Semi-automatic identification of independent components representing EEG artifact

Filipa Campos Viola [a,b], Jeremy Thorne [a,b], Barrie Edmonds [c], Till Schneider [d], Tom Eichele [e], Stefan Debener [a,b,*]

[a] *MRC Institute of Hearing Research, Southampton, UK*
[b] *Biomagnetic Center, Department of Neurology, University Hospital Jena, Germany*
[c] *MRC Institute of Hearing Research, Nottingham, UK*
[d] *Department of Neurophysiology and Pathophysiology, Center of Experimental Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany*
[e] *Department of Biological and Medical Psychology, University of Bergen, Norway*

See Editorial, pages 841–842

ARTICLE INFO

ABSTRACT

Objective: Independent component analysis (ICA) can disentangle multi-channel electroencephalogram (EEG) signals into a number of artifacts and brain-related signals. However, the identification and interpretation of independent components is time-consuming and involves subjective decision making. We developed and evaluated a semi-automatic tool designed for clustering independent components from different subjects and/or EEG recordings.
Methods: CORRMAP is an open-source EEGLAB plug-in, based on the correlation of ICA inverse weights, and finds independent components that are similar to a user-defined template. Component similarity is measured using a correlation procedure that selects components that pass a threshold. The threshold can be either user-defined or determined automatically. CORRMAP clustering performance was evaluated by comparing it with the performance of 11 users from different laboratories familiar with ICA.
Results: For eye-related artifacts, a very high degree of overlap between users (phi > 0.80), and between users and CORRMAP (phi > 0.80) was observed. Lower degrees of association were found for heartbeat artifact components, between users (phi < 0.70), and between users and CORRMAP (phi < 0.65).
Conclusions: These results demonstrate that CORRMAP provides an efficient, convenient and objective way of clustering independent components.
Significance: CORRMAP helps to efficiently use ICA for the removal EEG artifacts.

© 2009 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.

1. Introduction

For many years, electroencephalogram (EEG) recordings have been successfully used in clinical diagnosis and cognitive brain research. However, a key characteristic of scalp-recorded EEG signals is that they consist of a mixture of an unknown number of brain and non-brain contributions. In other words, the EEG signals suffer from the presence of various artifacts, which renders the identification and analysis of brain-related EEG activity difficult (Makeig et al., 2004a). Here we present a new approach to the identification of prominent EEG artifacts. In combination with independent component analysis (ICA), this approach provides an efficient, accurate and less subjective correction procedure for multi-channel EEG recordings.

Over the past few years, ICA has gained considerable popularity for the processing of EEG signals (e.g., Debener et al., 2006; Makeig et al., 2004a). ICA performs a linear un-mixing of multi-channel EEG recordings into maximally temporally independent statistical source signals, which are further referred to as independent components (ICs). ICA belongs to a larger family of blind source separation algorithms that separate mixed signals without the aid of detailed a priori information about the nature of these signals (Hyvärinen et al., 2001). Given the lack of knowledge about the exact nature, number, and configuration of neural and non-neural sources contributing to the scalp-recorded EEG, blind source separation algorithms are particularly well suited to the decomposition of EEG data. Indeed, several laboratories have successfully demonstrated that ICA can separate multi-channel EEG recordings into meaningful brain and non-brain processes. Typical examples in-

* Corresponding author. Address: Biomagnetic Center, Department of Neurology, University Hospital Jena, Germany. Tel.: +49 3641 9325770; fax: +49 3641 9325772.
E-mail address: stefan@debener.de (S. Debener).

clude the removal of artifacts, in particular eye blinks and lateral eye movements (Jung et al., 2000a,b); the removal of stimulus-locked electrical artifacts from cochlear implants (Debener et al., 2008a; Gilley et al., 2006); or the removal of residual ballistocardi-ogram and magnetic resonance imaging (MRI) gradient artifact from EEG data recorded inside the MRI (Debener et al., 2007, 2008b; Eichele et al., 2005; Feige et al., 2005; Onton et al., 2006). Moreover, ICA has been used for the identification of neuronal event-related oscillations (Makeig et al., 2002; Onton et al., 2005) and event-related potentials (Debener et al., 2005a,b). A thorough discussion of the concepts related to the application of ICA to EEG data is provided by Onton et al. (2006).

From a practical point of view, the efficient removal of EEG artifacts is very desirable, as a proper correction substantially improves the number of trials that can be retained for event-re-lated EEG analysis. Some studies have suggested (Debener et al., 2007; Joyce et al., 2004) that the removal of some EEG artifacts by means of ICA could be implemented as a fully automatic pro-cedure if a well defined criterion or template were provided. However, it is still necessary and often mandatory to visually in-spect and evaluate the quality of ICA decompositions before arti-fact processing.

Regarding eye blink artifacts, ICA-based correction compares favourably to more frequently used linear regression procedures (Joyce et al., 2004). The eye blink correction quality that can be achieved by means of ICA is illustrated in Fig. 1. As can be seen, ICA finds components that closely resemble the topography and time course of single, representative eye blinks, and thus can separate this artifact from other EEG activity. However, the user is left with the problem of component selection, interpretation and clustering, be-
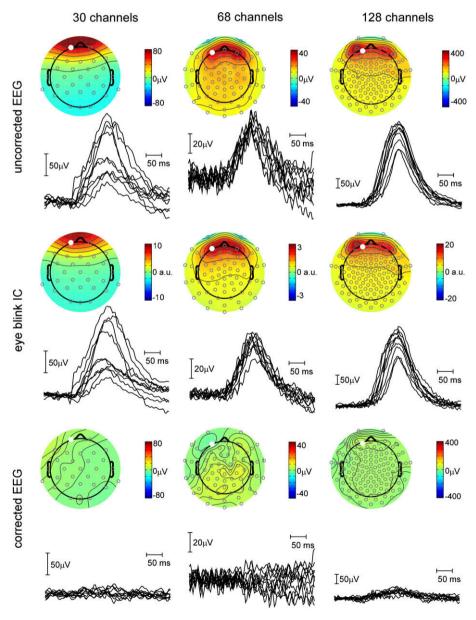

Fig. 1. ICA-based eye blink artifact correction, illustrated for three different datasets recorded in different laboratories and based on 30 (left), 68 (middle) and 128 (right) EEG channels, respectively. Top row shows 10 representative eye blinks at a channel in close proximity (68 and 128 channel datasets), or corresponding to (30 channel dataset) Fp1 of the international 10–20 system, and the mean voltage map for these eye blinks. Middle row shows the identified ICA eye blink component map (inverse weights, in arbitrary units) together with the 10 back-projected eye blinks at ~Fp1. Bottom row shows the result of the back-projection of all components except for the one shown in the middle row. Inspection of maps and voltage traces in the bottom row indicates near perfect eye blink correction for the 30 and the 68 channel datasets. Residual eye blink activity can be seen in the 128 channel dataset, illustrating our common observation that eye blinks can be represented by more than one ICA component in high-density EEG recordings.

cause ICA is usually applied to single subject datasets (for review, see Onton et al., 2006). For example, if 64-channel EEG were recorded from 20 subjects, 1280 components would require evaluation. A number of different methods can be used to guide the IC identification and selection process, such as visual inspection of IC properties (Debener et al., 2005a), a selection based on IC topographies and experimental condition effects (Debener et al., 2005b) or more formal cluster analysis procedures (e.g., Makeig et al., 2004b).

Formal cluster approaches based on the modified Mahalanobis distance are part of the EEGLAB open source environment (Delorme and Makeig, 2004). Types of IC information or features that can jointly be used for clustering comprise IC topographies (i.e., inverse ICA weights), event-related potentials (ERPs, i.e., component activation time-domain averages), spectra, time–frequency results, and source localization information. However, this approach leaves the user with a large number of parameters to determine by trial and error, as the dimensionality and relative weight for each of these features requires specification. Accordingly, clustering based on a joint consideration of multiple features is a time consuming and difficult task, regardless of the actual cluster algorithm used.

Even if an optimal configuration were to be found, an inevitable problem would still be the need to re-cluster or re-group the first level results, which would also be guided by subjective decision making rather than objective, data-driven criteria.

We developed a new, simple way of clustering, named CORRMAP, designed to identify certain prominent artifact ICs across subjects in a semi-automatic way with full user control but using a statistically guided cluster definition. We validated the performance of our template-correlation based cluster approach by comparing the results with the identification and classification of ICs representing various EEG artifacts from 11 different EEGLAB users who were familiar with ICA. This test data comprised 4256 ICs from three different studies recorded in three different laboratories.

2. Methods

2.1. CORRMAP description

CORRMAP is a semi-automatic ICA clustering tool. It requires as its main input a template map (inverse IC weights) and it operates
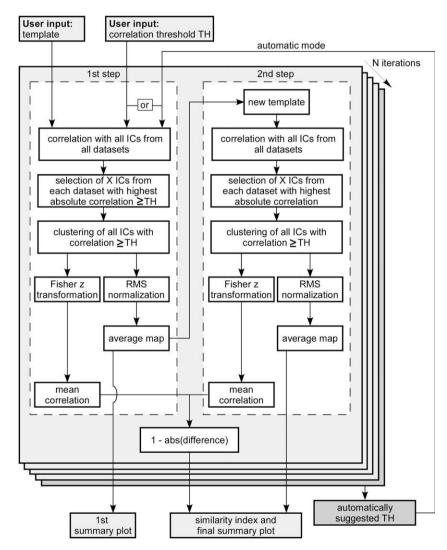

Fig. 2. Schematic flow chart of the CORRMAP tool. The main inputs are a template ICA component map selected by the user and a correlation threshold (TH) that can be selected or calculated by the tool (automatic mode). The template is compared with all component maps from all datasets by calculating a correlation value. All components with an absolute correlation equal to or greater than TH are selected to be part of the cluster and the mean correlation is calculated after Fisher *z* transformation. Up to X (usually 1–3) components per dataset are considered. This parameter can be changed by the user. An average map is calculated for the clustered components. The same procedure is repeated in a second step using this new map as the template (right column). A similarity index informs about the dependency of the result on the originally selected template. In automatic mode the procedure is repeated for *N* iterations utilizing different TH values and the TH which shows the maximum similarity index is suggested as the automatic correlation threshold.

in two different modes. In the *automatic mode*, CORRMAP identifies all ICs correlating with the template above an automatically determined threshold (see below). In the *manual mode*, CORRMAP identifies all ICs correlating with the template above a user specified threshold. A schematic illustration of all processing steps involved is shown in Fig. 2.

The core of the algorithm is a two-step loop. In the first step (Fig. 2, left) the inverse weights (i.e., IC maps) from a selected template IC are correlated with all ICs from all datasets. For each dataset, CORRMAP selects up to three ICs with the largest supra threshold correlation with the template. The maximum number of ICs selected can be changed by the user. This approach was chosen because in high-density EEG recordings, the same process (e.g., eye blinks) can be represented by more than one IC (e.g., Onton et al., 2006). Across all datasets, the selected ICs are then sorted in descending order of correlation. Here, absolute correlations are used to take into account the sign ambiguity problem (Onton et al., 2006). The mean correlation of a resulting cluster is then computed via Fisher's *z* transform, to account for the non-normal distribution of correlation values. Next, an average cluster map is calculated, after inversion of those ICs showing a negative correlation (sign ambiguity problem) and root mean square (RMS) normalization of each individual IC.

In the second step, the average cluster map obtained in the first step is then used as a new template and the same process is repeated (Fig. 2). This step evaluates the dependence of a cluster on the template IC initially selected. A similarity index (SI) was defined as one minus the absolute difference between the mean correlation values obtained from steps 1 and 2. A value close to 1 indicates that the resulting cluster is robust against the selection of the initial map, whereas a small value indicates that the initial template is not very representative of the cluster. For each of the two processing steps, a summary plot showing the template, the selected ICs, their correlations with the template and further cluster information, is produced.

The correlation threshold initially used can either be given as an input parameter (manual mode) or can be determined automatically using an iterative process (automatic mode). In automatic mode, this process consists of repeating the two core steps described above using a range of correlations from 0.95 to 0.80 in steps of 0.01. This range and step size (determined in pilot tests) results in 16 iterations returning 16 similarity indices. In cases where correlations below 0.80 are considered, CORRMAP calculates additional iterations ranging from 0.79 to 0.55 in steps of 0.01. The final correlation threshold is then determined by choosing the iteration that returned the maximum SI. This procedure is based on the rationale that, with a low correlation threshold, qualitatively different maps would be included in the clusters, resulting in a smaller SI.

The default ICA algorithm used by EEGLAB (Delorme and Makeig, 2004) is Infomax ICA, where the number of ICs is usually equal to the number of EEG channels, normally corresponding to the rank of the data. However, CORRMAP also accepts a different number of ICs per dataset (in case of rank-deficiency or prior dimensionality reduction), thus providing greater flexibility. CORRMAP can also deal with variations in EEG channel numbers within a dataset, such as happens in the case of defective channels. In this case, CORRMAP requests a channel configuration file, and the inverse weights for missing electrodes are then automatically replaced using a modified version of the EEGLAB function eeg_interp(). All CORRMAP functions are written in Matlab (The MathWorks, Inc., MA, USA) and designed as a plug-in for the EEGLAB toolbox (Delorme and Makeig, 2004). CORRMAP is available under the General Public Licence (GPL-Free Software Foundation, Inc., Boston, MA) and can be downloaded from http://www.debener.de.

2.2. Validation study

In order to evaluate CORRMAP, we compared its performance in the automatic mode with the visual identification and selection of artifactual ICs from EEGLAB users familiar with ICA (further referred to as 'users'). For that purpose, we used IC maps from three different EEG studies recorded in three different EEG laboratories and spanning 30–128 channels.

Study 1 is based on 16 resting EEG datasets from 4 different subjects, recorded inside (1.5, 3 and 7 T) and outside ($\sim$0 T) the MRI environment in Nottingham, UK, and published by Debener et al. (2008b). Briefly, the EEG data were recorded using a 30-channel MR-compatible EEG system (Brainamp MR, Brain Products GmbH, Munich, Germany) and an electrode cap with an extended 10–20 layout (Easycap, Herrsching, Germany). Extended Infomax ICA was performed on the continuous 30-channel EEG data. This study consisted of a total of 480 ICs.

Study 2 comprised auditory evoked potential recordings from 16 subjects, recorded in Southampton, UK (Hine and Debener, 2007). Continuous EEG data were recorded using a 68-channel infracerebral electrode cap (Easycap) connected to a Synamps2 amplifier (Compumedics, Charlotte, NC), and extended Infomax ICA was performed on the concatenated single-trial EEG data. This study consisted of a total of 1088 ICs.

Study 3 comprised 128-channel EEG data recorded in a cross-modal semantic priming paradigm from 21 subjects using a Brainamps MR plus amplifier system and an equidistant electrode cap provided by Easycap (Schneider et al., 2008). Data for this study were recorded in Hamburg, Germany, and consisted of a total of 2688 ICs. Further information on experimental and data processing details of the datasets used here are given in the respective publications of the three studies.

The inverse ICA weights (IC maps) from these three studies were sent to 16 users from 16 different EEG laboratories experienced with using ICA. Eleven users responded to our request and returned the classification information. The IC maps were provided as part of a Matlab program that displayed all maps in 2-D and required the user to input IC indices. For each dataset from each study, the IC indices representing three different types of artifacts, if present, had to be specified: eye blink ICs, lateral eye movement ICs and heartbeat artifact ICs. Note that users were provided only with the IC maps and did not have access to further information such as raw data or component activations. This was done to control for the information type that had to be used by the users for the classification. They received no further information except for the number of EEG channels used on each study. The maximum number of components they could select for each dataset and each artifact type was set to 3 (see above for rationale). A single example for each artifact type was provided. The selected indices were saved in a file for further analysis. Manual clustering was performed independently by the 11 users without time constraints. None of the users had access to the clusters selected by the others users. The users also indicated their experience with using ICA for removing artifacts on a Likert rating scale (from 1 = novice/beginner to 8 = expert).

2.3. Statistical analysis

CORRMAP was run in automatic mode, using as its input templates IC maps selected by visual inspection from the first dataset in each study. The output of CORRMAP was compared to the ICs selected by our users in three ways. First, we calculated the number of users that identified ICs also selected by CORRMAP for a given artifact type. Second, in order to evaluate whether users were significantly more liberal or conservative than CORRMAP, we calculated a paired *t*-test (i.e., the mean difference between the

number of ICs identified by CORRMAP and each user) for each dataset, study and artifact type. Note that this measure does not inform about the degree of overlap between the ICs identified. Thus, in a third step, we calculated the degree of overlap or association (phi) between the users, and between the users and CORRMAP. Phi represents the degree of association between two binary variables with values close to 1 representing a high degree of association, and values close to 0 representing a low association. The significance calculation of phi scores corresponds to the significance calculation used for parametric correlations.

We also calculated the proportion of ICs that were missed by the users. This is defined in respect of only those ICs picked by CORRMAP and is the ratio of the total number of ICs picked by the 11 users to the total possible (i.e., 11 multiplied by the number of ICs selected by CORRMAP). As a 'true' classification cannot be defined in real data, we used CORRMAP selection as the reference.

3. Results

In automatic mode on a typical PC (2.13 GHz CPU), it took CORRMAP between 11 s (Study 1, 480 ICs) and 44 s (Study 3, 2688 ICs) to compute the cluster and generate output figures for further inspection. We are not aware of another clustering tool capable of producing the same output within similar time parameters. Furthermore informal feedback provided by the users revealed that they required substantially more than 30 min for performing the same classification task.

The descriptive statistics and results for the significance tests for all three studies and the three artifact types analyzed are summarized in Table 1. The first three rows show the total number of ICs identified by CORRMAP for each type of artifact for each study, respectively.

In Study 1, the eye blink cluster consisted of 15 ICs from a total of 16 datasets (four subjects in four separate experimental conditions). For the eye blinks in the other two studies, the number of ICs selected by CORRMAP was greater than the total number of datasets (Study 2 = 16 datasets, Study 3 = 21 datasets), indicating that in some cases more than one IC per dataset contributed to the eye blink artifact. Fig. 3 shows a typical CORRMAP summary plot for the eye blink cluster in Study 1. Each IC map is depicted along with the absolute correlation with the template and information about the dataset to which it belongs. In this output, the mean map is shown enlarged, together with the correlation with the average map after the first iteration (below), and summary cluster information (above). The line plot in the upper right hand corner shows the sorted correlation values with the selected threshold indicated by a dashed line. A threshold value of $r = 0.94$ was automatically found by analyzing the similarity indices over a number of iterations. The similarity indices from all iterations are shown in the second line plot. A dashed line indicates the threshold used for the cluster depicted; it points towards the highest similarity index across all iterations performed.

For the other two artifact types analyzed, the total number of ICs selected per cluster by CORRMAP was smaller than the total number of datasets, except for the lateral eye movement cluster in Study 3. For this study there was one dataset that contributed more than one IC (not shown). In four out of the nine cases studied (3 types of artifact, 3 studies), a significant ($p < 0.05$, see Table 1) difference between the number of ICs selected by CORRMAP and the number of ICs selected by the users was observed. Differences were largest for heartbeat artifacts in Study 1 and eye blink artifacts in Study 3.

For the eye blink and eye movement artifacts in Studies 1 and 2 (30 and 68 channels, respectively), only a few ICs that were identified by CORRMAP were not selected by users (range between 1.2% and 11.7%, not shown) and vice versa. For Study 3 (128 channels) on the other hand, the ratio of missed ICs was 17.4% (lateral eye movements) and 25% (blinks). For the heartbeat artifact cluster this ratio ranged between 27.3% and 90.9%. This result reveals that only a few heartbeat ICs identified by CORRMAP were selected by some users, and the cluster of Study 1 includes a single IC that was not selected by any of the 11 users.

Table 2 summarizes the evaluation of the overlap between users and CORRMAP (first three rows) and across users (last three rows). High degrees of association between users and CORRMAP were found for ICs representing eye blinks (phi scores ranged between 0.83 and 0.99) and for ICs representing lateral eye movements (phi scores ranged between 0.85 and 0.91). Evaluation of the consistency across users also resulted in high phi scores for these artifact types, suggesting that independent users were similarly consistent in their classification between themselves as they were with CORRMAP. However, for ICs representing heartbeat artifacts phi score calculations revealed only low to moderate degrees of association both within users (range 0.19–0.65) and between CORRMAP and users (range 0.07–0.71). This suggests that the identification of heartbeat artifacts by ICs is more difficult than the identification of eye blinks or lateral eye movements.

The high degree of overlap between users and CORRMAP is illustrated for the eye blink cluster of Study 2 in Fig. 4. The number of users that indicated each IC is displayed on the top of each map. The ICs are sorted in descending order of correlation with the cluster average (not shown). In 19 out of the 24 ICs, a perfect match between users and CORRMAP was evident; that is, all 11 users identified these 19 maps as representing eye blink artifacts. Of the other five ICs selected by CORRMAP, only four were identified by fewer than five users, indicating a moderate discrepancy.

Fig. 5 illustrates two types of discrepancy between CORRMAP and users. Fig. 5A shows an example of two ICs selected by CORRMAP and both contributing to an eye blink artifact, but with only one being consistently identified by all users. Fig. 5B, on the other hand, shows one IC that was not selected by CORRMAP but was labelled as an eye blink by some of the users. In this case 4 out of 11 users mis-interpreted a possible brain event-related IC (cf. Delorme et al., 2007) as an eye blink. Topographically, this IC indeed resembled a typical eye blink, but did not actually contribute to eye blinks, as revealed by a comparison of the raw data with the respective IC time course.
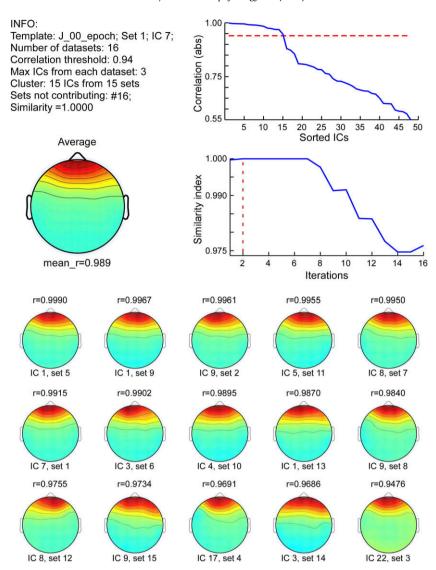
Table 1
Number of independent components (ICs) Identified by CORRMAP and by users for three artifact types for three studies.

Study	Artifact type					
	Blink	Lateral eye movements		Heartbeat		
	Number of ICs identified by CORRMAP					
(1) 30 Channels	15	13		4		
(2) 68 Channels	24	15		7		
(3) 128 Channels	47	22		7		
	Number of ICs identified by users					
	Mean	SD	Mean	SD	Mean	SD
(1) 30 Channels	15.27	0.47	16.45	3.45	12.55	10.11
(2) 68 Channels	23.73	3.52	17.55	2.77	9.00	3.22
(3) 128 Channels	38.63	3.98	22.45	3.96	8.82	5.55
	t-Test between users and CORRMAP (two-tailed)					
	$t(10)$	p Value	$t(10)$	p Value	$t(10)$	p Value
(1) 30 Channels	1.94	0.08	3.33	<0.01	2.80	0.02
(2) 68 Channels	−0.26	0.80	3.05	0.01	2.06	0.07
(3) 128 Channels	−6.97	<0.001	0.38	0.71	1.09	0.30

Note. p values <0.05 were considered significant.

INFO:
Template: J_00_epoch; Set 1; IC 7;
Number of datasets: 16
Correlation threshold: 0.94
Max ICs from each dataset: 3
Cluster: 15 ICs from 15 sets
Sets not contributing: #16;
Similarity =1.0000

Fig. 3. Example CORRMAP output figure showing the eye blink artifact component cluster from 16 datasets recorded with 30 EEG channels. The plot displays information about the cluster (top left) and, in the top right corner, the correlations sorted in descending order and with the correlation threshold used indicated in red. Below, the similarity indices are plotted, illustrating the result of the automatic mode threshold detection. The iteration picked by the automatic mode is indicated in red. Below, all component maps (inverse weights, in arbitrary units) identified as belonging to this cluster are shown, together with their correlation with the template and information about the original dataset and component index therein. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

The high degree of association found for lateral eye movements is illustrated by the cluster of Study 2 in Fig. 6A. Out of the 16 subjects, 15 contributed one IC each to the CORRMAP cluster. In 10 out of the 15, a perfect match between users and CORRMAP was evident, and only a single IC was selected by fewer than 10 users. Here, as in the cluster shown in Fig. 4, a very high similarity between the resulting IC maps was found, irrespective of the polarity reversal across ICs that can cause confusion. Fig. 6B, on the other

Table 2
Degree of association between CORRMAP clusters and users' identification of three artifact types in three studies.

Study	Artifact type					
	Blink		Lateral eye movements		Heartbeat	
	Mean	Range	Mean	Range	Mean	Range
	Association between CORRMAP and users					
(1) 30 Channels	0.99	[0.93 1.00]	0.91	[0.71 0.96]	0.07[a]	[−0.02 0.56]
(2) 68 Channels	0.89	[0.85 0.94]	0.89	[0.75 0.94]	0.62	[−0.01 0.84]
(3) 128 Channels	0.83	[0.76 0.87]	0.85	[0.61 0.95]	0.71[a]	[−0.01 0.85]
	Association between users					
(1) 30 Channels	0.99	[0.92 1.00]	0.93	[0.73 0.99]	0.19[a]	[0.02 0.33]
(2) 68 Channels	0.89	[0.82 0.93]	0.90	[0.71 0.97]	0.65	[0.07 0.76]
(3) 128 Channels	0.91	[0.83 0.98]	0.75	[0.55 0.82]	0.58[a]	[0.21 0.73]

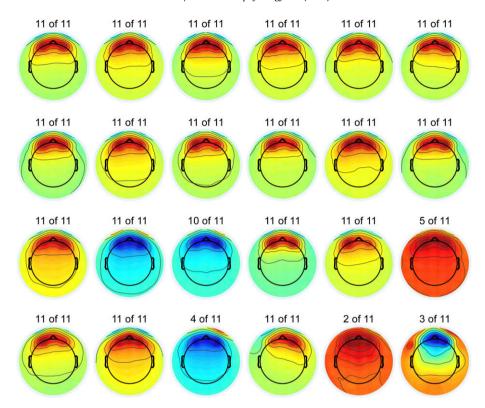[a] One user was excluded from the analysis.

Fig. 4. CORRMAP validation result for eye blink ICA components based on 16 subjects and 68 EEG channel recordings. The cluster was obtained by running CORRMAP in automatic mode, which selected 24 components with a correlation value equal to or greater than 0.87. The number of users that labelled these components as representing an eye blink artifact is represented at the top of each component map. Maps represent inverse weights in arbitrary units.
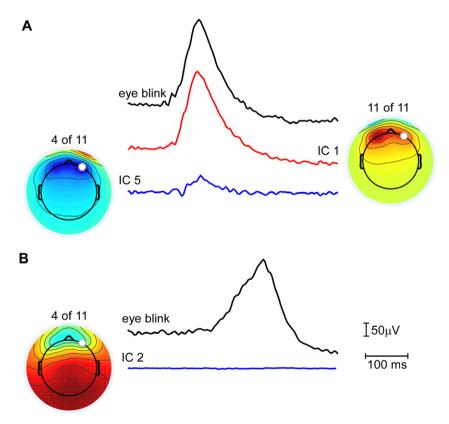

Fig. 5. Two examples showing inconsistencies between CORRMAP results and user selection. (A) Illustration of a representative eye blink artifact for a frontal channel (black), and back-projected activities at this channel for the two ICA component maps displayed (blue and red). Note that the left component was identified by only 4 out of 11 users, but shows a contribution to the eye blink. (B) An example where 4 out of 11 users have indicated an eye blink component not selected by CORRMAP. Inspection of the component activity (in blue) in comparison to a representative channel eye blink (black) does not support the interpretation of this component as representing eye blinks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

hand, illustrates the lower level of agreement found for the heartbeat cluster of Study 2. In this case, CORRMAP found only seven ICs from seven different datasets out of the 16 datasets in this study. Note that none of these ICs was identified by all users.

4. Discussion

The aim of the present study was to evaluate a simple and efficient procedure for the clustering of ICs representing EEG artifacts. ICA has become a popular and powerful choice for removing EEG artifacts (e.g., Jung et al., 2000a), but it requires the correct interpretation of ICs by the user. This interpretational step is required for brain-related as well as artifactual ICs, which, ideally, should be robust across independent observations (i.e., subjects). Component identification and evaluation is a time-consuming and potentially error-prone process, as a large number of ICs needs to be considered. Typically, the number of ICs in a study is given by the product of the number of EEG channels and the number of subjects. The EEGLAB plug-in CORRMAP developed here can help to screen large numbers of components quickly and objectively, and thus provides guidance for the identification and efficient removal of EEG artifacts such as eye blinks and lateral eye movements.

In contrast to other available clustering approaches (Delorme and Makeig, 2004), CORRMAP introduces a strategy that is focused on just a single feature (inverse ICA weights). This allowed us to code CORRMAP capabilities in a simple, quick, easy to revise and user friendly way, while keeping the number of subjective decisions to be performed by the user to a minimum: Users only need to choose one template IC map to initiate clustering. In the current version of CORRMAP, we have focused on the inverse IC weights as the single clustering parameter. It should be noted, however, that other features may be more useful for clustering other types of processes identified by ICA. ICA for example has been shown to disentangle mu rhythms from EEG alpha activity (e.g., Makeig et al., 2002), but this classification probably requires the consideration of spectral information in addition to, or instead of, topographical information (Makeig et al., 2004b).
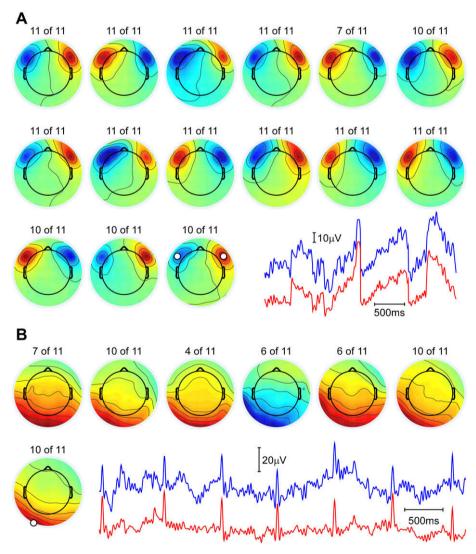

Fig. 6. CORRMAP validation result for lateral eye movement (A) and heartbeat artifact (B) ICA components based on 16 subjects and 68 EEG channel recordings. (A) The cluster was obtained in automatic mode, which selected 15 components with a correlation value equal to or greater than 0.91. An example of lateral eye movements is shown for the raw data (blue, linear derivation of left and right fronto-lateral channels) and the back-projected component (red, for the component with indicated electrode locations). (B) Similar plot for the heartbeat artifact IC cluster. CORRMAP automatic mode identified 7 components with a correlation value equal to or greater than 0.91. In A and B, the number of users that labelled the components as representing the respective artifacts is displayed on the top of each component map. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

It is our experience that a careful visual inspection of EEG raw data, and the ICA decomposition, helps to substantially improve the quality of the decomposition and ultimately the quality of the artifact correction and thus the signal quality that can be achieved. However, if the focus is on ICA-based artifact correction, CORRMAP quickly guides the visual inspection of ICA decompositions and reduces the time necessary for data evaluation to a minimum. It may be argued that, in order to maximize the performance of CORRMAP, the template selected should be representative of the type of artifact to be removed. This selection in itself requires some experience with ICA and the consistency of ICA decompositions across different recordings. The CORRMAP output facilitates the identification of representative ICs, and in the fully automatic mode, the resulting cluster is to a substantial extent independent of the exact template chosen, as long as the template belongs to the same group of ICs. It is possible to quickly and easily compare the effects of different templates on the clustering output of CORRMAP. This approach not only helps to select representative cluster templates, but also helps to build up experience in using and understanding the benefits and limitations of ICA in the processing of EEG data. Accordingly, CORRMAP also provides some potential for the teaching of lab members about the identification, consistency and interpretation of ICs.

In many situations it should be sufficient for the user to choose the automatic mode feature, allowing the tool to suggest the best correlation threshold. This approach would be particularly useful for less experienced ICA users, or for situations where CORRMAP is being used to evaluate the robustness of ICA by evaluating the presence of specific components. In our experience the automatic mode reveals reasonable results, in particular for eye blink and lateral eye movement IC clusters, but it is important to regard the automatic threshold as a first guiding value only. In some situations it may be necessary to adjust the threshold after inspection of the cluster initially obtained.

Importantly, by comparing the classification of 11 users with CORRMAP, we observed that there was a large overlap in the selection of ICs representing eye blinks and lateral eye movements, probably because all users are very experienced with these types of common EEG artifacts. The main benefit of artifact removal with CORRMAP is that it provides an objective, repeatable and quick method for identifying artifact-related ICs.

On a descriptive level, the overlap between users and CORRMAP was larger for studies comprising fewer channels and therefore fewer ICs. We attribute the low consistency observed for high-density data to the ICA 'over-fitting' problem that is more evident in high-density than low-density EEG recordings. With high-density recordings it is commonly observed that the same physiological process can be represented in a number of ICs (typically less than 4), making its identification more complicated and thus error-prone. As a result, several ICs that account for the same process can be included in the same decomposition, and the number of ICs to be attributed to the same process may thus vary across datasets and laboratories, causing some confusion. CORRMAP addresses this issue by allowing the selection of up to three ICs per dataset for any one artifact. On the other hand, users with less experience in analysing high-density data may have expected only one IC, or very few ICs, as representative of a physiological process such as eye blinks. In this case, we would conclude that using CORRMAP can result in a cluster of ICs more representative of the artifact in question than might be possible for an inexperienced user.

Much less prominent, and therefore less well known by EEG researchers (including many users that participated in the validation study), are heartbeat artifacts. The prominence of heartbeat artifacts in EEG data depends on the recording reference, with the nose-tip reference usually allowing for a better identification than linked earlobes or vertex. The other factor is the spatial sampling of the head sphere, and thus the recording montage used. The recording montage used in Study 2 (Hine and Debener, 2007) included infracerebral electrode sites, similar to the layout of the geodesic sensor net as provided by Electrical Geodesics Inc. (Eugene, OR), to improve the spatial sampling of the EEG. However, electrodes placed at the lower half of the head sphere are closer to the heart, and thus prone to pick up more electrical heartbeat activity by means of volume conduction. As a result, ICA decompositions of Study 2 included ICs reflecting a heartbeat artifact in most data sets, which was not the case for Studies 1 and 3. In Study 1 (Debener et al., 2008b), a scalp reference (Fcz) was used in combination with a 10–20 electrode layout, whereas in Study 3 (Schneider et al., 2008), although a nose-tip reference was used, electrode layout was similar to the 10–10 system only. Moreover, in Study 1, most ICs classified by users as heartbeat ICs in fact probably reflected residual ballistocardiogram activity, which is typical of EEG data recorded inside an MRI scanner (Debener et al., 2008b). The topographies of these ICs resemble those that can be attributed to electrical heartbeat activity, but, as only two users were familiar with analysing EEG data recorded inside an MRI scanner, a mis-attribution may have contributed to the rather poor overlap between CORRMAP and users. Furthermore, heartbeat artifact, and the related topography, is less well known among EEG researchers than, say, eye blinks, probably because it less frequently affects EEG recordings. Accordingly, the results also represent, to some extent, the familiarity of users with the different artifact topographies investigated, among which the heartbeat artifact topography is probably the least common.

While a 'true' best classification cannot be easily determined in real data, the examples discussed above highlight possible reasons for poor classification outcomes and poor inter-rater reliability. It should be noted, however, that a detailed investigation of the sensitivity of CORRMAP was beyond the scope of this study. Such a validation approach would require the use of artificial data, where the ground truth (i.e., the number and type of artifact ICs per dataset) is known. A study based on simulated data could be performed to examine, and further compare, the performance of users and software (such as CORRMAP), and would complement the current approach.

In conclusion, CORRMAP has proved to be efficient, quick, and at least as consistent as a group of 11 ICA users from different laboratories in the classification of eye blink and lateral eye movement ICs. This was made possible by focusing solely on topographic information as a single clustering parameter. Other types of information should of course be considered for the detailed examination of ICs, in particular those representing brain-related activity (e.g., Debener et al., 2005a,b; Makeig et al., 2002, 2004a; Onton et al., 2005) or more complex artifacts such as those caused by cochlear implants (Debener et al., 2008a). CORRMAP could be further optimized to take into account such parameters, making it potentially useful for clinical applications. However, if the focus is on EEG artifact removal, in particular eye blinks and lateral eye movements, then CORRMAP in combination with ICA provides a powerful, user-friendly approach.

References

Debener S, Makeig S, Delorme A, Engel AK. What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis. Cogn Brain Res 2005a;22(3):309–21.

Debener S, Ullsperger M, Siegel M, Fiehler K, von Cramon DY, Engel AK. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. J Neurosci 2005b;25(50):11730–7.

Debener S, Ullsperger M, Siegel M, Engel AK. Single-trial EEG/fMRI reveals the dynamics of cognitive function. Trends Cogn Sci 2006;10(12):558–63.

Debener S, Strobel A, Sorger B, Peters J, Kranczioch C, Engel AK, et al. Improved quality of auditory event-related potentials recorded simultaneously with 3-T fMRI: removal of the ballistocardiogram artefact. Neuroimage 2007;34(2):590–600.

Debener S, Hine J, Bleeck S, Eyles J. Source localization of auditory evoked potentials after cochlear implantation. Psychophysiology 2008a;45(1):20–4.

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 2008b;67(3):189–99.

Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 2004;134(1):9–21.

Delorme A, Westerfield M, Makeig S. Medial prefrontal theta bursts precede rapid motor responses during visual selective attention. J Neurosci 2007;27(44):11949–59.

Eichele T, Specht K, Moosmann M, Jongsma ML, Quiroga RQ, Nordby H, et al. Assessing the spatiotemporal evolution of neuronal activation with single-trial event-related potentials and functional MRI. Proc Natl Acad Sci USA 2005;102(49):17798–803.

Feige B, Scheffler K, Esposito F, Di Salle F, Hennig J, Seifritz E. Cortical and subcortical correlates of electroencephalographic alpha rhythm modulation. J Neurophysiol 2005;93(5):2864–72.

Gilley PM, Sharma A, Dorman M, Finley CC, Panch AS, Martin K. Minimization of cochlear implant stimulus artifact in cortical auditory evoked potentials. Clin Neurophysiol 2006;117(8):1772–82.

Hine J, Debener S. Late auditory evoked potentials asymmetry revisited. Clin Neurophysiol 2007;118(6):1274–85.

Hyvärinen A, Karhunen J, Oja E. Independent component analysis. New York: Wiley; 2001.

Joyce CA, Gorodnitsky IF, Kutas M. Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. Psychophysiology 2004;41(2):313–25.

Jung TP, Makeig S, Humphries C, Lee TW, McKeown MJ, Iragui V, et al. Removing electroencephalographic artifacts by blind source separation. Psychophysiology 2000a;37(2):163–78.

Jung TP, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin Neurophysiol 2000b;111(10):1745–58.

Makeig S, Westerfield M, Jung TP, Enghoff S, Townsend J, Courchesne E, et al. Dynamic brain sources of visual evoked responses. Science 2002;295(5555):690–4.

Makeig S, Debener S, Onton J, Delorme A. Mining event-related brain dynamics. Trends Cogn Sci 2004a;8(5):204–10.

Makeig S, Delorme A, Westerfield M, Jung TP, Townsend J, Courchesne E, et al. Electroencephalographic brain dynamics following manually responded visual targets. Plos Biol 2004b;2(6):747–62.

Onton J, Delorme A, Makeig S. Frontal midline EEG dynamics during working memory. Neuroimage 2005;27(2):341–56.

Onton J, Westerfield M, Townsend J, Makeig S. Imaging human EEG dynamics using independent component analysis. Neurosci Biobehav Rev 2006;30(6):808–22.

Schneider TR, Debener S, Oostenveld R, Engel AK. Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. Neuroimage 2008;42(3):1244–54.