation or fixation periods, both approximately $10\times$ higher than off-scanner (**Table 2**). Non-adaptive PA correction substantially reduced both values, yet remaining $4\times$ higher than off-scanner. For all correction techniques, the stimulation and fixation curves were generally similar to each other, with similar trends to the ssVEP RMS (**Fig. 8c**). K-means showed slightly more accentuated reductions in variability for $K < 10$, followed by milder decreases. In fact, a closer inspection of the individual subject curves for K-means revealed that both the PA center RMS and the ssVEP inter-trial variability often showed a sharper drop at smaller K followed by a smoother decrease thereafter; the “transition” value of K varied across subjects (roughly around 9–16 for RMS, 12–18 for ssVEP variability), which diluted this effect on the group-average curves.

Based on direct inspection of the ssVEPs recovered for each subject, the optimal PA correction parameters were found to be, across subjects, $N_a = 30\text{--}36$ for AAS, $N_o = 5\text{--}8$ for OBS, and $K = 10\text{--}17$ for K-means. These yielded clear ssVEP curves in occipital areas for most subjects, which were generally indiscernible without PA correction (**Fig. 9, Supp. Fig. 8**). Based on the performance measures obtained with these optimized parameters (summarized in **Table 2**), a number of observations can be drawn: (i) the three adaptive methods were generally superior to non-adaptive correction (i.e. lower EEG RMS, similar trial-average ssVEP RMS, lower ssVEP inter-trial variability), with AAS being the closest to the non-adaptive case; (ii) OBS achieved the strongest EEG denoising (lowest ssVEP variability), but appeared to reduce visual responses to a greater extent as well (lowest ssVEP RMS); (iii) K-means yielded an intermediate outcome relative to OBS and AAS. Overall, these differences across correction methods were fairly moderate (**Table 2**), and the outcomes remained substantially different from off-scanner data (e.g. both the EEG RMS and ssVEP variability measures after OBS remained approximately $3\times$ higher than off-scanner). Nevertheless, the described differences in performance were generally directly visible on the recovered ssVEPs of individual subjects (**Fig. 9**).

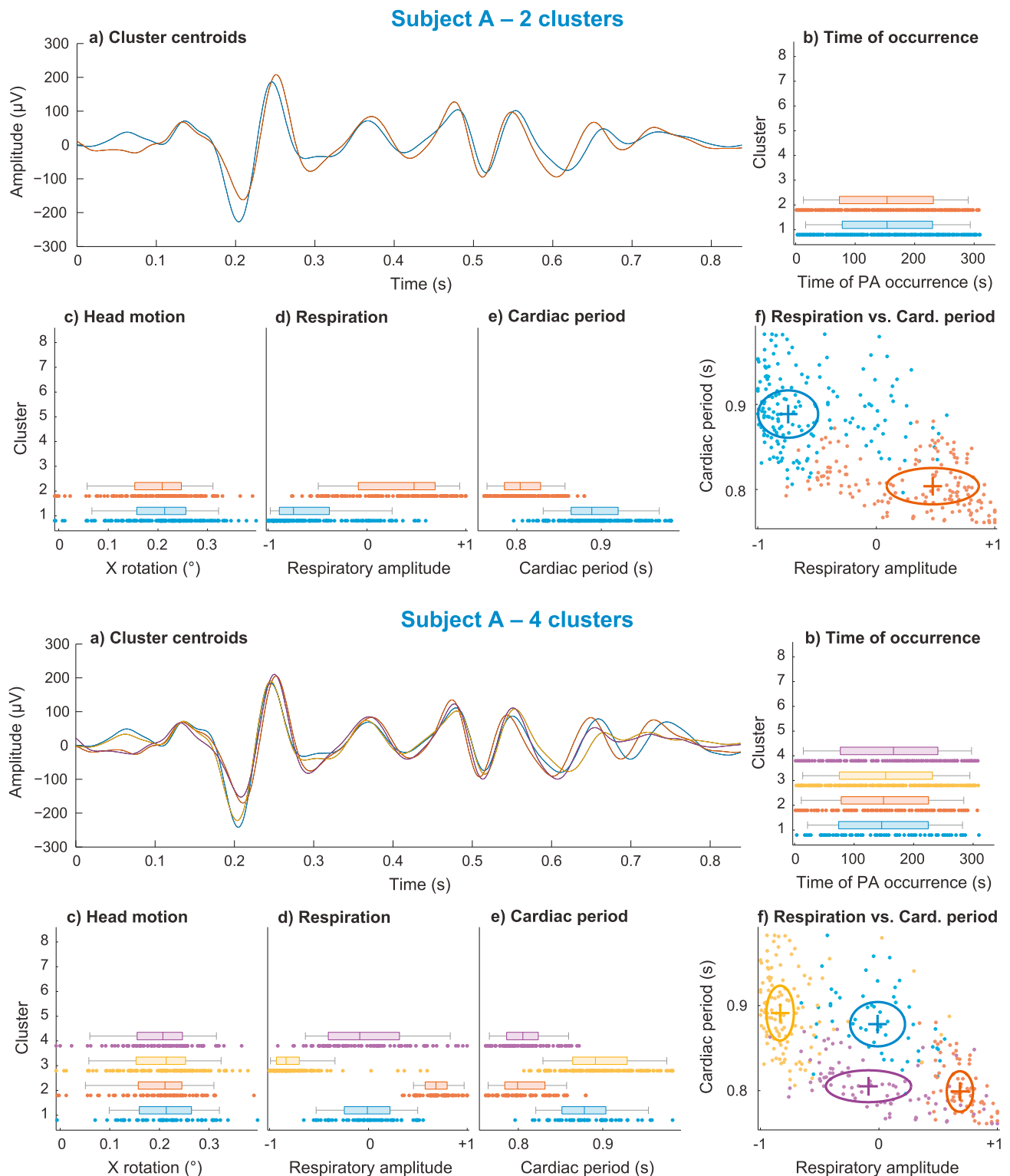


Fig. 6. Cluster-wise distribution of the PA-attributed variables (time of PA occurrence, respiratory amplitude, cardiac period and head motion) when clustering the PA epochs directly based on their EEG signal amplitudes, for channel Oz. The distributions are shown for an example subject, with 2 and 4 clusters. **a)** Cluster centroids, corresponding to the mean of all PA epochs assigned to each cluster. **b–e)** Cluster-wise distributions of different PA variables. The clusters are shown ordered according to their median time of PA occurrence, in ascending order (this was arbitrarily adopted as a common criterion for systematically ordering all clustering outputs, for visualization purposes). The distribution of each variable in each cluster is shown by its individual epochs, each represented by a dot, and also summarized as a box-plot (the central mark corresponding to the median of the distribution, the box edges to the 25th and 75th percentiles, and the whiskers to the 5th and 95th percentiles). Head motion (c) is shown for only one of its six parameters (X-rotation, i.e. nodding motion), to simplify the plots. **f)** 2D plot combining the respiration and cardiac period distributions. The 2D cluster distributions are summarized by ellipses delineating the 25th to the 75th percentiles in either dimension, with a central cross marking the median.

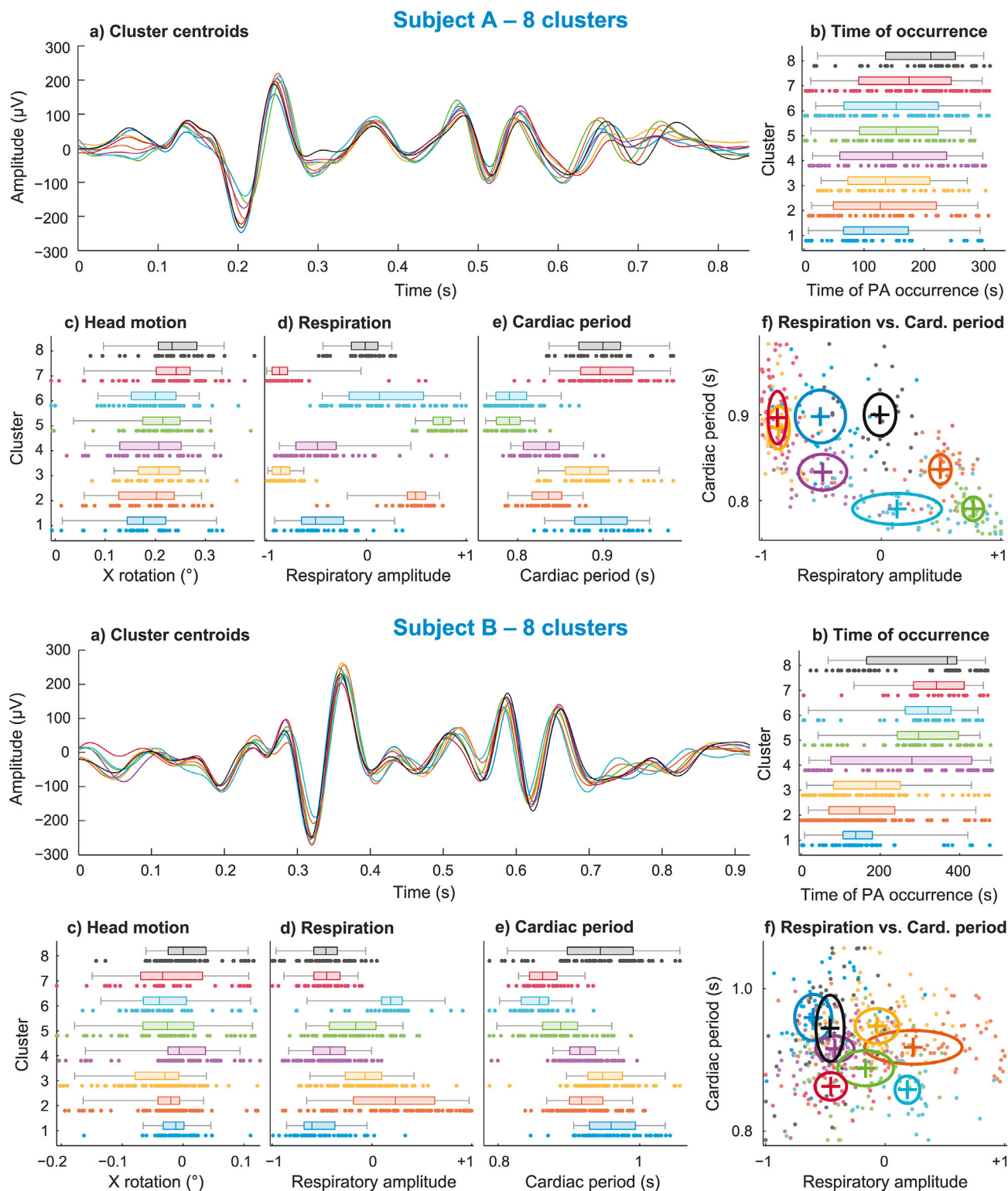


Fig. 7. Cluster-wise distribution of the PA-attributed variables (time of PA occurrence, respiratory amplitude, cardiac period and head motion) when clustering the PA epochs directly based on their EEG signal amplitudes, for channel Oz. The distributions are shown for two example subjects (subject A being the same as in Fig. 6), with 8 clusters. **a)** Cluster centroids, corresponding to the mean of all PA epochs assigned to each cluster. **b–e)** Cluster-wise distributions of different PA variables. The clusters are shown ordered according to their median time of PA occurrence, in ascending order (this was arbitrarily adopted as a common criterion for systematically ordering all clustering outputs, for visualization purposes). The distribution of each variable in each cluster is shown by its individual epochs, each represented by a dot, and also summarized as a box-plot (the central mark corresponding to the median of the distribution, the box edges to the 25th and 75th percentiles, and the whiskers to the 5th and 95th percentiles). Head motion (**c**) is shown for only one of its six parameters (X-rotation, i.e. nodding motion), to simplify the plots. **f)** 2D plot combining the respiration and cardiac period distributions. The 2D cluster distributions are summarized by ellipses delineating the 25th to the 75th percentiles in either dimension, with a central cross marking the median.

Table 2
PA correction performance with different correction techniques.

	EEG RMS (μV)		ssVEP RMS (μV)		ssVEP trial variability (μV)	
	PA center	PA periphery	Stimulation	Fixation	Stimulation	Fixation
Off-scanner (fixation only)		5.9 ± 0.6		0.2 ± 0.1		4.8 ± 0.4
Uncorrected	58 ± 4	22 ± 3	2.9 ± 0.3	3.3 ± 0.5	48 ± 3	50 ± 4
Non-adaptive ^a	24 ± 3	22 ± 3	1.7 ± 0.2	0.7 ± 0.1	21 ± 2	23 ± 3
AAS _(opt) ^b	24 ± 3	22 ± 3	1.6 ± 0.2	0.6 ± 0.1	20 ± 2	22 ± 3
OBS _(opt) ^b	17 ± 3	20 ± 3	1.4 ± 0.2	0.7 ± 0.1	15 ± 2	16 ± 3
K-means _(opt) ^b	21 ± 3	20 ± 3	1.6 ± 0.2	0.7 ± 0.1	18 ± 2	20 ± 3

Notes. The values shown correspond to averages across subjects, and the error margins to their standard error.

^a Non-adaptive average PA subtraction (equivalent to K-means with $K = 1$).

^b Adaptive correction approaches employed with optimal parameters (N_o, N_s, K), which were obtained by direct inspection of the ssVEPs recovered for each subject.

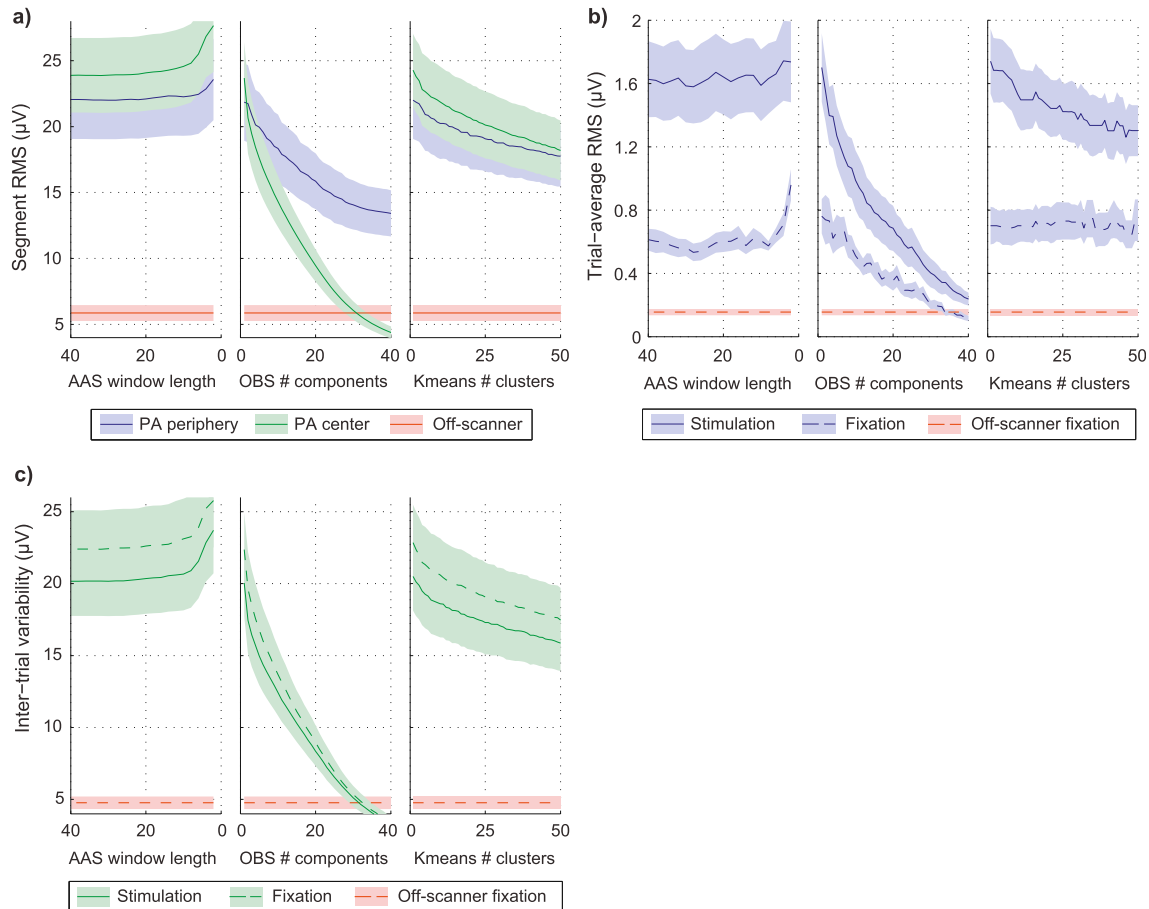


Fig. 8. Pulse artifact correction performance of three different approaches: AAS, OBS and K-means clustering, tested for a range of different values of their adaptability parameters. The curves shown correspond to averages across subjects, and the colored margins to their standard error. For clearer visualization, only the parameter sub-ranges where relevant performance variations were observed are shown. **a)** EEG signal RMS, averaged within temporal intervals located at the periphery of the PA epochs, comprising the outermost 2% of each epoch window (blue), or their center, comprising the innermost 98% of each epoch window (green); the red curve corresponds to the average RMS in subject off-scanner resting-state data. **b)** RMS of the trial-averaged ssVEP, obtained from the actual checkerboard stimulation periods (full lines) or from the same triggers after being time-shifted by 15s to the fixation periods (dotted lines), to serve as control; the red curve corresponds to the same estimate in off-scanner (resting-state fixation) data. **c)** ssVEP inter-trial variability, estimated as the standard deviation across trials; this curve was also obtained from the actual stimulation periods (full lines) or from triggers that were time-shifted to the fixation periods (dotted lines), as well as in the off-scanner (resting-state fixation) data (red).

4. Discussion

This study investigated the variability of EEG pulse artifacts across heartbeats, in simultaneous EEG-fMRI. While previous reports have pointed to the existence of PA variability, presumably caused by changes in head position and orientation throughout the recordings (Allen et al., 1998; Bonmassar et al., 2002), the present study is, to our knowledge, the first to objectively assess the magnitude and properties of PA variability,

and to investigate its underlying mechanisms, including, but not restricted to, head motion. PA variability was found to contribute to EEG variance significantly and substantially, with over $500 \mu\text{V}^2$ at 7T. Clustering analyses revealed that PA variability not only is linked to variations in head position/orientation, but also, and more importantly, to the respiratory cycle and to heart rate fluctuations. The study also investigated the impact of such variability on PA correction, an essential processing step in EEG-fMRI analysis. Overall, techniques capable of

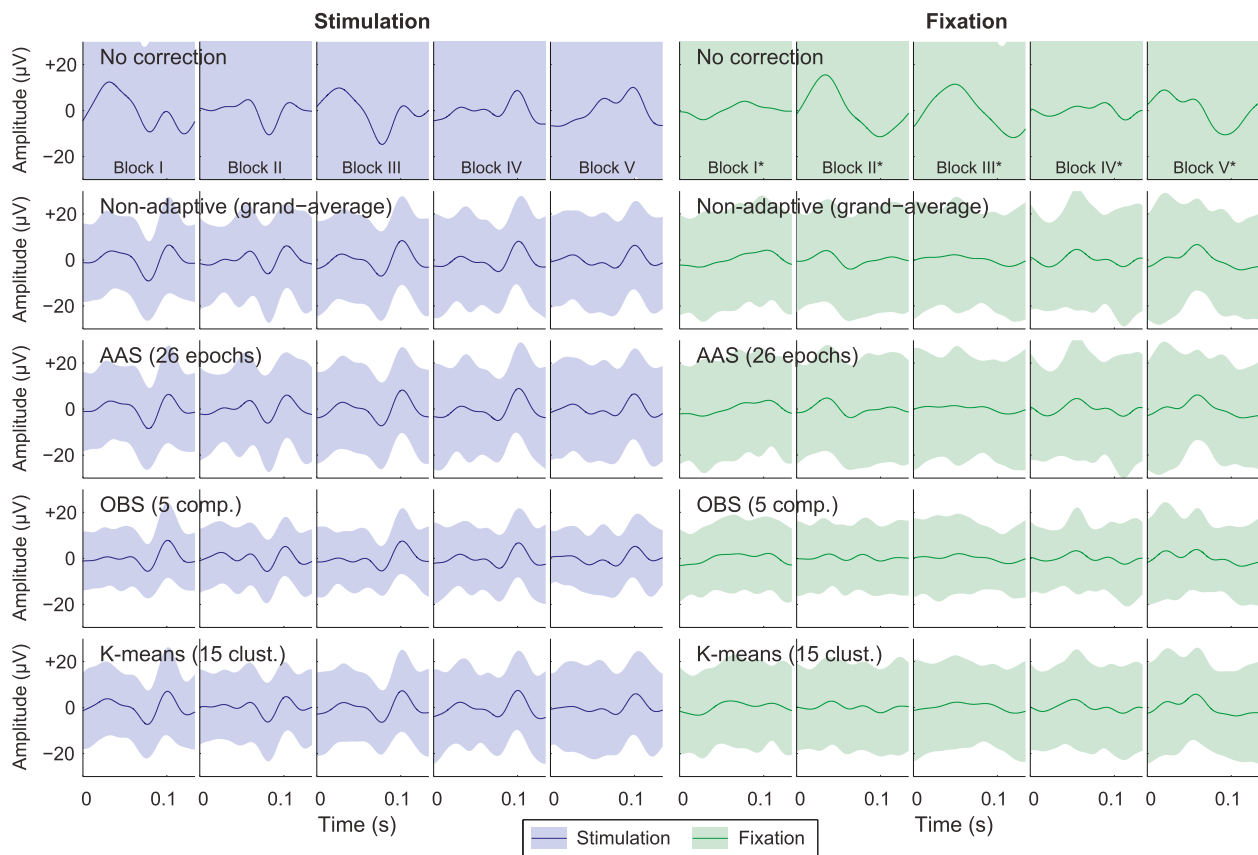


Fig. 9. Block-averaged ssVEPs captured by channel Oz of an example subject, as obtained before and after PA correction using each of the three tested approaches, at their optimal parameters. The results obtained with grand-average PA subtraction (i.e. non-adaptive approach) are also shown for comparison (2nd row). Each block-averaged response corresponds to an average over 75 ssVEP trials. For easier visualization, the block-averaged ssVEPs are shown from only five of the ten paradigm blocks (labelled I–V for stimulation, I*–V* for fixation periods) – the chosen blocks are the same for all correction approaches, and were selected as those with the lowest RMS after non-adaptive correction. The curves shown were obtained from the actual stimulation periods (blue, left column) as well as from the same triggers after being time-shifted by 15s to the fixation periods (green, right column), serving as control. The error margins represent the standard deviation across trials (i.e. checkerboard reversals).

accounting for short-term beat-to-beat variability, such as OBS and K-means clustering, tended to outperform AAS, and especially non-adaptive (grand-average) PA subtraction.

4.1. Pulse artifact variability assessment

The results obtained in this work provide evidence that PA variability across heartbeats is significant, and on average can contribute to EEG signal variance with more than $500 \mu V^2$, at 7T. This was far larger than the variability associated to the He coldheads, for instance, which has previously been found important at 7T (Mullinger et al., 2008), yet was close to negligible in the present setup (Fig. 3), possibly helped by previous improvements in signal transmission (Jorge et al., 2015b). Assuming that, like its amplitude, PA variability scales linearly with B_0 (Debener et al., 2008), this 7T estimate translates into more than $92 \mu V^2$ for the more conventional case of 3T systems. Such variability can easily overshadow various contributions from neuronal activity, thus posing an important confound for EEG analysis. Across the scalp, the distribution found for the mean PA power was consistent with previous reports (Mullinger et al., 2013), with occipital and lateral channels being the most affected (Fig. 3). The PA variability showed a fairly similar distribution, supporting the notion that channels with larger PAs are also affected by larger variance across PAs.

4.2. Pulse artifact variability mechanisms

The results of the clustering analyses demonstrate that beyond head

motion, as hypothesized, both respiration and heart rate (here analyzed in terms of cardiac period) are significantly linked to PA variability, and in fact to a greater extent than head motion per se. This is, to our knowledge, the first study that has investigated and identified these relationships. It is important to note that the respiratory amplitude and cardiac period are not themselves fully independent, given physiological mechanisms such as respiratory sinus arrhythmia (Hirsch and Bishop, 1981), which can couple the two. Indeed, in various subjects, a strong correlation between the two variables could be observed (e.g. Fig. 6), which hinders the determination of their individual contributions to PA variability, and could even question whether one variable might be linked only indirectly to PA variability, through its correlation with the other. Nevertheless, on the group level, the correlation between the two indices was fairly low, and various instances were found where variability could be clearly linked to one, but not the other (e.g. Fig. 6, bottom). Moreover, on the scalp maps obtained for intra-cluster variability reduction (Fig. 5), there were evident differences between the cardiac period (more concentrated in occipital channels) and respiratory amplitude (both occipital and lateral channels). These observations show that respiration and heart rate are both linked to PA variability, and these links are, at least to some extent, independent. This notion is also consistent with the existence of various distinct mechanisms by which the two processes may be able to affect the PA, as discussed in section 1.

Although at a lower degree of importance, changes in head position/orientation across time were nonetheless found to contribute significantly to PA variability, supporting the main assumptions of most previous studies addressing PA correction (Allen et al., 1998). Crucially, a

strong correlation was found between the head motion parameters and the time of PA occurrence, which was due to the fact that most motion patterns found in this study were dominated by slow, close-to-linear drifts across time. Slow head drifts are indeed a very commonly observed motion pattern in compliant volunteers during MRI (Gallichan et al., 2016), which explains that observation and supports that the present results should apply fairly well to studies in healthy volunteers in general. On the other hand, in more specific subject groups, such as children, head motion could acquire a more dominant role in PA variability.

Beyond the variables included in this study, it is likely that additional mechanisms may contribute to PA variability as well. For instance, the reduction in intra-cluster variability obtained when clustering the PA epochs directly based on their EEG signals (18.9%) was substantially larger than those achieved with any of the individual variables, or even their sum (Fig. 4, Table 1). Moreover, in various cases, the signal-based clustering yielded some clusters that were not clearly differentiated with respect to any of the variables (e.g. Fig. 7 bottom). This suggests that additional sources of variability are likely at play, which remain unidentified. Potential candidates would include arterial CO₂, for example, which is a strong vasodilator and can vary over periods of a few seconds in humans at rest (Wise et al., 2004). More generally, the numerous mechanisms involved in the reflex and humoral controls of circulation can affect heart function, as well as vascular tone throughout the body (Guyton and Hall, 2006), which could lead to changes in the PA across time. Given the complexity of these mechanisms, it appears unlikely that PA variability can be modelled and predicted to its full extent.

Another relevant question regarding variability mechanisms is how the different sources of the PA may be involved in its variability. One potentially useful test would be to include concurrent information from optical motion tracking systems. Some of these methods have been found sensitive enough to directly observe the small bulk head movements that are induced by cardiac pulsation (Maclaren et al., 2012), which are thought to be one of the main sources of the PA (Mullinger et al., 2013). It would therefore be valuable to investigate whether the pulsation-related bulk motion itself also exhibits substantial variability across heartbeats, and whether that variability is also linked to the PA variables under study, such as respiratory amplitude. Such insights could potentially offer further clues into how these mechanisms cause PA variability, and specifically on its different contributing sources.

4.3. Potential confounds

As PA variability was investigated directly on real EEG measurements, in the MRI scanner with concurrent fMRI acquisition, the analyses here performed were naturally susceptible to a number of confounds that could pose as apparent variability across PA epochs, while not being PA-related. Important examples are gradient artifact residuals and real EEG activity. The former confound was necessary to allow monitoring head motion; the latter could have been avoided by acquiring EEG with an insulating layer between the electrodes and the scalp (e.g. a swimming cap), but this would remove Hall effect contributions as well, which are an important component of the PA (Mullinger et al., 2013). Given these trade-offs, we opted instead to include “confound variables” for these putative contributions throughout the analysis, and they were found to have negligible effects on the observed PAs. Another potentially important confound is the presence of artifacts from spontaneous sudden subject motion, which at 7T can be comparable in power to real EEG activity (Jorge et al., 2015a); these contributions were in principle minimized by the fairly restrictive PA outlier exclusion step. The presence of other artifacts related to the MRI environment (e.g. He cold-heads) is also expected to have had a negligible impact on the analysis, given their minute contributions for the in-scanner phantom data (Fig. 3). Finally, eye blinks can often create important artifacts in frontal electrodes, and may have possibly contributed to the slightly enhanced variability reductions observed in frontal areas when direct PA

signal-based clustering was applied (Fig. 5 bottom-right). Nevertheless, these did not noticeably affect the scalp distributions observed for the other variables, and most of the more detailed analyses were focused on channels Oz and TP9, which are relatively far from the frontal areas.

On the processing side, an important aspect that could contribute to apparent PA variability is the actual triggering of PA epochs (Iannotti et al., 2015; Marino et al., 2018). Namely, our ECG recordings were performed with a single channel (as is common in EEG-fMRI setups), and at 7T the conventional ECG features become “hidden” by a larger, slower oscillation (similar to those shown by Debener et al. (2008)), likely caused by Hall effects. This artifact waveform repeats cyclically with every heartbeat, allowing for beat detection, but is substantially inaccurate for PA triggering. Even small timing mismatches can then yield large amplitude differences across the EEG PA epochs, and influence their clustering, as previously reported (Goncalves et al., 2007). In the present work, this issue was overcome by fine-tuning the ECG-derived triggers with an EEG-based moving-window correlation-maximization step, which effectively adjusts the triggers so as to minimize the variability across PA epochs (as exemplified in Supp. Fig. 1). In agreement with this, in all tests that used direct signal-based clustering (e.g. Figs. 6 and 7, Supp. Fig. 5–7), no cluster centroids were found that differed from each other by a simple constant offset – the prevalent differences, as described, occurred in peak amplitudes and temporal scaling (“dilation/contraction”).

Altogether, the results here obtained constitute a novel and relevant contribution to the current body of work investigating the PA, which has so far been more focused on spatiotemporal dynamics (Debener et al., 2008) and generating mechanisms (Yan et al., 2010; Mullinger et al., 2013), but mostly not addressing inter-epoch variability. It is important to note, nonetheless, that these results only establish correlational, not causal, links between the different variables and PA variability. Future steps to investigate causality could explore tasks of respiratory modulation, for example, or controlled changes in head position/orientation. Nevertheless, the complexity of the underlying physiological mechanisms, and their interdependencies, would remain a challenge for asserting direct causal relationships.

4.4. Pulse artifact correction techniques

Given the spatiotemporal properties of PAs, their correction is essential to the vast majority of EEG-fMRI applications (Mullinger and Bowtell, 2011). Beyond the techniques explored in this work, other relatively popular approaches include independent component analysis (Benar et al., 2003; Mantini et al., 2007; Abreu et al., 2016) and methods based on independent reference signals (Bonmassar et al., 2002; LeVan et al., 2013; Chowdhury et al., 2014), for example. These more recent techniques are often used in combination with AAS, and generally found to produce superior results, but all have their own limitations as well, including suboptimal assumptions regarding the spatiotemporal properties of the PA, and/or the underlying EEG activity (Debener et al., 2008; Mullinger et al., 2013), or the need for additional hardware, for example. The problem of PA modeling and correction is therefore not fully solved, and remains a focus of active research efforts. Understanding the mechanisms underlying PA variability could prove highly valuable to this challenge.

Exploiting the cyclic nature of cardiac activity, AAS remains one of the most widely used techniques for PA correction. Crucially, AAS assumes that the PA is stationary over periods of at least a few consecutive epochs, varying on a slower timescale putatively related to drifts in head position and orientation. The findings of this work confirm that motion-related variations are indeed present, and tend to be slow. On the other hand, PA variations due to respiration and heart rate were found to be more important, and as these variations can occur across consecutive PA epochs, AAS is bound to perform sub-optimally (e.g. Fig. 1). This is further demonstrated by the results observed with PA-based clustering, where the resulting clusters often led to the separation of epochs that occurred very close to each other in time (e.g. see time of PA occurrence

plots in Fig. 6, Supp. Fig. 5,6), while their differences in terms of respiratory amplitude and/or cardiac period were markedly clear. Consistent with this, in the data acquired in this work, AAS was found to improve data quality to some extent, but not substantially different from a non-adaptive (grand-average) subtraction, when using AAS averaging window sizes (N_o) larger than approximately 8 epochs. Moreover, for smaller windows, AAS was actually found to *increase* signal variability (Fig. 8). This is likely associated to the faster-timescale PA variability, which cannot be sufficiently averaged out when using too small windows, and could in fact contaminate neighboring epochs.

In principle, OBS and K-means can account for short-timescale PA variability, and indeed, both approaches were found more capable of reducing data variability than AAS (Fig. 8a,c, Table 2). Tying these results to those of the PA variability analysis, it is clear that K-means, for instance, can effectively account for breathing and heart rate-related variability, since the observed clusters were consistently associated with differences in respiratory amplitude and cardiac period (Figs. 6 and 7, Supp. Fig. 5–7). On the other hand, the increased flexibility of OBS and K-means carried increased risks of overcorrection, i.e. removing real EEG information. OBS, especially, was found to reduce ssVEP amplitudes fairly aggressively when increasing N_o (Fig. 8b – considering that the difference between the stimulation and fixation curves should correspond to the contribution of the actual ssVEPs); this was also observed with K-means, albeit to a smaller extent. The risk of overcorrection has been previously identified for OBS (Grouiller et al., 2007). At subject level, the optimal number of OBS components for ssVEP detection ranged from 5 to 8, although in at least 4 subjects there was no clear threshold marking a transition to overcorrection – thus making this risk harder to identify and avoid.

Overall, the performance of K-means can be considered an intermediate between AAS and OBS, both in terms of PA reduction and EEG signal preservation (Table 2). As in OBS, increasing the adaptability of K-means produced steady decreases in signal RMS and ssVEP variability, while the concomitant decreases in ssVEP amplitude were more restrained, and even close to stationary for $K < 8$. At the subject level, the optimal K tended to be higher (10–17), but generally yielded relatively marginal improvements beyond $K \approx 10$. Overall, in agreement with previous reports (Goncalves et al., 2007; de Munck et al., 2013), the performance exhibited by K-means suggests that this may be a rather favorable option for PA correction. Moreover, as with AAS, the approach can be combined with other correction methods such as ICA and reference sensor-based techniques. The fact that the epoch window can be adapted for each cluster according to its specific population is also a relevant advantage compared to OBS, especially given the observed importance of heart rate for PA variability. A limitation of K-means is that, conceptually, it describes PA variability within a “discrete domain” (defined by the number of clusters), while the evidence points to a more continuous domain of PA variations, for which OBS may be better fitting – assuming that a linear combination model can effectively capture that continuum.

One other important aspect is the suitability of these techniques for online (on-the-fly) PA correction, which is essential for real-time applications such as neurofeedback and epileptic activity monitoring (Allen et al., 1998; Zotev et al., 2014). AAS can be applied online in a rather straightforward manner, by using averaging windows containing only past epochs. OBS has recently been adapted for online correction as well, by estimating the basis (PCA decomposition) on a “buffer” set of past epochs, and then regressing out the most important principal components from the current epoch (Wu et al., 2016). To our knowledge, K-means has not been reported in such a manner to date, yet it is likely that the same principle could be applied, i.e. performing the clustering analysis on a set of past epochs, and then assigning the current epoch to the closest cluster, and subtracting its centroid. As with OBS, the complexity of the algorithm (e.g. number of past epochs in the “buffer”, number of clusters K) would have to be carefully considered with regards to processing speed.

4.5. Correction performance measures

Overall, the adopted performance measures were found suitable to describe the differences between correction methods and their impact on EEG data, particularly in light of the problem of PA variability. Nevertheless, they did reveal a number of limitations. For instance, the estimation of the signal RMS in PA center vs. periphery segments was designed taking into account the properties of the PA, and the periphery measure did yield reasonable estimates for baseline (PA-free) EEG RMS in individuals with longer cardiac periods, where PA fluctuations practically ceased in-between epochs (e.g. Supp. Fig. 1). In other cases, however, the periphery segments could not be deemed PA-free, and their RMS was reduced in PA correction as well. The ssVEP trial-average RMS in fixation, expected to comprise only artifactual contributions, was substantially higher in-scanner than off-scanner (Fig. 8b), suggesting that residual artifacts did not fully average out across the visual trials. As a result, even for the stimulation periods, this measure does not fully correspond to the ssVEP amplitude alone, and its reductions are not necessarily a symptom of overcorrection. The difference between the (in-scanner) curves of stimulation vs. fixation, however, should contain a fair indication of the actual ssVEP amplitude.

The ssVEP inter-trial variability was not significantly different between in-scanner stimulation and fixation periods (Fig. 8c), confirming that its dominant component is not the natural variability across neuronal responses, but rather ongoing artifacts. Notably, while substantially reduced by PA correction, this variability remained fairly high compared to off-scanner data, for the majority of parameters tested (except for OBS with high N_o – most likely a regime of heavy overcorrection). It is likely that at least part of the remaining variability may be due to other artifact contributions, separate from the PA. Gradient artifact residuals and environment artifacts are expected to play a fairly minor role, given the band-pass filtering range used, the results from variable-based analyses (Fig. 4), and from the phantom data (Fig. 3). In contrast, spontaneous motion remained unaccounted for, and has previously been found to cause important artifacts at 7T (Jorge et al., 2015a), thereby constituting a likely source of variability. Future work would be valuable to clarify the interplay between pulse and spontaneous motion artifacts, possibly relying on motion artifact information given by additional reference sensors.

The limitations of the adopted performance measures, discussed in the previous two paragraphs, reflect the complexity of the PA itself, in its spatiotemporal and spectral dynamics, substantial overlap with EEG activity of interest, subject specificity, and interplay with other artifacts. As seen here, as well as in previous studies (Vanderperren et al., 2010; Abreu et al., 2018), the combined use of multiple measures appears to be crucial to attain a more comprehensive picture of PA correction performance, to capture both the effects of denoising (artifact reduction), as well as preservation of the underlying EEG signals. Such comprehensive measures can then be valuable to the task of parameter optimization (e.g. choosing the averaging window size for AAS), to quantitatively indicate the most effective correction parameters among a tested range (Abreu et al., 2016). While the choice of performance measures remains an unsettled question in the field, the measures adopted in this work did succeed in providing useful insights on the performance of AAS, OBS and K-means, particularly in light of the problem of PA variability, which should prove valuable to most EEG-fMRI users.

5. Conclusion

We conclude that the variability of pulse artifacts across heartbeats has an important impact on EEG data quality acquired in the MRI environment. PA variability is linked to changes in head position and orientation, as previously hypothesized, but also, and more importantly, to respiration and to heart rate. The latter mechanisms are associated with short-timescale variability that cannot be fully captured by PA correction techniques based on windowed averaging, such as AAS, which

tended to be outperformed by more flexible approaches such as OBS and K-means clustering-based subtraction. Overall, these results offer fully novel insights into the dynamics of pulse artifacts, with important consequences for their correction, which should be taken into consideration for most EEG-fMRI applications.

Disclosure/conflicts of interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by Centre d'Imagerie BioMédicale (CIBM) of the UNIL, UNIGE, HUG, CHUV, EPFL and the Leenaards and Jeantet Foundations.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.neuroimage.2019.02.021>.

References

- Abreu, R., Leal, A., Figueiredo, P., 2018. EEG-informed fMRI: a review of data analysis methods. *Front. Hum. Neurosci.* 12, 29.
- Abreu, R., Leite, M., Jorge, J., Grouiller, F., van der Zwaag, W., Leal, A., Figueiredo, P., 2016. Ballistocardiogram artifact correction taking into account physiological signal preservation in simultaneous EEG-fMRI. *Neuroimage* 135, 45–63.
- Allen, P.J., Josephs, O., Turner, R., 2000. A method for removing imaging artifact from continuous EEG recorded during functional MRI. *Neuroimage* 12, 230–239.
- Allen, P.J., Polizzi, G., Krakow, K., Fish, D.R., Lemieux, L., 1998. Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. *Neuroimage* 8, 229–239.
- Benar, C., Aghakhani, Y., Wang, Y., Izenberg, A., Al-Asmi, A., Dubeau, F., Gotman, J., 2003. Quality of EEG in simultaneous EEG-fMRI for epilepsy. *Clin. Neurophysiol.* 114, 569–580.
- Bonmassar, G., Purdon, P.L., Jaaskelainen, I.P., Chiappa, K., Solo, V., Brown, E.N., Belliveau, J.W., 2002. Motion and ballistocardiogram artifact removal for interleaved recording of EEG and EPs during MRI. *Neuroimage* 16, 1127–1141.
- Chang, C., Cunningham, J.P., Glover, G.H., 2009. Influence of heart rate on the BOLD signal: the cardiac response function. *Neuroimage* 44, 857–869.
- Chowdhury, M.E., Mullinger, K.J., Glover, P., Bowtell, R., 2014. Reference layer artefact subtraction (RLAS): a novel method of minimizing EEG artefacts during simultaneous fMRI. *Neuroimage* 84, 307–319.
- de Munck, J.C., van Houdt, P.J., Goncalves, S.I., van Wegen, E., Ossenblok, P.P., 2013. Novel artefact removal algorithms for co-registered EEG/fMRI based on selective averaging and subtraction. *Neuroimage* 64, 407–415.
- Debener, S., Mullinger, K.J., Niaz, R.K., Bowtell, R.W., 2008. Properties of the ballistocardiogram artefact as revealed by EEG recordings at 1.5, 3 and 7 T static magnetic field strength. *Int. J. Psychophysiol.* 67, 189–199.
- Debener, S., Strobel, A., Sorger, B., Peters, J., Kranczioch, C., Engel, A.K., Goebel, R., 2007. Improved quality of auditory event-related potentials recorded simultaneously with 3-T fMRI: removal of the ballistocardiogram artefact. *Neuroimage* 34, 587–597.
- Ellingson, M.L., Liebenthal, E., Spanaki, M.V., Prieto, T.E., Binder, J.R., Ropella, K.M., 2004. Ballistocardiogram artifact reduction in the simultaneous acquisition of auditory ERPs and fMRI. *Neuroimage* 22, 1534–1542.
- Gallichan, D., Marques, J.P., Gruetter, R., 2016. Retrospective correction of involuntary microscopic head movement using highly accelerated fat image navigators (3D FatNavs) at 7T. *Magn. Reson. Med.* 75, 1030–1039.
- Glover, G.H., Li, T.Q., Ress, D., 2000. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn. Reson. Med.* 44, 162–167.
- Goldman, R.I., Stern, J.M., Engel Jr., J., Cohen, M.S., 2000. Acquiring simultaneous EEG and functional MRI. *Clin. Neurophysiol.* 111, 1974–1980.
- Goncalves, S.I., Pouwels, P.J., Kuijjer, J.P., Heethaar, R.M., de Munck, J.C., 2007. Artifact removal in co-registered EEG/fMRI by selective average subtraction. *Clin. Neurophysiol.* 118, 2437–2450.
- Gretsch, F., Marques, J.P., Gallichan, D., 2018. Investigating the accuracy of FatNav-derived estimates of temporal B0 changes and their application to retrospective correction of high-resolution 3D GRE of the human brain at 7T. *Magn. Reson. Med.* 80, 585–597.
- Grouiller, F., Jorge, J., Pittau, F., van der Zwaag, W., Iannotti, G.R., Michel, C.M., Vulliemoz, S., Vargas, M.I., Lazeyras, F., 2016. Presurgical brain mapping in epilepsy using simultaneous EEG and functional MRI at ultra-high field: feasibility and first results. *MAGMA* 29, 605–616.
- Grouiller, F., Vercueil, L., Krainik, A., Segebarth, C., Kahane, P., David, O., 2007. A comparative study of different artefact removal algorithms for EEG signals acquired during functional MRI. *Neuroimage* 38, 124–137.
- Guyton, A.C., Hall, J.E., 2006. *Textbook of Medical Physiology*, eleventh ed. Elsevier Saunders, Philadelphia.
- Hirsch, J.A., Bishop, B., 1981. Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate. *Am. J. Physiol.* 241, H620–H629.
- Iannotti, G.R., Pittau, F., Michel, C.M., Vulliemoz, S., Grouiller, F., 2015. Pulse artifact detection in simultaneous EEG-fMRI recording based on EEG map topography. *Brain Topogr.* 28, 21–32.
- Jenkinson, M., Bannister, P., Brady, M., Smith, S., 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17, 825–841.
- Jorge, J., Gretsch, F., Gallichan, D., Marques, J.P., 2018. Tracking discrete off-resonance markers with three spokes (trackDOTS) for compensation of head motion and B0 perturbations: accuracy and performance in anatomical imaging. *Magn. Reson. Med.* 79, 160–171.
- Jorge, J., Grouiller, F., Gruetter, R., van der Zwaag, W., Figueiredo, P., 2015a. Towards high-quality simultaneous EEG-fMRI at 7 T: Detection and reduction of EEG artifacts due to head motion. *Neuroimage* 120, 143–153.
- Jorge, J., Grouiller, F., Ipek, O., Stoermer, R., Michel, C.M., Figueiredo, P., van der Zwaag, W., Gruetter, R., 2015b. Simultaneous EEG-fMRI at ultra-high field: artifact prevention and safety assessment. *Neuroimage* 105, 132–144.
- Jorge, J., van der Zwaag, W., Figueiredo, P., 2014. EEG-fMRI integration for the study of human brain function. *Neuroimage* 102 Pt 1, 24–34.
- LeVan, P., Maclaren, J., Herbst, M., Sostheim, R., Zaitsev, M., Hennig, J., 2013. Ballistocardiographic artifact removal from simultaneous EEG-fMRI using an optical motion-tracking system. *Neuroimage* 75, 1–11.
- MacKay, D.J.C., 2003. *An Example Inference Task: Clustering*. Information Theory, Inference, and Learning Algorithms. Cambridge University Press, Cambridge, UK ; New York, pp. 284–292.
- Maclaren, J., Armstrong, B.S., Barrows, R.T., Danishad, K.A., Ernst, T., Foster, C.L., Gumus, K., Herbst, M., Kadashevich, I.Y., Kusik, T.P., Li, Q., Lovell-Smith, C., Prieto, T., Schulze, P., Speck, O., Stucht, D., Zaitsev, M., 2012. Measurement and correction of microscopic head motion during magnetic resonance imaging of the brain. *PLoS One* 7, e48088.
- Mantini, D., Perrucci, M.G., Cugini, S., Ferretti, A., Romani, G.L., Del Gratta, C., 2007. Complete artifact removal for EEG recorded during continuous fMRI using independent component analysis. *Neuroimage* 34, 598–607.
- Marino, M., Liu, Q., Del Castello, M., Corsi, C., Wenderoth, N., Mantini, D., 2018. Heart-brain interactions in the MR environment: characterization of the ballistocardiogram in EEG signals collected during simultaneous fMRI. *Brain Topogr.* 31, 337–345.
- Moosmann, M., Schonfelder, V.H., Specht, K., Scheeringa, M., Nordby, H., Hugdahl, K., 2009. Realignment parameter-informed artefact correction for simultaneous EEG-fMRI recordings. *Neuroimage* 45, 1144–1150.
- Mulert, C., Lemieux, L., 2010. *EEG-fMRI: Physiological Basis, Technique, and Applications*. Springer-Verlag, Berlin ; Heidelberg.
- Mullinger, K., Bowtell, R., 2011. Combining EEG and fMRI. *Methods Mol. Biol.* 711, 303–326.
- Mullinger, K., Brooks, M., Stevenson, C., Morgan, P., Bowtell, R., 2008. Exploring the feasibility of simultaneous electroencephalography/functional magnetic resonance imaging at 7 T. *Magn. Reson. Imaging* 26, 968–977.
- Mullinger, K.J., Havenhand, J., Bowtell, R., 2013. Identifying the sources of the pulse artefact in EEG recordings made inside an MR scanner. *Neuroimage* 71, 75–83.
- Neuner, I., Warbrick, T., Arrubla, J., Felder, J., Celik, A., Reske, M., Boers, F., Shah, N.J., 2013. EEG acquisition in ultra-high static magnetic fields up to 9.4 T. *Neuroimage* 68, 214–220.
- Niaz, R.K., Beckmann, C.F., Iannetti, G.D., Brady, J.M., Smith, S.M., 2005. Removal of fMRI environment artifacts from EEG data using optimal basis sets. *Neuroimage* 28, 720–737.
- Sijbers, J., Van Audekerke, J., Verhoye, M., Van der Linden, A., Van Dyck, D., 2000. Reduction of ECG and gradient related artifacts in simultaneously recorded human EEG/fMRI data. *Magn. Reson. Imaging* 18, 881–886.
- Ullsperger, M., Debener, S., 2010. *Simultaneous EEG and fMRI: Recording, Analysis, and Application*. Oxford University Press, New York.
- van Gelderen, P., de Zwart, J.A., Starewicz, P., Hinks, R.S., Duyn, J.H., 2007. Real-time shimming to compensate for respiration-induced B0 fluctuations. *Magn. Reson. Med.* 57, 362–368.
- Vanderperren, K., De Vos, M., Ramautar, J.R., Novitskiy, N., Mennes, M., Asseconci, S., Vanrumste, B., Stiers, P., Van den Bergh, B.R., Wagemans, J., Lagae, L., Snaert, S., Van Huffel, S., 2010. Removal of BCG artifacts from EEG recordings inside the MR scanner: a comparison of methodological and validation-related aspects. *Neuroimage* 50, 920–934.
- Vincent, J.L., Larson-Prior, L.J., Zempel, J.M., Snyder, A.Z., 2007. Moving GLM ballistocardiogram artifact reduction for EEG acquired simultaneously with fMRI. *Clin. Neurophysiol.* 118, 981–998.
- Wise, R.G., Ide, K., Poulin, M.J., Tracey, I., 2004. Resting fluctuations in arterial carbon dioxide induce significant low frequency variations in BOLD signal. *Neuroimage* 21, 1652–1664.
- Wu, X., Wu, T., Zhan, Z., Yao, L., Wen, X., 2016. A real-time method to reduce ballistocardiogram artifacts from EEG during fMRI based on optimal basis sets (OBS). *Comput. Methods Progr. Biomed.* 127, 114–125.
- Yan, W.X., Mullinger, K.J., Geirsdottir, G.B., Bowtell, R., 2010. Physical modeling of pulse artefact sources in simultaneous EEG/fMRI. *Hum. Brain Mapp.* 31, 604–620.
- Zotef, V., Phillips, R., Yuan, H., Misaki, M., Bodurka, J., 2014. Self-regulation of human brain activity using simultaneous real-time fMRI and EEG neurofeedback. *Neuroimage* 85 Pt 3, 985–995.