

Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects

Tzyy-Ping Jung^{a,b,*}, Scott Makeig^{a,c}, Marissa Westerfield^{a,d}, Jeanne Townsend^{a,d},
Eric Courchesne^{a,d}, Terrence J. Sejnowski^{a,b}

^aUniversity of California, San Diego, La Jolla, CA 92093, USA

^bHoward Hughes Medical Institute and Computational Neurobiology Lab, The Salk Institute, P.O. Box 85800, San Diego, CA 92186-5800, USA

^cNaval Health Research Center, P.O. Box 85122, San Diego, CA 92186-5122, USA

^dChildren's Hospital Research Center, La Jolla, CA 92037, USA

Accepted 23 June 2000

Abstract

Objectives: Electrical potentials produced by blinks and eye movements present serious problems for electroencephalographic (EEG) and event-related potential (ERP) data interpretation and analysis, particularly for analysis of data from some clinical populations. Often, all epochs contaminated by large eye artifacts are rejected as unusable, though this may prove unacceptable when blinks and eye movements occur frequently.

Methods: Frontal channels are often used as reference signals to regress out eye artifacts, but inevitably portions of relevant EEG signals also appearing in EOG channels are thereby eliminated or mixed into other scalp channels. A generally applicable adaptive method for removing artifacts from EEG records based on blind source separation by independent component analysis (ICA) (Neural Computation 7 (1995) 1129; Neural Computation 10(8) (1998) 2103; Neural Computation 11(2) (1999) 606) overcomes these limitations.

Results: Results on EEG data collected from 28 normal controls and 22 clinical subjects performing a visual selective attention task show that ICA can be used to effectively detect, separate and remove ocular artifacts from even strongly contaminated EEG recordings. The results compare favorably to those obtained using rejection or regression methods.

Conclusions: The ICA method can preserve ERP contributions from all of the recorded trials and all the recorded data channels, even when none of the single trials are artifact-free. © 2000 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Artifact removal; Electrooculographic; Independent component analysis; Single-trial event-related potentials; Event-related potential

1. Introduction

Single-trial event-related potentials (ERPs) consist of brief epochs of electroencephalographic (EEG) activity time-locked to experimental events of interest. These recordings are usually averaged prior to analysis to increase their signal/noise ratio. Here, 'noise' includes non-phase-locked EEG signals and non-neural artifacts such as eye blinks and eye movements. However, ERP averaging may not cancel some artifacts induced by eye movements or blinks if they are time-locked to experimental events. These artifacts may seriously interfere with correct ERP analysis and interpretation. In addition, data from frontal and temporal electrodes located near the eyes or scalp

muscles are often discarded since these are more heavily contaminated by artifacts than central scalp channels. Another common strategy is to reject all EEG epochs containing artifacts larger than some arbitrarily selected EEG voltage value. However, when limited data are available, or when blinks and muscle movements occur too frequently as in children and some patient groups, the amount of data lost to artifact rejection may be unacceptable. For example, Small (1971) reported a visual ERP experiment conducted on autistic children who produced electrooculographic (EOG) artifacts in nearly 100% of the trials. In this case, the presence of large background EEG signals not time- and phase-locked to experimental events may make ERP averages of the few artifact-free trials too unstable to permit useful analysis.

One approach to reducing contamination from eye movement artifacts is to regress out reference signals collected near the eyes. Regression methods have been proposed

* Corresponding author. Institute for Neural Computation, University of California, San Diego, Department 0523, La Jolla, CA 92093-0523, USA. Tel.: +1-619-453-4100, ext. 1455; fax: +1-619-587-0417.

E-mail address: jung@salk.edu (T.-P. Jung).

using both time domain (Hillyard and Galambos, 1970; Verleger et al., 1982) and frequency domain techniques (Whitton et al., 1978; Woestenburg et al., 1983). All regression methods, whether in time or frequency domains, depend on having one or more clean reference channels (e.g. one or more 'EOG' channels) which cannot be further analyzed after regression. However, these methods share an inherent weakness, in that both eye movements and EEG signals propagate to periocular ('EOG') sites. Therefore, regression-based artifact removal procedures also eliminate neural activity common to the reference electrodes and to other frontal electrodes. Because the regression coefficients are determined largely by the typically large EOG signals, regression methods may also introduce neural activity projecting to the reference channel into other sites (Jung et al., 2000).

Principal component analysis (PCA) has been proposed as a method to remove eye artifacts from multichannel EEG (Berg and Scherg, 1991). However, PCA cannot completely separate eye artifacts from brain signals, especially when they have comparable amplitudes (Lagerlund et al., 1997; Jung et al., 1998b, 2000). By combining PCA, multiple source models for EOG and EEG, and an artifact-aligned averaging method (Lins et al., 1993), Berg and Scherg (1994) demonstrated a more effective PCA-based approach to correct eye artifacts. However, the accuracy of their method depends on the availability of separate and accurate inverse source solutions for EEG and EOG. Building an accurate EEG source model requires a priori knowledge of event-related brain activity following not only stimuli, but also blinks and saccades. The method also relies on the amount and quality of separately recorded calibration data which are needed to provide estimates of the source vectors and transmission coefficients in the EOG model.

Makeig et al. (1996) proposed an approach to the analysis of EEG and ERP data based on an unsupervised neural network learning algorithm that takes a logistic infomax approach to performing independent component analysis (ICA) (Bell and Sejnowski, 1995). They showed that this ICA algorithm can be used to separate neural activity from muscle and blink artifacts in spontaneous EEG data and reported its use for finding independent components of EEG and ERP data and for tracking changes in alertness (Makeig et al., 1997; Jung et al., 1998c). Subsequent independent work (Vigário, 1997) based on a related approach verified that different artifacts can also be detected in multichannel magnetoencephalographic (MEG) recordings. However, this study did not attempt to remove the identified artifacts. Jung et al. (1998a,b, 2000) introduced an ICA-based method based on an extended infomax ICA algorithm (Girolami, 1998; Lee et al., 1999). This method can be used to detect and remove a wide variety of artifacts (including eye blinks, muscle noise, heart signal, and line noise) from spontaneous EEG data.

This study demonstrates, through analysis of sample data sets collected in a visual spatial selective attention task, that

the ICA-based method can also be used to remove stimulus-induced eye artifacts from single-trial ERP records. The method uses spatial filters derived by the ICA algorithm, avoiding the need for separate reference channels for each artifact source, and allowing analysis of neural ERP activity projecting to periocular (EOG) channels. Here, we analyze experimental data collected from 28 normal controls and from 22 clinical subjects (10 autistic and 12 brain lesion subjects) who had difficulty in inhibiting unwanted eye movements toward peripheral target stimuli.

2. Materials and methods

2.1. Subjects

Data were collected from 28 normal controls, 10 high-functioning autistic and 12 brain lesion subjects. All subjects had normal or corrected-to-normal vision. The control subjects had no history of substance abuse, special education, major medical or psychiatric illness, developmental or neurologic disorder. The autistic subjects met DSM-III-R (American Psychiatric Association, 1987) criteria for autistic disorder, as well as criteria from the Autism Diagnostic Interview, the Autism Diagnostic Observation Schedule (Lord et al., 1994) and the Childhood Autism Rating Scale (Schopler et al., 1980). Lesion sites in the patients were verified by neuroradiological examination of magnetic resonance images. The stroke and autistic subjects had no additional neurologic or psychiatric diagnoses.

2.2. EEG data collection

EEG data were recorded from 31 scalp electrodes, 29 placed at locations based on a modified International 10–20 system, one placed below the right eye (VEOG) and one placed at the left outer canthus (HEOG). All 31 channels were referred to the right mastoid and were digitally sampled for analysis at 256 Hz. Subjects participated in a 2 h visual spatial selective attention task in which they were instructed to attend to filled circles flashed in random order in 5 boxes laterally arrayed 0.8 cm above a central fixation point. Stimulus locations were outlined by 5 evenly spaced 1.6 cm blue squares displayed on a black background at visual angles of 0° , $\pm 2.7^\circ$ and $\pm 5.5^\circ$ from fixation. One (attended) location was marked by a green square throughout each 72 s experimental block. Subjects were instructed to maintain fixation on the central cross and to press a response button as quickly as possible each time they saw a filled circle appear in the attended location. The location of the attended square was counterbalanced across experimental blocks (for further details see Makeig et al., 1999a; Townsend and Courchesne, 1994). Prior to analysis, we rejected those trials in which the subject blinked or moved their eyes at the moment the visual stimulus was presented

since the eye movement might have impaired the perception of, and response to, the stimulus.

2.3. Independent component analysis

Independent component analysis (ICA) (as defined by Comon, 1994) is a method for solving the *blind source separation* problem: to recover N independent source signals, $\mathbf{s} = \{s_1(t), s_2(t), \dots, s_N(t)\}$ (e.g. different voices, music, or other sound sources) from N linear mixtures, $\mathbf{x} = \{x_1(t), \dots, x_N(t)\}$, modeled as the result of multiplying the matrix of source activity waveforms, \mathbf{s} , by an unknown square matrix \mathbf{A} (i.e. $\mathbf{x} = \mathbf{A}\mathbf{s}$). Given almost no advance knowledge of the nature of the sources or of the mixing process, the task is to recover a version, \mathbf{u} , of the original sources, identical to \mathbf{s} , save for scaling and source order. To do this, it is necessary to find a square matrix, \mathbf{W} , specifying filters that linearly invert the mixing process, i.e. $\mathbf{u} = \mathbf{W}\mathbf{x}$. The key assumption used to distinguish sources from mixtures is that sources, s_i , are statistically independent, while their mixtures, x_i , are not. In contrast with decorrelation techniques such as PCA, which ensure only that output pairs are uncorrelated ($\langle u_i u_j \rangle = 0, \forall ij$), ICA imposes a much stronger criterion, statistical independence, which occurs when the multivariate probability density function (p.d.f.) factorizes: e.g.

$$f_{\mathbf{u}}(\mathbf{u}) = \prod_{i=1}^N f_{u_i}(u_i)$$

Statistical independence requires that all second-order and higher-order correlations of the u_i are zero, while decorrelation only seeks to minimize second-order statistics (covariance or correlation).

Bell and Sejnowski (1995) proposed a simple neural network ‘infomax’ algorithm that blindly separates mixtures, \mathbf{x} , of independent sources, \mathbf{s} , using information maximization (infomax). They showed that maximizing the joint entropy, $H(\mathbf{y})$, of the output of a neural processor minimizes the mutual information among the output components, $\mathbf{y}_i = g(\mathbf{u}_i)$, where $g(\mathbf{u}_i)$ is an invertible bounded non-linearity and $\mathbf{u} = \mathbf{W}\mathbf{x}$. Recently, Lee et al. (1999) extended the ability of the infomax algorithm to perform blind source separation on linear mixtures of sources having either sub- or super-Gaussian distributions based on ideas from Girolami (1998) and Lee et al. (1999). For further details, see these sources and Jung et al. (2000).

2.4. Applying ICA to single-trial EEG data

Use of ICA for blind source separation of EEG data is based on two plausible premises: (1) EEG data recorded at multiple scalp sensors are linear sums of temporally independent components arising from spatially fixed, distinct or overlapping brain or extra-brain networks; (2) the spatial spread of electric current from sources by volume conduction does not involve significant time delays. For further

details regarding ICA assumptions underlying EEG analysis, see Makeig et al. (1997, 1999a) and Jung et al. (1998a).

Fig. 1 presents a schematic illustration of the procedure. In EEG analysis, the rows of the input matrix, \mathbf{x} , are EEG signals recorded at different electrodes and the columns are measurements recorded at different time points (Fig. 1A, left). ICA finds an ‘unmixing’ matrix, \mathbf{W} , which decomposes or linearly unmixes the multichannel scalp data into a sum of temporally independent and spatially fixed components, $\mathbf{u} = \mathbf{W}\mathbf{x}$. The rows of the output data matrix, \mathbf{u} , are time courses of activation of the ICA components. The columns of the inverse matrix, \mathbf{W}^{-1} , give the relative projection strengths of the respective components at each of the scalp sensors (Fig. 1A, right). These scalp weights give the scalp topography of each component, and provide evidence for the components’ physiological origins (e.g. eye activity should project mainly to far frontal sites). The projection of the i th independent component onto the original data channels is given by the outer product of the i th row of the component activation with the i th column of the inverse matrix, and is in the original units (e.g. μV). Scaling information and polarity are distributed between the activation waveforms and the maps, and the true size of a component is given only by the size of its projection. In this paper, all scalp maps were interpolated from 31 EEG channels and referred to the original right-mastoid reference. For each component, the amplitudes of scalp maps (individually scaled color bars, Fig. 1A, right panels) were the projection of the component (in μV) at a given time point (vertical blue line).

Artifact-free event-related brain signals were obtained by projecting selected non-artifactual ICA components back onto the scalp, $\mathbf{x}_0 = \mathbf{W}^{-1}\mathbf{u}_0$, where \mathbf{u}_0 is the matrix, \mathbf{u} , of activation waveforms with rows representing activations of artifactual (or irrelevant) sources set to zero (Fig. 1B).

2.5. Numerical methods

Extended-ICA decomposition was performed on 31 channel, 1 s data epochs from 500 to 700 target stimulus trials for each of 28 normal controls and the 22 clinical subjects using routines coded in MATLAB 5 and C (available from <http://www.cnl.salk.edu/~scott/ica.html>) running on a Pentium II 400 MHz PC with 512 MB RAM. Only target trials in which the subject pressed the response button within the allowed (200–1000 ms) time window were analyzed. Previous results (Makeig et al., 1997, 1999a,b; Jung et al., 2000) and those we report here show that ICA decomposition is relatively insensitive to the exact choice of learning rate or batch size.

Here, we decomposed all the target epochs at once to study the ICA decomposition of the small evoked responses themselves (Makeig et al., 1999a; Jung et al., 1999, 2000). However, if artifact removal is the ultimate goal, it is not necessary to decompose all 500–700 1 s epochs recorded over 2 h in an experiment at once. Earlier (Jung et al.,

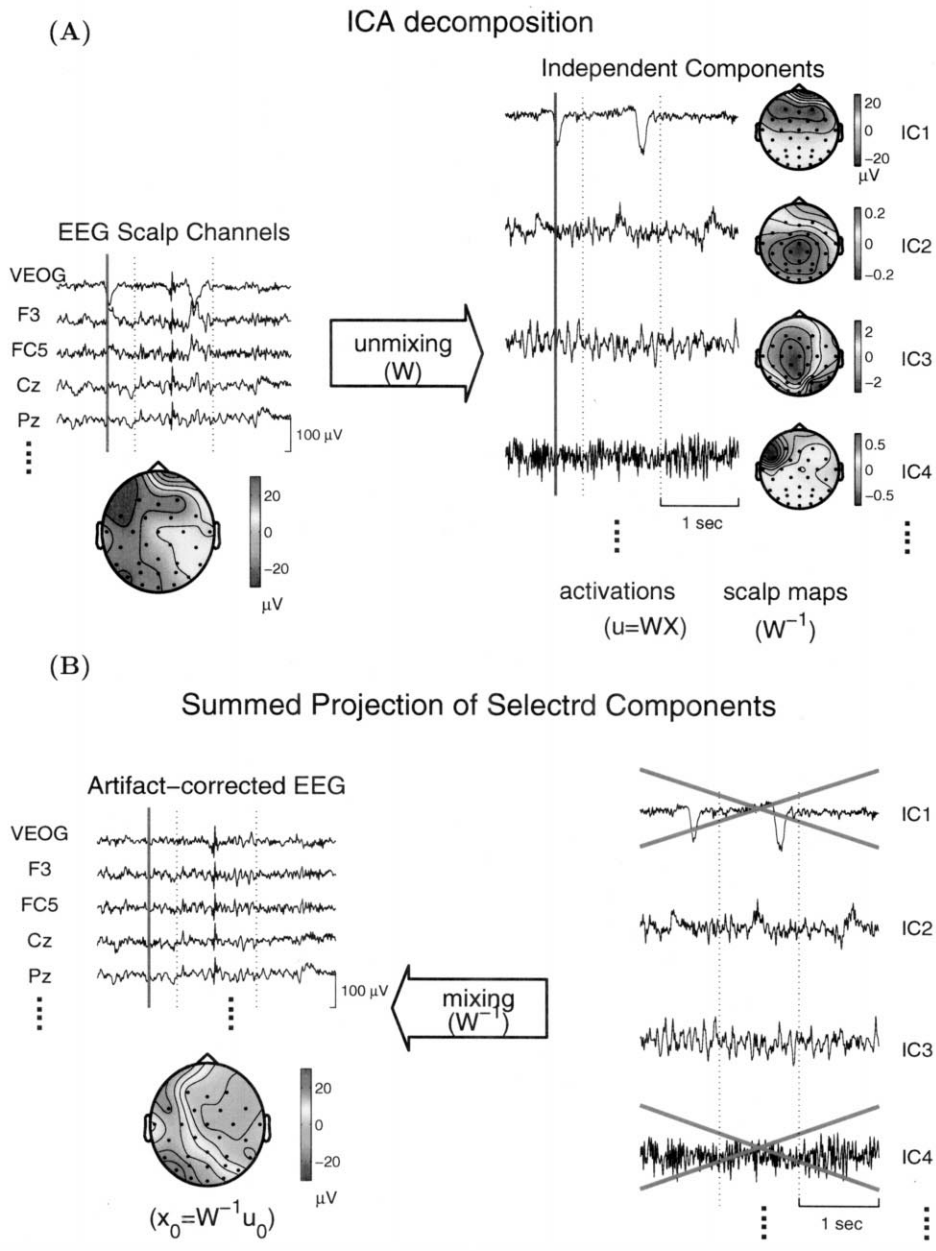


Fig. 1. Schematic overview of ICA applied to EEG data. (A) A matrix of single-trial EEG data, \mathbf{x} , recorded at multiple scalp sites (only 4 are shown), is used to train an ‘unmixing’ weight matrix, \mathbf{W} , to minimize the statistical dependence of the equal number of outputs, $\mathbf{u} = \mathbf{W}\mathbf{x}$ (4 shown here). After training, ICA components consist of time series (the rows of \mathbf{u}) giving the time courses of activation of each component, plus fixed scalp topographies (the columns of \mathbf{W}^{-1}) giving the projections of each component onto the scalp sensors. (B) Some ICA components account exclusively or predominantly for artifactual activity, for example component IC1, generated by blinks, or IC4, generated by temporal muscle activity. Others account for various evoked and/or spontaneous EEG activity (e.g. IC2, IC3). Artifact-free EEG signals, \mathbf{x}_0 , can be obtained by mixing and projecting back onto the scalp channels selected non-artifactual ICA components (IC2 + IC3) by multiplying the selected activation waveforms, \mathbf{u}_0 , by the inverse mixing matrix, \mathbf{W}^{-1} .

1998a,b, 2000), we showed that eye, muscle and line-noise artifacts may be separated from brain activity by training ICA on as little as 10 s of EEG recording, taking less than 1 min to converge.

2.6. Regression analysis

To compare the relative effectiveness of ICA for artifact removal to previous methods, we implemented the multiple-

lag regression model of Kenemans et al. (1991). In this model, the effect of the EOG channel signal on the other EEG channels at each sampling time t is given by

$$eeg(t) = EEG(t) - \sum_{g=0}^T (\beta_g eog(t - g)), \quad \text{where } \beta_g = SS^{-1} sp_g$$

Here EEG denotes the ‘true’ EEG (without eye artifacts), while eeg and eog are the recorded EEG and EOG signals,

respectively, and T is the maximum time lag. The sequence of lagged regression coefficients, β_g , describes the instantaneous and delayed effects of the EOG on the EEG. The vector sp_g of length $(T + 1)$ contains the inner products of $eeg(t)$ and $eog(t - g)$ ($g = 0, \dots, T$), while SS is the $(T + 1) \times (T + 1)$ matrix of inner products of $eog(t - g)$. This method adjusts for both frequency- and phase-dependent differences in EOG-to-EEG transfer functions (Kenemans et al., 1991).

2.7. Identifying artifactual components

Most of the 31 independent components derived by ICA applied to 31 channel single-trial target response data accounted for distinct activities arising from different brain or extra-brain networks. We found *blink-related component(s)* among them by the following procedure. (1) We first viewed the 31 channel ICA activations on a scrolling display and searched for the component(s) with time courses resembling blink activity (brief, large monopolar potentials). (2) We verified the nature of the candidate components by plotting their scalp topographies, which provided further evidence as to their physiological origin (using the heuristic that eye activity projects most strongly to far frontal sites).

Other independent components accounting for *eye movements* were found using the following procedure. (1) We separately averaged the single-trial EEG records time-locked to target stimuli presented at 5 different locations. (2) We applied the spatial filters derived by ICA based on all the single-trial EEG data to the 5 resulting 31 channel ERPs and searched for components whose time courses varied systematically with stimulus locations. Our assumption here was that involuntary saccadic movements following stimulus presentations would systematically vary in amplitude and direction as a function of the distance and direction from the fixation point to the target location. We found such a systematic relationship for components whose scalp maps suggested they accounted for saccadic eye movements, but not for other components with different scalp maps accounting for other brain (or extra-brain) activities. (3) We verified the nature of the candidate eye movement components by examining their scalp topographies. Again, we reasoned that components accounting for lateral eye movements should project most strongly to far frontal sites, and should show a polarity difference between the two periocular sites.

3. Results

We present here the analysis of representative data from two normal (30- and 31-year-old) male control subjects, one 32-year-old autistic subject and one 55-year-old female stroke patient whose lesion involved the right frontal-temporal-parietal cerebral cortex. Results for the remaining 46 subjects can be seen at <http://www.cnl.salk.edu/~jung/ERPartifact.html>.

For each subject, ICA decomposition was performed on 500–700 (31 channel, 1 s) data epochs time-locked to target stimulus presentations. We first identified components accounting for blink and eye movement artifacts according to the procedures described above. The remaining derived independent components displayed a variety of distinct relations to task events. The activations of some components were clearly time- and phase-locked to stimulus onsets, while others were time- and phase-locked to button presses (Jung et al., 1999). Still other components captured various types of oscillatory or other background EEG phenomena; these will be detailed elsewhere.

3.1. Removing blink and eye movement artifacts

Example 1. ICA was applied to a target response data set collected from a 32-year-old male autistic subject performing the visual selective attention experiment. Despite our request that he minimize blinks, about 50% of the trials were contaminated by eye blinks as judged by the common convention of detecting and rejecting trials of EEG voltages at periocular sites that exceeded a pre-set threshold. Extended-ICA successfully isolated blink artifacts to a single independent component (Fig. 2A) whose contributions could be removed from the EEG records by subtracting its component projection from the data (Jung et al., 1998a,b). Though the subject was instructed to fixate the central cross during each 72 s block, the technician watching the video monitor noticed that the subject's eyes also tended to move slightly towards target stimuli presented at peripheral locations. A second ICA component accounted for these small horizontal eye movements (Fig. 2B).

Fig. 2B (5 traces) shows separate ERP averages (at periocular site HEOG) of responses to targets presented at the 5 different visual field locations. The size of the prominent eye movement-related component is proportional to the angle between the stimulus location and the fixation point. Its scalp pattern is also consistent with the scalp pattern expected for lateral eye movements. Note the overlap in scalp topography between the two independent components accounting for blinks (Fig. 2A) and for lateral eye movements (Fig. 2B). ICA component maps need not be orthogonal and may even be nearly spatially coincident.

The left panel in Fig. 2C shows all 641 single-trial ERP epochs recorded at HEOG. Here, we use a recently developed visualization tool, the 'ERP image' (Jung et al., 1999, 2000), to show all 641 single-trial ERP epochs time-locked to onsets of target stimuli (left vertical line). Single-trial event-related responses are plotted as gray-scaled horizontal traces (see scale bar) sorted by the subject reaction time (thick black line). The ERP average of these trials is plotted below the ERP image. The middle panel in Fig. 2C shows the summed signals in each trial identified as blink and eye movement artifacts. The right panel in Fig. 2C shows the corrected single-trial records obtained by subtracting the components accounting for artifacts (middle panel) from

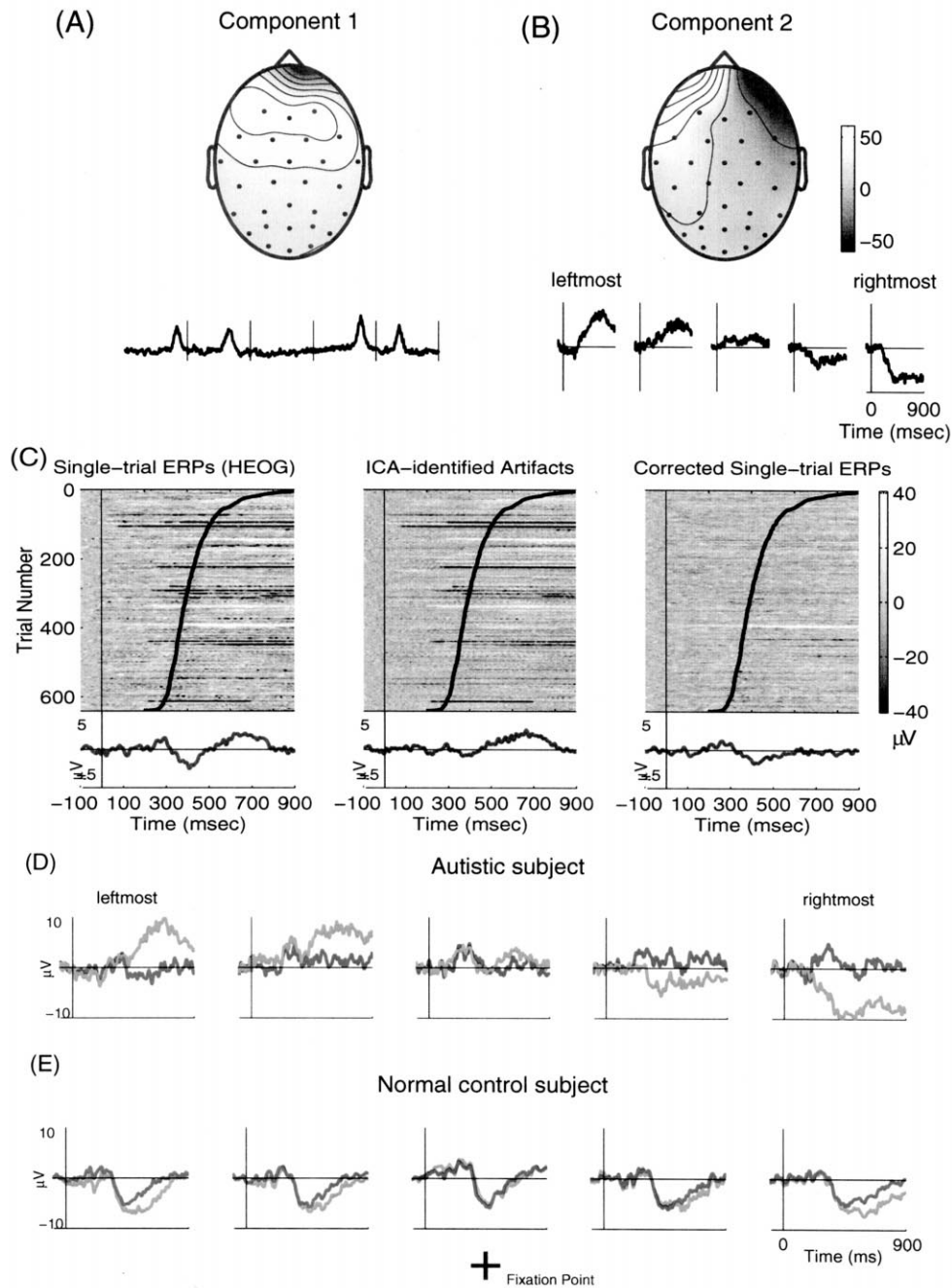


Fig. 2. Removal of eye blink and eye movement artifacts in normal and autistic subjects. (A) Scalp topography and 5 consecutive 1 s target response epochs showing the time course of activation of an ICA component accounting for blink artifacts. This component was separated by ICA from 641 target response trials recorded from an adult autistic subject in the visual selective attention experiment. (B) The scalp topography of a second component (an eye movement component) and its averaged activation time courses in response to target stimuli presented at the 5 different attended locations. (C) (Left) ERP images of single-trial ERPs at periocular site HEOG (same autistic subject) time-locked to 641 targets presented at all 5 attended locations, sorted by response time (thick black line). (Middle) Summed projections at HEOG of the two ICA components accounting for artifacts. (Right) Corrected single-trial ERPs obtained by subtracting the ICA-extracted artifacts (right) from the original data (left). (D) Averaged ERPs at site HEOG to targets presented at each of 5 attended locations, before (faint traces) and after (solid traces) artifact removal. Removing blink and eye movement artifacts from the single-trial ERPs revealed the relative independence of the remaining small visual response from stimulus location. Note that the differences in the target ERPs before removal of the two ICA components accounting primarily for eye movements are progressively larger in response to stimuli presented away from the central fixation point, consistent with a tendency for this subject to move his eyes towards the stimulus locations. (E) The same ICA-based artifact removal procedure applied to single-trial data collected from a normal control subject. Here, the differences between the uncorrected and corrected ERPs were progressively larger in responses to stimuli presented away from the central fixation point.

the original records (left panel). A large number of blink and eye movement artifacts (center panel, horizontal line segments) were removed from the records by this procedure.

While most ERP reports ignore data collected at periocular locations, ICA allows ERP data at periocular electrode sites to be separated into components generated by eye activity and components generated by brain activity. Fig. 2D shows the averaged ERPs from a patient at site HEOG in response to stimuli presented at the 5 different attended locations before (faint traces) and after (solid traces) artifact removal. The artifact-corrected averaged periocular responses to stimuli presented at the 5 different locations revealed a small periocular positivity near 250 ms that did not depend on stimulus location.

Fig. 2E shows the same artifact removal procedure applied to data collected from a normal control subject. Conventional threshold-based methods for rejecting contaminated EEG epochs could not wholly avoid summing ocular artifacts, since in this example saccade amplitudes were often smaller than typical (50–75 μV) thresholds for rejecting blinks and eye movements. After artifact correction by ICA, the responses at site HEOG to targets presented at the 5 different locations (solid traces) were again quite similar and showed strong inferior frontal activity recently identified with a fronto-parietal early-P300 subcomponent (P3f) peaking at the moment of motor command in faster-responding subjects (Makeig et al., 1999a). If, alternatively, the periocular data channels (including HEOG) had been used as reference signals to regress out contributions to signals at adjacent sites (as in many studies, for example Hillyard and Galambos, 1970; Woestenburger et al., 1983), cerebral activity expressed in those channels would have been subtracted from every scalp site, and the reference sites themselves would have become silent.

3.2. Verifying the artifactual nature of the removed signals

After artifactual signals were removed from the original EEG data, we needed to verify that (1) the signals removed by the ICA-based method did not contain cerebral activity, and (2) the remaining signals contained no or minimal artifacts. To this end, we first viewed the histogram of the maximum potentials (positive or negative) at two periocular channels (VEOG and HEOG) for each subject, and used local minima in the histogram as thresholds for splitting all the single target response trials (from 100 ms before to 900 ms after stimulus onsets) into 3 subsets: least contaminated, moderately contaminated, and most heavily contaminated, according to their absolute maximum potential values at the two periocular channels. Using this method, the threshold voltage separating least from moderate and moderate from heavily contaminated trials varied between individual subjects. Across all 50 subjects, the threshold separating least contaminated from moderately contaminated trials ranged from 25 to 40 μV ; the threshold separat-

ing moderately from heavily contaminated trials ranged from 70 to 100 μV .

We then applied the spatial filters derived by ICA to the data for each subject and removed activities accounting for artifacts as described in the previous section. Our underlying assumptions were as follows. (1) If the signals removed using ICA did not contain any cerebral activity, the average of the least contaminated trials should differ little before and after artifact removal. The average of the heavily contaminated trials, on the other hand, should show significant differences. (2) If all the eye movement artifacts were effectively removed from the EEG records, the corrected ERP averages should have very similar waveforms and scalp maps, since they were collected from the same scalp locations from the same subject performing the same task on the same day.

Example 2. Fig. 3 shows the ERP averages of the least (351 trials), moderately (132 trials), and heavily contaminated (72 trials) target trials from a normal control subject before (Fig. 3A) and after (Fig. 3B) artifact removal. Before artifact removal (Fig. 3A), the averaged ERPs differed largely because of the contamination from blinks and eye movements. After artifact removal (Fig. 3B), they differed little, presumably because they now contained only electrical activity of the brain. Note also that the averaged ERPs after artifact removal were very similar to the average of least contaminated trials, and the moderately and heavily contaminated trial averages showed larger pre-removal and post-removal differences, respectively.

Fig. 3C shows the scalp topographies at the P300 peak (505 ms) for the 3 trial averages before (top 3 panels) and after (lower 3 panels) artifact removal. Note that the scalp maps for the 3 ‘corrected’ ERPs (lower 3 panels) are remarkably similar and strongly resemble the artifact-free scalp map of the least contaminated trials (upper left panel). This result indicates that ICA may preserve all of the recorded EEG signals even if the single trials and raw averages are heavily contaminated by blinks and eye movements. Note that if a simple rejection method were used on these data (e.g. discarding all trials with maximum EOG potentials above 75 μV), the averaged ERPs would be the averages of only the least and moderately contaminated trials, and that these would still contain some artifacts even after removal of 13% of the trials.

3.3. ICA-based artifact removal versus artifact rejection

Example 3. We used 545 single trials collected from a patient with frontal-temporal-parietal lesion to compare the effectiveness of artifact removal using the ICA-based method with results of the standard artifact rejection method. The artifact rejection procedure employed in the data collection laboratory was to discard blink contaminated trials containing potentials exceeding 75 μV at either sites VEOG or HEOG. This procedure rejected all but 49 of 545 trials of the patient’s data. Since blinks and muscle move-

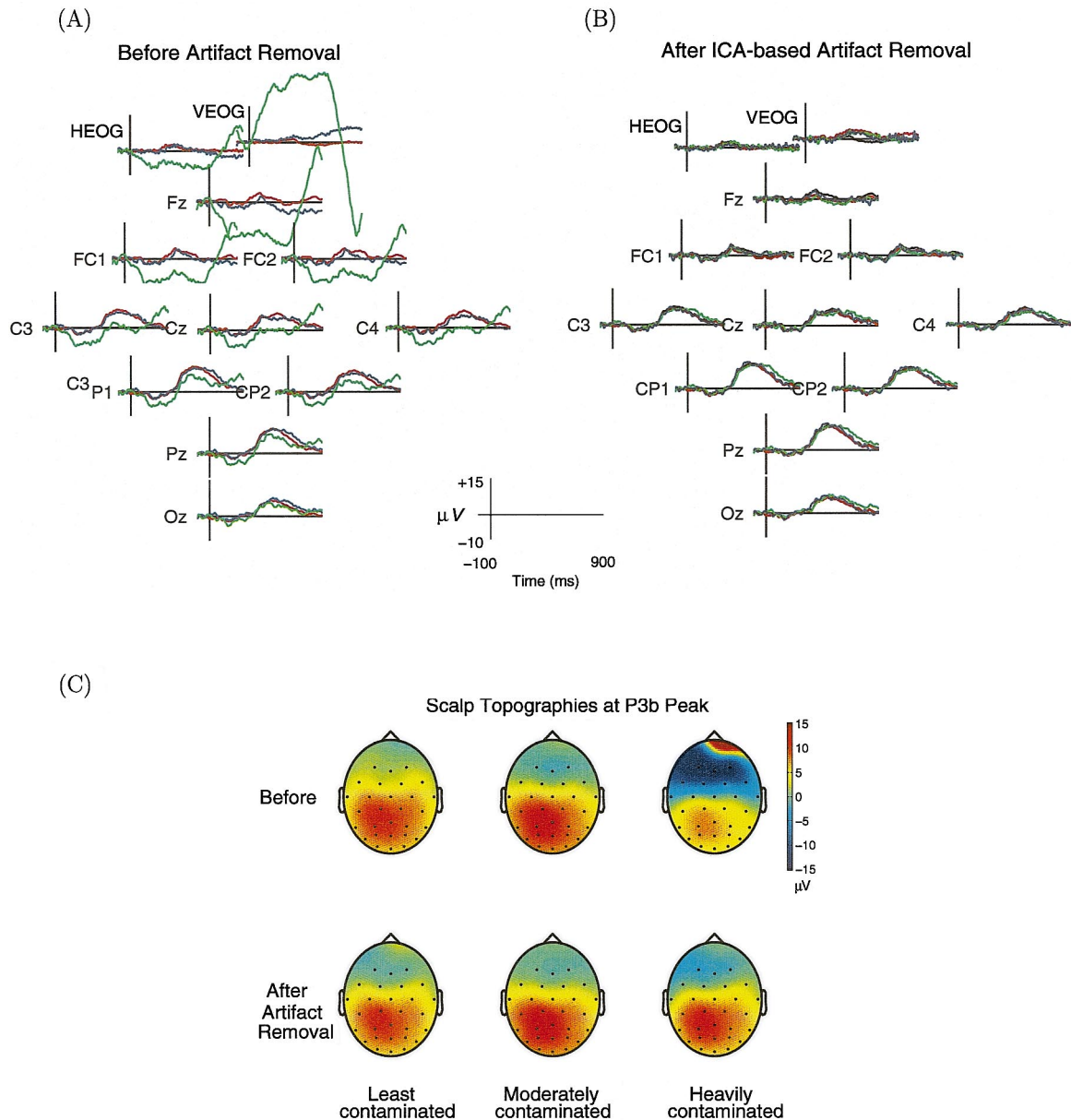


Fig. 3. Removal of eye blink and eye movement artifacts from single trials. (A) Averages of least (red traces, 351 trials), moderately (blue traces, 132 trials) and heavily (green traces, 72 trials) contaminated single-trial target response epochs from a 31 year-old normal control subject. (B) Averages of artifact-corrected ERPs for the 3 trial groups. (C) Scalp topographies at the P300 peak (505 ms) for the least (left), moderately (center) and heavily (right) contaminated trial groups before (top) and after (bottom) artifact removal.

ments occurred so frequently, the amount of data lost to standard artifact rejection would be unacceptable, and so use of an effective artifact removal procedure was crucial to the use of this patient's data for further analysis.

Fig. 4A shows the averaged target responses of the lesion subject before (blue traces) and after (red traces) ICA-based artifact removal and standard artifact rejection (green traces). Our assumption underlying this comparison was that after an effective artifact removal procedure, the corrected averages would resemble the average of 'artifact-free' trials (if enough of these were recorded).

In the figure, the average of the 49 remaining trials after

artifact rejection appears highly contaminated by large alpha-frequency EEG activity. However, at frontal sites its (low-frequency) contour resembles the average of all 545 'ICA-corrected' trials more than that of the uncorrected trials, supporting the inference that the signals removed by the artifact removal procedure were indeed artifacts. At posterior sites which were relatively uncontaminated by eye movements, problems resulting from averaging an insufficient number of 'artifact-free' (e.g. 49) trials are apparent. The average is overlaid with strong remnants of alpha EEG activity that clearly compromise accurate measures of ERP peak amplitude and latency. The ICA-

corrected average is free both of the large eye movement artifacts present in the uncorrected average of all trials, and of the non-phase-locked background EEG present in the average of the few ‘artifact-free’ trials.

Fig. 4B shows the scalp maps at the ‘P300’ peak (near 457 ms) in averages of all 545 uncorrected trials (left), the 49 ‘artifact-free’ trials (center) and of all 545 trials after artifact removal (right). As expected, the P300 scalp map for the uncorrected trial average included a strong negativity at far frontal sites produced by blinks and eye movements. The artifact-free average showed no recognizable scalp

pattern because of alpha contamination, while the corrected trial average generally resembled the ‘P300’ scalp topography found for normal control subjects (Makeig et al., 1999a), albeit with smaller than normal amplitude.

Fig. 4C shows the averaged ERPs for this patient at site HEOG in response to target stimuli presented at the 5 different attended locations, before (blue traces) and after (red traces) ICA-based artifact removal. As has been observed in patients with similar lesions (Hécaen, 1962; Ládvavas et al., 1997), the patient clearly found it difficult to inhibit ipsilesional eye movements, as indicated by the progres-

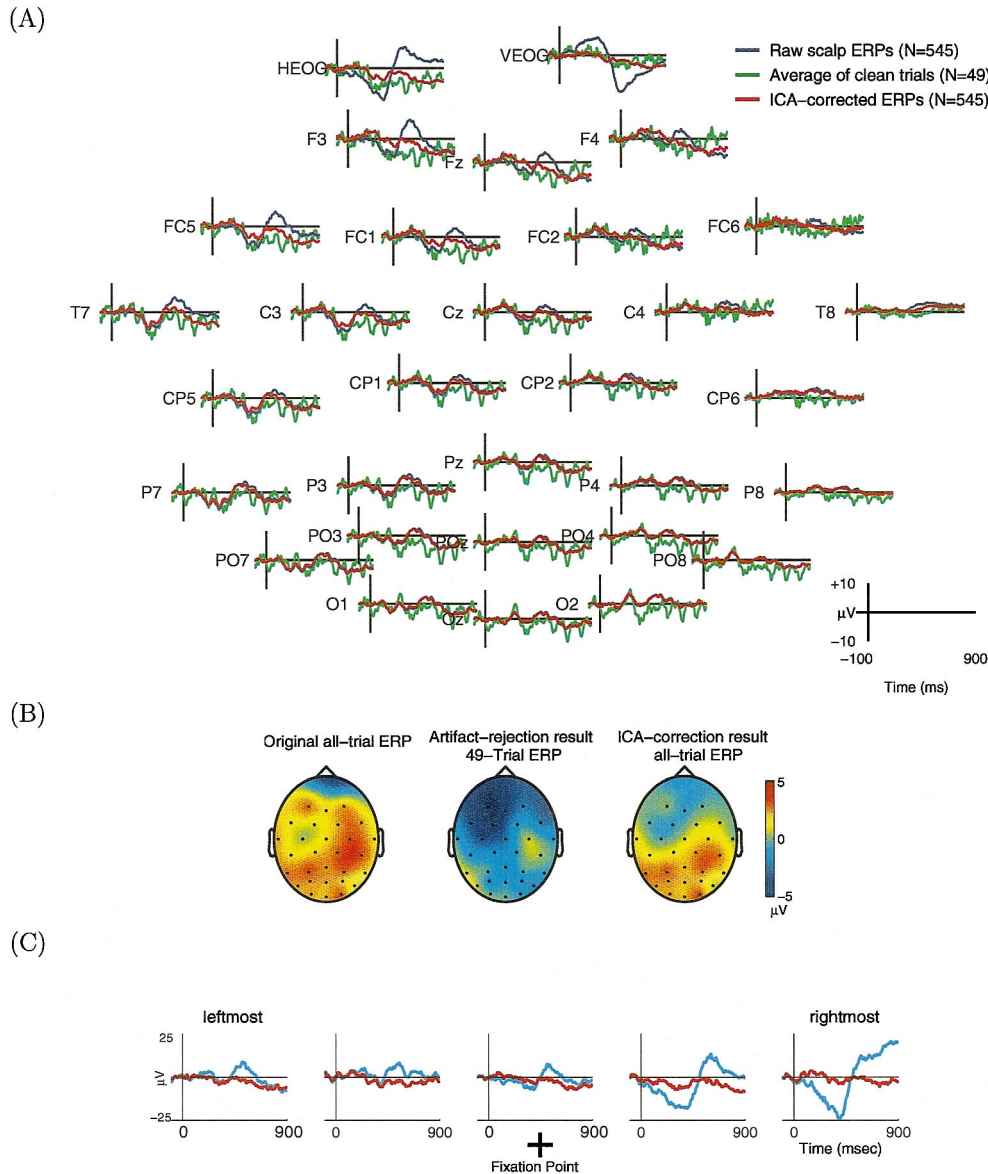


Fig. 4. ICA correction versus artifact rejection in a patient with right-frontal-temporal-parietal lesion. (A) Averages of responses to target stimuli before (blue traces) and after (red traces) ICA-based artifact correction. The (red traces) average of all 545 artifact-corrected trials strongly resembles the (green traces) average of 49 trials resulting from the conventional artifact rejection method, while avoiding the residual EEG ‘contamination’ remaining in the 49 trial averages because of insufficient cancellation of strong non-phase-locked alpha and other EEG processes. (B) Scalp topographies at the P300 peak (457 ms) of the (left) uncorrected average, (center) artifact-rejected average, and (right) ICA-corrected response average. (C) Averaged ERPs at site HEOG to targets presented at each of 5 attended locations, before (blue traces) and after (red traces) ICA-based artifact removal. Note that the subject was unable to inhibit her eye movements towards targets ipsilateral to the lesion. ICA-corrected averages (red traces) show no such lateralized differences.

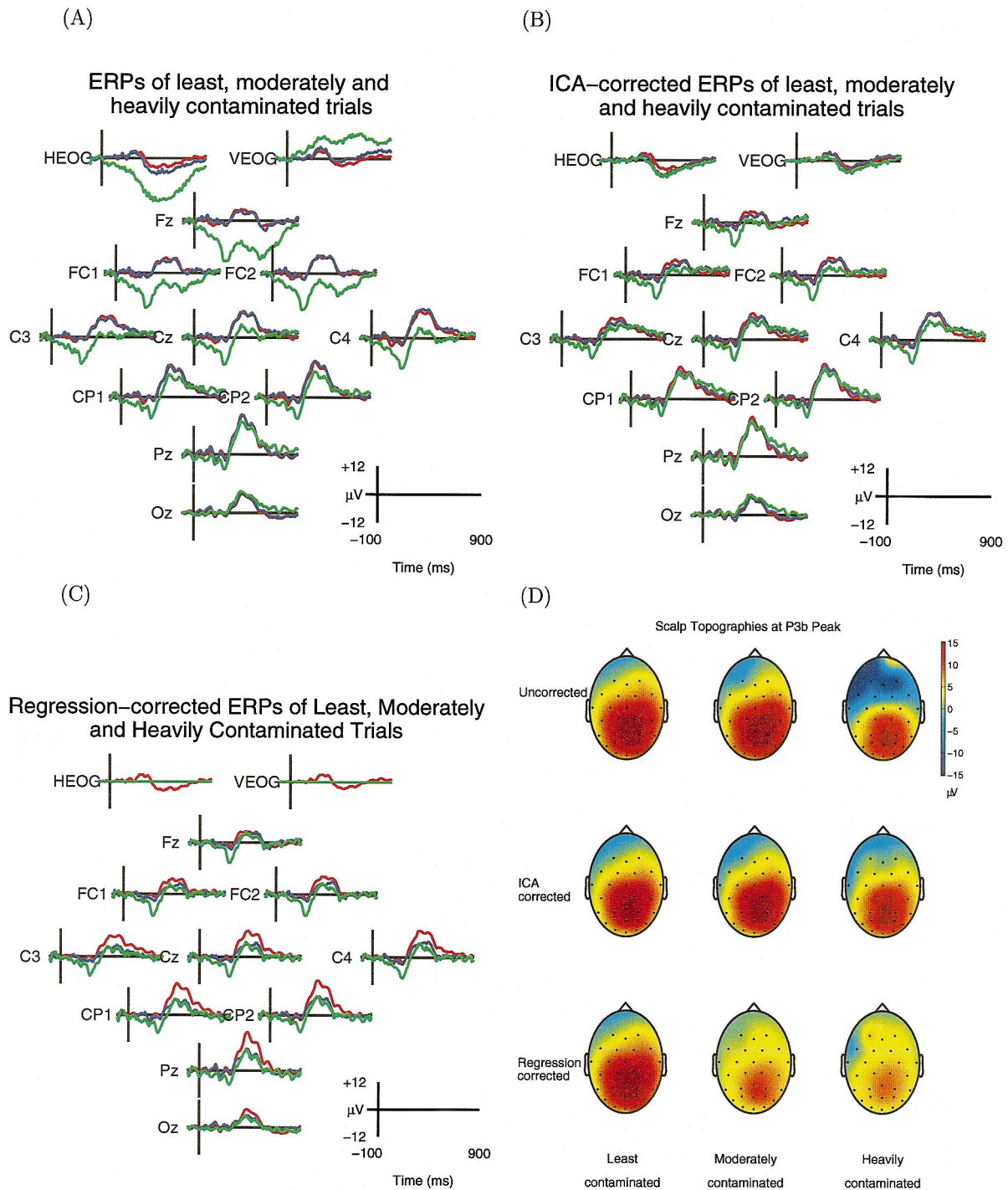


Fig. 5. ICA correction versus regression correction in removing artifacts from ERPs. (A) The raw averages of least (red traces, 397 trials), moderately (blue traces, 105 trials) and heavily (green traces, 64 trials) contaminated trials from a normal control subject before any artifact correction, (B) after ICA correction, and (C) after correction by multiple-lag regression. Note that the least contaminated trials were not subjected to regression here. (D) The scalp topographies at the P300 peak (423 ms) of least (left), moderately (middle) and heavily (right) contaminated trials before (top) and after (middle) artifact correction by ICA and regression (bottom).

sively larger eye movement artifacts in the uncorrected averaged responses to target stimuli presented to the right of fixation. The corrected ERP averages (red traces) show

that the ICA-based artifact removal identified and removed these mostly unilateral eye movement artifacts, revealing underlying neuronal phenomena consistent across stimulus

locations that would be difficult or impossible to measure accurately using other analysis methods.

3.4. ICA-based artifact removal versus regression

Example 4. This example uses a data set from a normal control to compare the use of ICA with multiple-lag regression to correct for eye movement artifacts. Again, we first viewed a histogram of the maximal potentials in each epoch at the two periocular channels and then split all single target response trials into 3 groups: those least ($\max|\mu V| \leq 25$), moderately ($25 < \max|\mu V| \leq 35$), and most heavily ($\max|\mu V| > 35$) contaminated by eye movements.

The regression method employed the recorded signals at sites VEOG and HEOG as reference channels. As recommended, regression was performed only for single-trial epochs containing maximal absolute potentials greater than 25 μV . Fig. 5A shows the averages of the 3 groups of trials before artifact removal. The averages of trials ‘corrected’ by ICA and by regression for the least (red, 397 trials), moderately (blue, 105 trials) and heavily (green, 64 trials) contaminated trial groups are shown in Fig. 5B,C, respectively. Fig. 5D shows the scalp topographies at the P300 peak (423 ms) for the 3 ERP averages before (top 3 panels) and after (middle 3 panels) artifact removal by ICA and by regression (bottom 3 panels). The 3 ICA-corrected averages had very similar waveforms and scalp topographies, while the regression-corrected averages had smaller potentials for the moderately (blue) and heavily (green) contaminated trials compared to the averages of the least contaminated (‘artifact-free’) trials (red traces). After removal of eye movement artifacts by regression (middle trace), the neural signals contributing to the reference signals (VEOG and HEOG) were also eliminated from the other frontal scalp sites. Also note that the averages of moderately and heavily contaminated trials contain no signals at sites VEOG and HEOG, whereas the ICA-corrected averages reveal the neural activity contributing to the recorded signals at these sites.

4. Discussion

The goal of the present study was to determine whether ICA could be used to remove artifacts of non-neural origin from single ERP data trials particularly in clinical subjects that are heavily contaminated with eye movement artifacts, thereby preserving the recorded event-related brain activity. Here, ICA was applied to single-trial target response records from a total of 50 (28 normal, 10 autistic and 12 brain lesion) subjects in a visual selective attention experiment. For each subject, ICA derived spatial filters or maps that decomposed EEG data recorded at multiple scalp sensors into a sum of components with fixed scalp distributions and maximally independent time courses. The spatial filters clearly separated eye blink and eye movement artifacts

into separate components with physiologically plausible scalp maps.

4.1. ICA correction versus artifact rejection

Rejecting contaminated EEG segments using threshold voltage criteria, the most commonly used method for dealing with artifacts in research settings, can be unacceptable when blinks and muscle movements occur too frequently. In contrast, ICA-based artifact removal can effectively detect and separate contaminations arising from a wide variety of artifactual sources in EEG records without losing neural signals recorded at frontal and even periocular sites. This method can preserve most or all of the recorded trials for analysis (Fig. 4, red traces), even when few or none of the single trials are artifact-free. In our example, ICA also recovered from the artifact-laden data of clinical subjects the same ‘P300’ scalp topography (Fig. 4B, right panel) generally obtained from normal control subjects. Before artifact removal, this pattern was overwhelmed by a strong artifactual negativity at frontal sites and was thus very difficult to identify either in the scalp map of the raw data averages or the regression-corrected averages (Fig. 4B, left and middle panels). Additionally, the result of ICA-based artifact removal illustrates the importance of having an effective artifact removal method when ERPs from clinical populations, whose data may be heavily contaminated by involuntary eye movements or muscle tension, are to be measured and interpreted.

4.2. ICA correction versus regression correction

The current alternative to artifact rejection is regression in the time or frequency domain. This is performed using conceptually distinct sets of EEG and EOG channel data to derive parameters characterizing the appearance and spread of EOG artifacts in the EEG channels. However, because EEG and ocular activity mix bidirectionally (Peters, 1967; Oster and Stern, 1980), regressing out eye artifacts inevitably involves subtracting relevant neural signals from each record as well as ocular activity. Regression methods become even more problematic when appropriate regressing channels are not available. For example, regressing out muscle artifacts using this method would require a reference channel selective for each contributing muscle. ICA-based methods, on the other hand, can be used to separate and remove multiple muscle artifacts as well as ocular artifacts, as we have shown elsewhere (Jung et al., 2000).

The present results (Fig. 5) showed that regression-based artifact removal of moderately and heavily contaminated trials consistently overcorrected and removed neural activity from electrodes located over frontal and periocular sites. ICA, on the other hand, recovered the EEG activity without relying on the availability of one or more ‘uncontaminated’ reference channels, even when applied to heavily contaminated trials. The resulting corrected response averages were

in each case remarkably similar to the averages of the least contaminated trials (Fig. 5).

4.3. ICA correction versus PCA correction

Compared to PCA that minimizes second-order correlations among the output channels, ICA imposes a much stronger criterion, statistical independence. Statistical independence requires that all second-order and higher-order correlations reduce to zero. PCA finds orthogonal directions of greatest variance in the data, whereas ICA component maps could be non-orthogonal. In general, there is no reason why neurobiologically distinct artifact and EEG sources should be spatially orthogonal to one another. Since ICA does not impose any condition on the spatial filters, it can collect concurrent activity arising from spatially overlapping artifact and EEG source distributions. We have found that ICA produces components whose scalp maps have only a few spatial maxima, consistent with relatively compact or simply connected source generators. In comparison, PCA produces components that mostly have more complex spatial patterns (Silberstein and Cadusch, 1992; Jung et al., 1998b, 2000).

Elsewhere, we have shown that ICA can effectively remove contamination from a wide variety of artifactual sources in spontaneous EEG data with results comparing favorably to those obtained using regression and PCA methods (Jung et al., 1999, 2000). We have also presented evidence for the behavioral consistency of ICA decomposition applied to ERP averages (Makeig et al., 1999a,b).

4.4. ICA limitations

Although ICA appears to be quite useful for EEG analysis, it is important to keep in mind that it also has some inherent limitations.

First, infomax ICA can decompose at most N sources from N scalp electrodes. Usually, the effective number of temporally-independent signals contributing to the scalp EEG is unknown, and it is likely that observed brain activity arises from more physically separable effective sources than the available number of EEG electrodes. We have explored and discussed this issue elsewhere (Makeig et al., 2000a).

Second, the assumption of temporal independence used by ICA cannot be satisfied when the training data set is too small, or when separate topographically distinguishable phenomena always occur concurrently in the data. In the latter case, simulations show that ICA may derive a component accounting for their joint occurrence, plus separate components accounting for their periods of solo activation. Such confounds imply that converging behavioral or other evidence must be obtained before concluding that spatio-temporally overlapping ICA components measure neurophysiologically or functionally distinct activities.

Third, ICA assumes that the physical sources of artifactual and neural activity contributing to EEG signals are spatially stationary through time. In general, there is no

reason to believe that cerebral and artifactual sources in the spontaneous EEG necessarily remain fixed over time or occurrences. Examples of non-fixed sources may include spreading sleep spindles (McKeown et al., 1998). However, in our studies of averaged and unaveraged data from normal control subjects in these experiments (Jung et al., 1999), the relatively small numbers of obtained components showing stimulus-locked, response-locked, and non-phase-locked categories, each accounting for activity occurring across sets of 500 or more 1 s trials, suggests that the brain areas generating our data were primarily fixed. This supposition is concordant with repeated observations in functional brain imaging experiments that discrete, spatially restricted areas of cortex are activated during task performance (Friston et al., 1998), and has been further verified for these data in a computationally intense moving-average ICA study (Makeig et al., 2000b).

4.5. Advantages of ICA-based artifact removal

Our results show that ICA has at least 3 advantages compared with other artifact removal methods. (1) ICA simultaneously separates EEG signals including artifacts into independent components based on the characteristics of the data, without relying on the availability of one or more ‘clean’ reference channels for each type of artifact. This avoids the problem of mutual contamination between regressing and regressed channels. (2) ICA-based artifact removal can preserve all of the recorded trials, a crucial advantage over rejection-based methods when limited data are available, or when blinks and muscle movements occur too frequently, as in some patient groups. (3) Unlike regression methods, ICA-based artifact removal can preserve data at all scalp channels, including frontal and periocular sites.

ICA uses spatial filtering to achieve artifact removal; the filters derived from a brief EEG recording portion (e.g. 5–10 s) might even be applied to multichannel EEG by a simple matrix multiplication to successive EEG epochs to create an artifact-reduced EEG in real time. More exhaustive ICA decomposition may be needed only when the spatial filters fail to effectively remove artifacts with very different spatial distributions. We thus believe it may be possible to perform artifact removal in routine clinical EEG in near real time using an appropriate hardware and software implementation.

In addition to artifact removal, ICA decomposition can be highly useful for enhancing the amount and quality of information in event- or response-related brain signals that can be extracted from ERP data (Makeig et al., 1997, 1999a), and examining systematic variations from trial to trial within subjects (see Jung et al., 1999). The analysis of single-trial data is particularly important in studies of orienting, habituation, or associative learning. However, most or all of the information available in single trials is usually sacrificed to increase signal-to-noise ratio through within- or between-subject averaging (Kenemans et al., 1989).

Removing EEG artifacts from brain activity by ICA can increase the amount of behaviorally and neurophysiologically relevant information available in ERP data, and may also give researchers the ability to examine trial to trial response variations. Similar ICA approaches should be equally applicable to other types of multichannel biomedical data for which linear superposition of the contributing signals can be assumed (e.g. electromyography, magnetoencephalography, electrocardiography, etc.).

Acknowledgements

This report was supported in part by a grant from the Swartz Foundation, and grants from the Office of Naval Research ONR.Reimb.30020.6429 (S.M.), from the Howard Hughes Medical Institute (T.J.S.), and from the National Institute of Health NIMH 1-R01-NH-36840 and NINDS 1R01-NS34155-01 (E.C. and J.T.). The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, or the US Government.

References

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 3rd ed. revised. Washington, DC: American Psychiatric Association, 1987.
- Bell AJ, Sejnowski TJ. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation* 1995;7:1129–1159.
- Berg P, Scherg M. Dipole models of eye activity and its application to the removal of eye artifacts from the EEG and MEG. *Clin Physiol Meas* 1991;12(Suppl A):49–54.
- Berg P, Scherg M. A multiple source approach to the correction of eye artifacts. *Electroenceph clin Neurophysiol* 1994;90(3):229–241.
- Comon P. Independent component analysis – a new concept? *Signal Processing* 1994;36(3):287–314.
- Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R. Event-related fMRI: characterizing differential responses. *Neuroimage* 1998;7(1):30–40.
- Girolami M. An alternative perspective on adaptive independent component analysis algorithms. *Neural Computation* 1998;10(8):2103–2114.
- Hécaen H. Clinical symptomatology in right and left hemisphere lesions. In: Mountcastle VB, editor. *Interhemispheric relation and cerebral dominance*, Baltimore, MD: John Hopkins University Press, 1962.
- Hillyard SA, Galambos R. Eye-movement artifact in the CNV. *Electroenceph clin Neurophysiol* 1970;28(2):173–182.
- Jung T-P, Humphries C, Lee T-W, Makeig S, McKeown MJ, Iragui V, Sejnowski TJ. Extended ICA removes artifacts from electroencephalographic data. *Adv Neural Info Processing Systems* 1998a;10:894–900.
- Jung T-P, Humphries C, Lee T-W, Makeig S, McKeown MJ, Iragui V, Sejnowski TJ. Removing electroencephalographic artifacts: comparison between ICA and PCA. *Neural Networks Signal Processing* 1998b;VIII:63–72.
- Jung T-P, Makeig S, Bell AJ, Sejnowski TJ. Independent component analysis of electroencephalographic and event-related potential data. In: Poon P, Brugge J, editors. *Auditory processing and neural modeling*, New York: Plenum Press, 1998c. pp. 189–197.
- Jung T-P, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ. Analyzing and visualizing single-trial event-related potentials. *Adv Neural Info Processing Systems* 1999;11:118–124.
- Jung T-P, Makeig S, Humphries C, Lee T-W, McKeown MJ, Iragui V, Sejnowski TJ. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 2000;37:163–178.
- Kenemans JL, Verbaten MN, Roelofs JW, Slangen JL. “Initial-” and “change-orienting reactions”: an analysis based on visual single-trial event-related potentials. *Biol Psychol* 1989;28(3):199–226.
- Kenemans JL, Molenaar P, Verbaten MN, Slangen JL. Removal of the ocular artifact from the EEG: a comparison of time and frequency domain methods with simulated and real data. *Psychophysiology* 1991;28(1):114–121.
- Ládavas E, Zeloni G, Zaccara G, Gangemi P. Eye movements and orienting of attention in patients with visual neglect. *J Cog Neurosci* 1997;9(1):67–74.
- Lagerlund TD, Sharbrough FW, Busacker NE. Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition. *J Clin Neurophysiol* 1997;14(1):73–82.
- Lee T-W, Girolami M, Sejnowski TJ. Independent component analysis using an extended infomax algorithm for mixed sub-Gaussian and super-Gaussian sources. *Neural Computation* 1999;11(2):606–633.
- Lins OG, Picton TW, Berg P, Scherg M. Ocular artifacts in recording EEG and event-related potentials II: source dipoles and source components. *Brain Topogr* 1993;6(1):65–78.
- Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Aut Dev Disord* 1994;24(5):659–685.
- Makeig S, Bell AJ, Jung T-P, Sejnowski TJ. Independent component analysis of electroencephalographic data. *Adv Neural Info Processing Systems* 1996;8:145–151.
- Makeig S, Jung T-P, Bell AJ, Ghahremani D, Sejnowski TJ. Blind separation of event-related brain responses into independent components. *Proc Natl Acad Sci USA* 1997;94:10979–10984.
- Makeig S, Westerfield M, Jung T-P, Covington J, Townsend J, Sejnowski TJ, Courchesne E. Independent components of the late positive response complex in a visual spatial attention task. *J Neurosci* 1999a;19(7):2665–2680.
- Makeig S, Westerfield M, Townsend J, Jung T-P, Courchesne E, Sejnowski TJ. Independent components of early event-related potentials in a visual spatial attention task. *Phil Trans Biol Sci* 1999b;354:1135–1144.
- Makeig S, Jung T-P, Ghahremani D, Sejnowski TJ. Independent component analysis of simulated ERP data. In: Nakada T, editor. *Human higher function I: advanced methodologies*, 2000a, in press.
- Makeig S, Enghoff S, Jung T-P, Sejnowski TJ. Moving-window ICA decomposition of EEG data reveals event-related changes in oscillatory brain activity. *The 2nd International Workshop on Independent Component Analysis and Signal Separation*, 2000b, in press.
- McKeown MJ, Humphries C, Achermann P, Borbely AA, Sejnowski TJ. A new method for detecting state changes in the EEG: exploratory application to sleep data. *J Sleep Res* 1998;7(Suppl 1):48–56.
- Oster PJ, Stern JA. Measurement of eye movement electrooculography. In: *Matin I, Venables PH, editors. Techniques in psychophysiology*, Chichester: Wiley, 1980. pp. 275–309.
- Peters JF. Surface electrical fields generated by eye movement and eye blink potentials over the scalp. *J EEG Technol* 1967;7:27–40.
- Schopler E, Reichler RJ, De Velis RF, Daly K. Toward objective classification of childhood autism: childhood autism rating scale (CARS). *J Aut Dev Disord* 1980;10:91–103.
- Silberstein RB, Cadusch PJ. Measurement processes and spatial principal components analysis. *Brain Topogr* 1992;4(4):267–276.
- Small JG. Sensory evoked responses of autistic children. *Infantile Autism* 1971:224–239.
- Townsend J, Courchesne E. Parietal damage and narrow “spotlight” spatial attention. *J Cog Neurosci* 1994;6:220–232.

- Verleger R, Gasser T, Möcks J. Correction of EOG artifacts in event-related potentials of EEG: aspects of reliability and validity. *Psychophysiology* 1982;19(4):472–480.
- Vigário RN. Extraction of ocular artifacts from EEG using independent component analysis. *Electroenceph clin Neurophysiol* 1997;103:395–404.
- Whitton JL, Lue F, Moldofsky H. A spectral method for removing eye-movement artifacts from the EEG. *Electroenceph clin Neurophysiol* 1978;44:735–741.
- Woestenburg JC, Verbaten MN, Slangen JL. The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain. *Biol Psychol* 1983;16:127–147.