Psychophysiology, 47 (2010), 1142–1150. Wiley Periodicals, Inc. Printed in the USA.
Copyright © 2010 Society for Psychophysiological Research
DOI: 10.1111/j.1469-8986.2010.01015.x

Validation of ICA as a tool to remove eye movement artifacts from EEG/ERP

MAARTEN MENNES,[a] HEIDI WOUTERS,[a] BART VANRUMSTE,[b,c] LIEVEN LAGAE,[a] AND PETER STIERS[a,d]

[a]Department of Woman & Child, Section Paediatric Neurology, K.U. Leuven, Leuven, Belgium
[b]SCD/ESAT, K.U. Leuven, Leuven, Belgium
[c]MOBILAB, Katholieke Hogeschool Kempen, Geel, Belgium
[d]Department of Neuropsychology and Psychopharmacology, Maastricht University, Maastricht, The Netherlands

Abstract

Eye movement artifacts in electroencephalogram (EEG) recordings can greatly distort grand mean event-related potential (ERP) waveforms. Different techniques have been suggested to remove these artifacts prior to ERP analysis. Independent component analysis (ICA) is suggested as an alternative method to "filter" eye movement artifacts out of the EEG, preserving the brain activity of interest and preserving all trials. However, the identification of artifact components is not always straightforward. Here, we compared eye movement artifact removal by ICA compiled on 10 s of EEG, on eye movement epochs, or on the complete EEG recording to the removal of eye movement artifacts by rejecting trials or by the Gratton and Coles method. ICA performed as well as the Gratton and Coles method. By selecting only eye movement epochs for ICA compilation, we were able to facilitate the identification of components representing eye movement artifacts.

Descriptors: EEG/ERP, Artifact removal, Validation, Blink artifact, Independent component analysis

The increasing popularity of the electroencephalogram (EEG) and event-related potentials (ERP) for research and clinical purposes has resulted in better acquisition devices minimizing the amount of noise picked up during an EEG recording. However, an important source of noise that cannot be avoided by better amplifiers or electromagnetically shielded rooms is the electrical activity that is associated with eye movements. Among eye movements, blinks cause the largest distortions, mainly because of the movement of the eyelids across the surface of the eyes, but also saccades can cause large distortions in EEG signals and ERP waveforms, particularly at frontal electrodes (Iwasaki et al., 2005).

Because these artifacts can hamper correct interpretation, it is best to remove them from the data. In most ERP studies this is done by rejecting trials containing these artifacts. Such trials can be detected by setting an absolute amplitude threshold. However, the amount of data lost by rejecting trials containing eye movements can be unacceptably high, especially when one is working with clinical populations or children, for whom it is difficult to refrain from blinking. A method that allows the cleaning up of the contaminated trials by filtering or removing the eye movement artifacts from the data would, therefore, be highly beneficial in terms of research effort efficiency. Consequently, several methods that try to remove as much of the eye movement artifacts without reducing data quality have been introduced (Berg & Scherg, 1994; Croft & Barry, 2000; Gratton, Coles, & Donchin, 1983).

More recently, independent component analysis (ICA) was introduced as a method for separating artifactual data (blinks, eye movements, and muscle and line noise) from useful brain activity in EEG recordings (Flexer, Bauer, Pripfl, & Dorffner, 2005; Ford, Sands, & Lew, 2004; Frank & Frishkoff, 2007; Jung et al., 2000a, 2000b; Vigario, 1997). ICA is a data-driven blind source separation method that is applied on biomedical data such as EEG, ERP, magnetoencephalography, and functional magnetic resonance imaging (fMRI; Bell & Sejnowski, 1995; Stone, 2002; Vigario, Sarela, Jousmaki, Hamalainen, & Oja, 2000). It is used to reduce large sets of research data to a small number of independent components in order to ease interpretation or to locate possible brain sources for a measured signal (Esposito et al., 2003; Iidaka, Matsumoto, Nogawa, Yamamoto, & Sadato, 2006; Kansaku et al., 2005; Makeig et al., 2004; Zeki, Perry, & Bartels, 2003).

It appears however, that only few research groups effectively apply ICA or other blind source separation algorithms (e.g., SOBI, JADE) to remove eye movement artifacts from their EEG

The authors thank Maarten De Vos, Nikolai Novitski, Jennifer Ramautar, Stefan Sunaert, Bea Van den Bergh, Katrien Vanderperren, Sabine Van Huffel, Anneleen Vergult, and Johan Wagemans for helpful discussions. This work is supported by grants from Fonds voor Wetenschappelijk Onderzoek, Vlaanderen (# G.0211.03 and #7.0008.03) and K.U. Leuven (IDO/05/010 EEG-fMRI).

Address reprint request to: Maarten Mennes, Phyllis Green and Randolph Cowen Institute for Pediatric Neuroscience, NYU Child Study Center, 215 Lexington Ave, 14th Floor, New York, NY 10016, USA. E-mail: maarten.mennes@nyumc.org

PSYCHOPHYSIOLOGY

data (Debener, Hine, Bleeck, & Eyles, 2008; Huang, Jung, Delorme, & Makeig, 2008; Jung et al., 2001; Joyce, Gorodnitsky, & Kutas, 2004; Slagter et al., 2007). We believe that the reason for this is, in some respect, inherent to ICA. Because it is a blind source decomposition method, the resulting independent components are statistically valid, but their physiological meaning is not always clear. It is up to the user to decide which components represent artifactual data and which brain activity of interest. However, ICA can be applied in several ways, which differ in the selection of data from the continuous data set entered into the decomposition analysis (Ille, Berg, & Scherg, 2002). Because this affects the salience of components representing the targeted artifact, the appropriate selection of data used for ICA provides a way to improve component selection and interpretation and, hence, the quality of artifact removal. In addition, entering only a part of the data into the ICA analysis can considerably reduce the computation time, thus compensating one of the drawbacks when one uses ICA to remove artifacts from continuous EEG data sets. Especially when one uses a large number of electrodes, the computation time and computer memory required to run ICA on a complete EEG data set could exceed commonly available resources. Short computation times would also allow ICA to be used for near real-time correction of EEG data.

Here, we introduce a procedure for easily identifying eye movement artifacts among independent components. By applying ICA only on sections of data that were contaminated with eye movement artifacts (i.e., epochs selected around the eye movements; Verleger, Gasser, & Möcks, 1982), we anticipated capturing the eye movement artifacts in easily identifiable components, leading to straightforward criteria to select components to remove, making quantitative selection of components viable. We compared the results of this procedure to the application of ICA on 10 s of EEG and on the complete continuous EEG data. In all these procedures, the resulting components were subsequently used to remove the eye movement artifacts from the continuous data. To underline the validity of our component selection procedure and the usefulness of ICA in general for the removal of eye movement artifacts, we compared the results of the eye artifact removal by ICA to (a) data uncorrected for eye movements, (b) data corrected for eye movements by rejecting EEG epochs with an amplitude exceeding $\pm$ 100 μV, and (c) to the results of eye artifact removal by the well-established Gratton and Coles method (Gratton et al., 1983). Various aspects of the performance of these procedures were evaluated. (1) We visually evaluated the cleaned continuous data. (2) Heavily contaminated trials were compared to trials with only little contamination by eye movements. (3) The impact of the artifact removal on trials with little contamination was assessed. (4) The variation across trials after removal of the artifacts was measured. (5) The statistical sensitivity of the data to experimentally induced effects was evaluated.

Methods

Participants

Data are presented for 6 right-handed adults (2 female; age: 24–34 years). None had a history of substance abuse or psychiatric or neurological disorder. All participants had normal or corrected-to-normal vision. Data were collected in the course of ongoing studies within the laboratory. The study was approved by the local ethics committee, and all participants gave their informed consent.

EEG Recordings

Nineteen electrodes were applied using the standard 10–20 system of electrode placement: Fp1, Fp2, F3, F4, F7, F8, C3, C4, T3, T4, P3, P4, T5, T6, O1, O2, Fz, Cz, and Pz. A ground electrode was placed on the forehead above the nose. Additionally two electrodes were placed on the outer canthi (horizontal electrooculogram [HEOG]) and two above and below the right eye (vertical electrooculogram [VEOG] = below eye-above eye) to detect horizontal and vertical eye movements. All electrodes were referenced to linked left and right mastoids, and all electrode impedances were below 5 kΩ at the start of the recording session. Sampling rate was 1000 Hz, with an analog bandpass of 0.095–70 Hz. Prior to removal of the eye movement artifacts, data were filtered off-line using a 30-Hz digital low-pass filter. Off-line analysis of the data, including removal of the eye artifacts, was performed using the EEGLAB v4.5 toolbox (Delorme & Makeig, 2004) under Matlab v7.0 (Mathworks, Natick, MA), except for the removal of the eye movement artifacts using the Gratton and Coles method, which was done with Brain Vision Analyzer 2 (Brain Products, Germany). Recently, a plug-in implementing the Gratton and Coles method also became available for EEGLAB (http://pinguin.uni-psych.gwdg.de/~ihrke/wiki/index.php/Ocular_Correction_EEGlab_Plugin).

Cued Attention Paradigm

EEGs were recorded during a cued attention paradigm based on paradigms of Corbetta, Kincade, and Shulman (2002) and Posner (1980). Participants were asked to respond to a total of 150 stimuli appearing in the left or right visual field and were either unaware (no cue, 20%) or informed (cue, 80%) about the location at which the stimulus would appear. In 20% of all cued trials the cue was invalid, that is, the stimulus appeared on the side opposite to the one indicated by the cue. For ERP analysis, task-relevant epochs were extracted from 150 ms before to 800 ms after target stimulus onset. During the task, central fixation had to be maintained, and participants were encouraged to make as few eye movements as possible, including blinks.

Procedures for Eye Movement Artifact Removal

The correction for eye movement artifacts was done for each subject separately. An individual approach to eye movement artifacts is needed, as, for instance, differences in blink artifacts exist between participants because of differences in eyelid size and electrode position (Iwasaki et al., 2005). Figure 1 gives an overview of the six procedures to remove or correct for eye movement artifacts used in the current study. First, data were used in their "raw" format, with only the 30-Hz low-pass filter applied (RAW). Second, task-specific epochs were extracted from the continuous raw data. Subsequently epochs in which the maximum amplitude exceeded $\pm$ 100 μV were rejected for further analysis (REJECT). Third, the Gratton and Coles ocular correction algorithm as implemented in Brain Vision Analyzer 2 was applied to the raw continuous data (GC). In this algorithm event-related potentials of interest are first subtracted from the raw EEG and EOG signals. Afterward, the proportion of EOG that is represented in each of the EEG channels is calculated first for blinks and subsequently for saccades and subtracted from the EEG recording. Finally, the subtracted event-related potentials are again added to the recording. Details can be found in Gratton et al. (1983).

In procedures four, five, and six, we used independent component analysis to remove eye movement artifacts. We used the

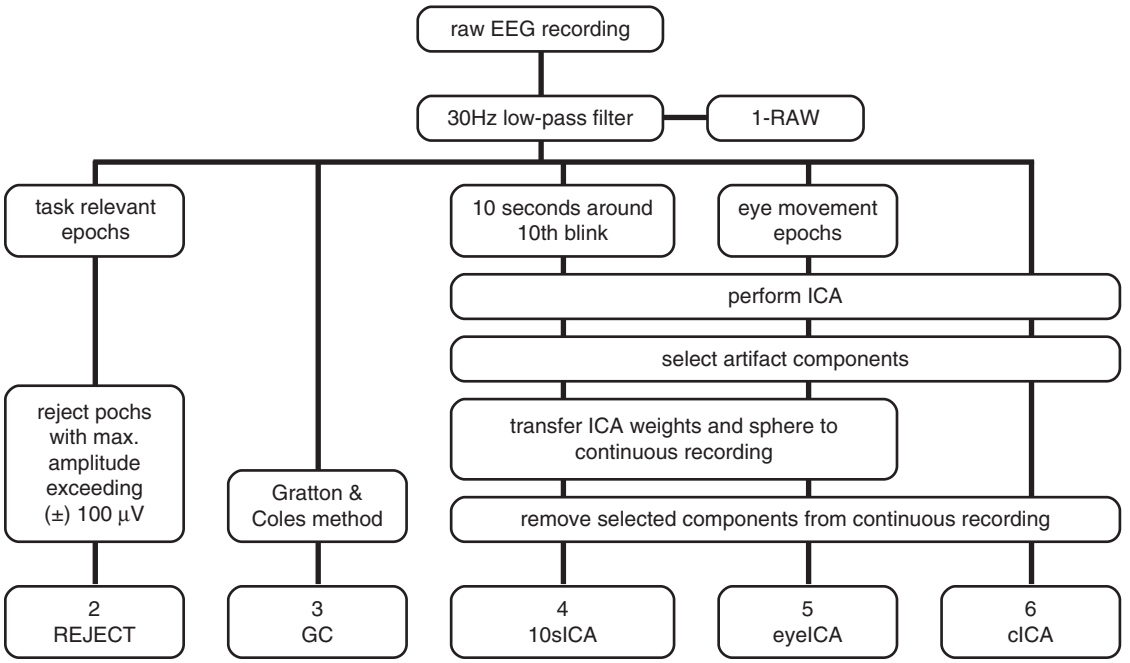

Figure 1. Diagram of the procedures used to remove eye movement artifacts from the EEG recording.

standard runICA command integrated in EEGLAB, which uses the logistic infomax ICA algorithm of Bell and Sejnowski (1995). Before applying ICA for each individual subject, we first marked all eye movements in the EEG using a simple detection algorithm. We used the fast and extreme rise in amplitude typical for the representation of the different eye movements in the VEOG and HEOG channels compared to normal EEG signals to mark the occurrence of all individual eye movements (left, right, blink, down). Vertical eye movements were marked by searching VEOG. These included blinks, as blink artifacts had similar, although more extreme, characteristics as an upward directed eye movement. Left and right directed eye movements were marked by searching HEOG. Other EEG artifacts might also exhibit a fast and extreme rise in amplitude. Such artifacts were also marked as eye movements by our algorithm, but were easily identifiable as non-eye movements and removed from the final selection of eye movements upon visual inspection of the markers placed by our algorithm. For procedure four, we selected a 10 s long EEG epoch (Jung et al., 1998a, 1998b, 2000a) around the 10th blink present in the raw data. After performing ICA on this 10 s epoch and selecting components representing eye movement artifacts (see below), we transferred the calculated ICA weights and sphere to the raw data and removed the selected components by subtracting the projection of the artifactual components from the original data of the subject. The fifth procedure applied ICA on all eye movement artifacts. For each individual subject, the continuous EEG recording was epoched based on the marked occurrences of eye movements. Epochs were extracted from 500 ms before onset of an eye movement to 600 ms thereafter. The length of these epochs was chosen in order to include the whole eye movement plus a small portion of non-eye-movement data. This portion was included to provide ICA with data from sources other than the highly prominent eye movement artifact sources. Preliminary analyses indicated that including more then 500 ms of non-eye-movement artifact data did not result in an improved separation of the artifactual data. Using less than 500 ms resulted in a less optimal separation of artifacts and cognitive sources in separate components. Subsequently, ICA was applied to these eye movement epochs. The mean total length of EEG data provided for ICA was 2.5 min, with a minimum of 1.1 min. The maximum of 4.2 min was still a lot shorter than the mean complete raw EEG (7 min). After the selection of components to remove, the ICA weights and sphere were transferred to the complete raw data and the selected components were removed (eyeICA). In the sixth and final procedure, we applied ICA to the complete raw data, subsequently subtracting components that were identified as representing eye movement artifacts (cICA). The mean length of the complete raw data was 7 min (minimum = 6.9 min, maximum = 7.2 min).

Selection of Eye Artifact Components from the ICA

For each ICA method we used the same procedure to identify components that represented eye movement artifacts. First, we selected components that each explained at least 15% of the variance of the EEG signal in VEOG and whose summed explained variance in that channel exceeded 90%. Usually one or two components were selected. In addition, we selected all other components that individually explained more than 15% of the variance of the EEG signal in HEOG and together with the components already selected from VEOG explained over 90% of the variance in HEOG. In most cases one additional component was selected, resulting in a total of two or three components that were removed from the data. We first looked at VEOG because components explaining most variance in this channel were easy to recognize, as they included the easily identifiable eyeblinks. However, eyeblinks are also registered in HEOG and, although registered to a much lesser extent, they sometimes cause crosstalk between components. The example below will clarify this issue and the selection procedure. For one of the participants the variance explained by the first four components was for the VEOG channel 82.4%, 16.5%, 0.2%, and 0.1% and for the HEOG channel 29.6%, 14.6%, 0.2%, and 53.9%. The first two

components reflected eyeblinks picked up by both VEOG and HEOG. Following the above selection procedure, we first selected Components 1 and 2, which together explained 82.4%+16.5% = 98.9% of the variance in VEOG. Second, Component 4 was selected based on the explained variance in HEOG. Together with Components 1 and 2, Component 4 explained 98.03% of the variance in HEOG (29.59%+14.57% +53.87%). Thus, in this example, Components 1, 2, and 4 were removed from the continuous EEG data.

The identification of eye movements, the application of ICA to the data, and the selection of components to remove were all integrated in a Matlab script. The only user input needed consisted of the evaluation and confirmation of the ICA components that were identified for removal. User confirmation of the artifactual ICA components was preserved to increase control over the correctness of the proposed solution. Scripts are available on request.

Validation Procedures

The quality of the removal of the eye movement artifacts was evaluated in five procedures as indicated in Table 1. First (Comparison 1, Table 1), grand average event-related potentials of the eye movement artifacts were calculated and compared across removal methods. The inspection of the raw average is suggested as a valid, initial form of evaluation (Croft & Barry, 2000; Verleger et al., 1982). The rejection method could not be included in this comparison.

For the next two comparisons, we used a method previously described by Jung et al. (2000b). Using each trial's absolute maximum potential values at the two EOG channels as an indication of the amount of contamination with eye movement artifacts, all validly cued trials of the administered cued attention paradigm were categorized as least, moderately, and most contaminated with eye movement artifacts. The categorization was based on the local maxima in a histogram of the maximal amplitude in all trials. The threshold separating the least and moderately contaminated trials ranged between 20 and 35 μV, and the threshold for separating the moderately and most contaminated trials ranged between 55 and 100 μV. Based on this method, two comparisons were made. In Comparison 2 we tested how efficient the removal procedures were by statistically comparing the least and most contaminated trials after removal of the eye movement artifacts according to the different removal procedures. Upon adequate removal of the eye movement artifacts, the grand average of the trials labeled as least and most contaminated should not differ statistically. In Comparison 3, we tested whether the removal procedures affected data without eye movements, because the correction methods were applied to the complete raw EEG recording, including the least contaminated trials. We statistically compared the least contaminated trials before removal of the eye movement artifacts with the least contaminated trials after removal of these artifacts. It is assumed that the least contaminated trials contain almost no eye movements. Therefore, they should be barely affected by the different artifact removal methods. Differences in this comparison could indicate overcorrection of the data. For these two comparisons, all removal procedures were included except the rejection method, as this method resulted in the rejection of all most contaminated trials.

The last two evaluations concerned the ERP characteristics of the data before and after removal of the eye movement artifacts. The first assessed the reduction of eye movement artifact-related variance in the data. By removing the eye movement artifacts, the standard deviation of the task-relevant epochs should decrease, because a large portion of noise is removed. This evaluation does not consider whether the different removal procedures might have distorted the underlying neural potentials (this is evaluated in Comparison 3). Here, we compared the standard deviation for each removal procedure calculated across all validly cued trials and all time points within the first second (1000 ms) after stimulus presentation (Comparison 4). A decrease in standard deviation while the number of trials remains constant should increase the statistical power of an analysis. This consideration was assessed in the final evaluation (Comparison 5) by statistically comparing the validly and invalidly cued trials of the cued attention paradigm after eye movement artifact removal. Comparisons 4 and 5 allowed including the REJECT procedure next to all other procedures for removing the eye movement artifacts.

Table 1. *Procedures Used to Evaluate the Performance of the Different Eye Artifact Removal Methods*

Evaluation procedure	Inspection
1. Grand average of eye movements	Visual
2. Least contaminated trials *after* removal of eye movement artifacts versus most contaminated trials *after* removal of eye movement artifacts	Statistical (SPM)
3. Least contaminated trials *before* removal of eye movement artifacts versus least contaminated trials *after* removal of eye movement artifacts	Statistical (SPM)
4. Standard deviation across 0–1000 ms after stimulus presentation	Statistical (ANOVA)
5. Valid cue trials versus invalid cue trials	Statistical (SPM)

Note: In Steps 1–4, only valid trials from the cued attention paradigm were included. In Steps 1, 4, and 5, all removal procedures were compared. The REJECT procedure could not be included in Steps 2 and 3.

Statistical Analysis

All statistical comparisons for the ERPs were made using statistical parametrical mapping (SPM), which is widely used in fMRI research. The benefit of this procedure compared to traditional statistics used in ERP analysis is that it allows for a simultaneous statistical comparison across all time points included in the EEG epochs and across all channels included in the EEG recording. This avoids the subjective selection of time points and electrodes for statistical analysis. Instead, the SPM analysis objectively indicates *where* and *when* significant effects occur. SPM is a mass univariate approach in which spatiotemporal neuroimaging data are modeled within the statistical framework of the general linear model. The software package used in this study was SPM5 (Wellcome Department of Cognitive Neurology, University College, London, UK). EEG epochs were entered into the analysis as space–time volumes with the anterior–posterior and left–right dimensions of each electrode's position as a two-dimensional spatial array. Time was entered as a third dimension. This implied that each data point (i.e., the EEG signal measured at a particular millisecond at a particular electrode) was entered into the analysis while preserving its spatiotemporal relationship to other data points. Taking into account the three-dimensional structure of the data justifies calling these data points "voxels." As such, voxels were defined containing the amplitude information at one electrode (with X and Y dimensions) at 1 ms (Z dimension).

Because the ERP data were low-pass filtered during preprocessing, no additional filtering was performed in the SPM analyses to control for possible serial autocorrelations in the data. Likewise, no global scaling was performed, because baseline corrections were applied in the preprocessing of the ERP epochs.

A fixed effects design was used in which conditions were modeled separately for each subject. Sessions per subject, on the other hand, were not modeled separately. For the least versus most contaminated comparison (#2) a separate analysis was done for each removal method (RAW, GC, 10sICA, eyeICA, cICA). Statistical differences between the least and most contaminated trials were assessed with an F contrast. For the comparison between the least contaminated trials without correction for eye movement artifacts and the least contaminated trials after removal of the eye artifacts (Comparison 3) all removal methods were modeled in the same SPM analysis. Four F contrasts were defined to assess the difference between the least contaminated uncorrected and least contaminated trials corrected according to method X (i.e., GC, 10sICA, eyeICA, or cICA). Finally, the cued-attention paradigm valid versus invalid comparison (#5) was assessed using a t contrast for valid < invalid in a separate fixed effects SPM analysis for each removal method (RAW, REJECT, GC, 10sICA, eyeICA, cICA).

Differences between conditions were evaluated at a significant alpha level of .05, corrected for multiple comparisons based on the random field theory (Friston, Firth, Liddle, & Frackowiak, 1991; Worsley, Evans, Marrett, & Neelin, 1992). Because we expected no differences in the comparisons involving the contaminated conditions, the significance level was in a second step lowered to $p < .001$ uncorrected for multiple comparisons to show that even at this liberal threshold no significant voxels were found. Clusters should contain at least 20 voxels to be regarded as significant.

Results

The amount of eye movements registered in each subject during the total duration of the cued attention paradigm ranged from 59 to 229 (mean = 137). Only 10% of all eye movements were horizontal. The number of trials categorized as least contaminated ranged between 14 and 45 (mean = 30), compared to 19 and 69 (mean = 35) for the most contaminated trials.

Grand Average of Eye Movements

Figure 2a shows the ERP of the blinks using the eye movements as marked by our algorithm (see above). It is clear that the different removal procedures had a different impact on the blink-related ERP measured at the frontal channels. The smaller blink-related ERP measured for the eyeICA and cICA methods suggests better correction compared to the GC and 10sICA methods. The impact of correction on horizontal eye movements was less clear (Figure 2b). In addition, the impact of horizontal eye movements on the EEG channel was limited, as evident from the RAW horizontal eye movement ERP, especially when compared to the impact of blinks on the frontal channels.

Least Contaminated versus Most Contaminated Trials after Eye Artifact Removal

ERPs synchronized on the presentation of validly cued trials and subsequently categorized as least and most contaminated by eye movements are presented in Figure 3, before (Figure 3a,b) and after removal of the eye movement artifacts with the various procedures (Figure 3c–f). It is evident from Figure 3 that obvious differences that existed at the frontal electrodes before removal of the eye movement artifacts between the trials categorized as least and most contaminated have disappeared after applying the removal methods. The average ERP waveforms of the least and most contaminated trials groups have become almost identical. This is true for each of the removal methods and is confirmed by the SPM analyses. No significant differences were found when comparing the least with the most contaminated trial groups after removing the eye movement artifacts. Even at the uncorrected, voxel-wise significance level of $p < .001$, this comparison yielded no significant differences for any of the removal procedures.

Least Contaminated RAW Trials versus Least Contaminated Trials after Eye Artifact Removal

The SPM analysis and subsequent F contrasts comparing the least contaminated valid trials before removal of the eye movement artifacts to the least contaminated valid trials after removal of the eye movement artifacts yielded no significant differences for any of the removal methods. This was also true when we lowered the threshold for significance to uncorrected $p < .001$. As shown in Figure 4, none of the removal methods induced con-

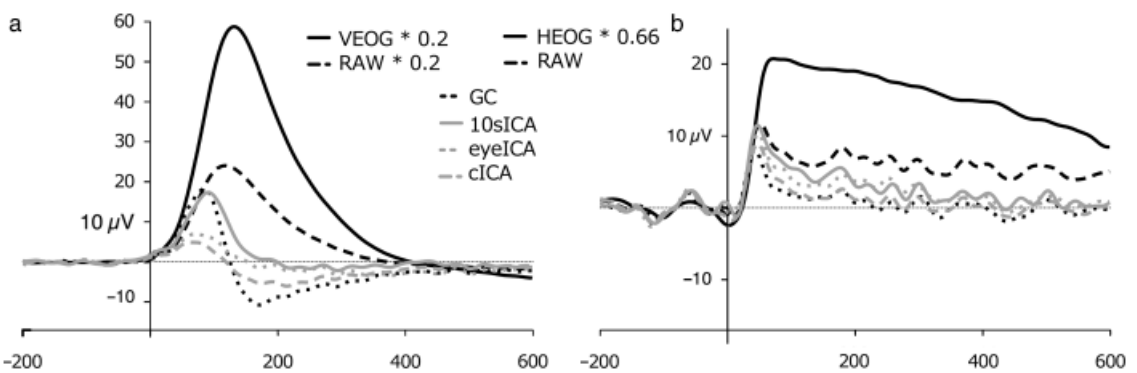

Figure 2. (a) Grand-average waveforms of the blink artifacts at electrodes Fp (calculated as the mean of Fp1 and Fp2) and VEOG. The signal at VEOG was multiplied by 0.2 to improve comparison with the blink ERP as observed on Fp in the continuous data after removal of the eye artifacts with the different procedures. The RAW signal was multiplied by 0.2. (b) Grand-average waveforms of horizontal eye movement artifacts at electrodes F7/8 and HEOG. HEOG shows the mean of left and right (multiplied by –1) horizontal eye movements at the HEOG channel multiplied by 0.66. All other signals show the mean signal across electrodes F7 and F8 (multiplied by –1).

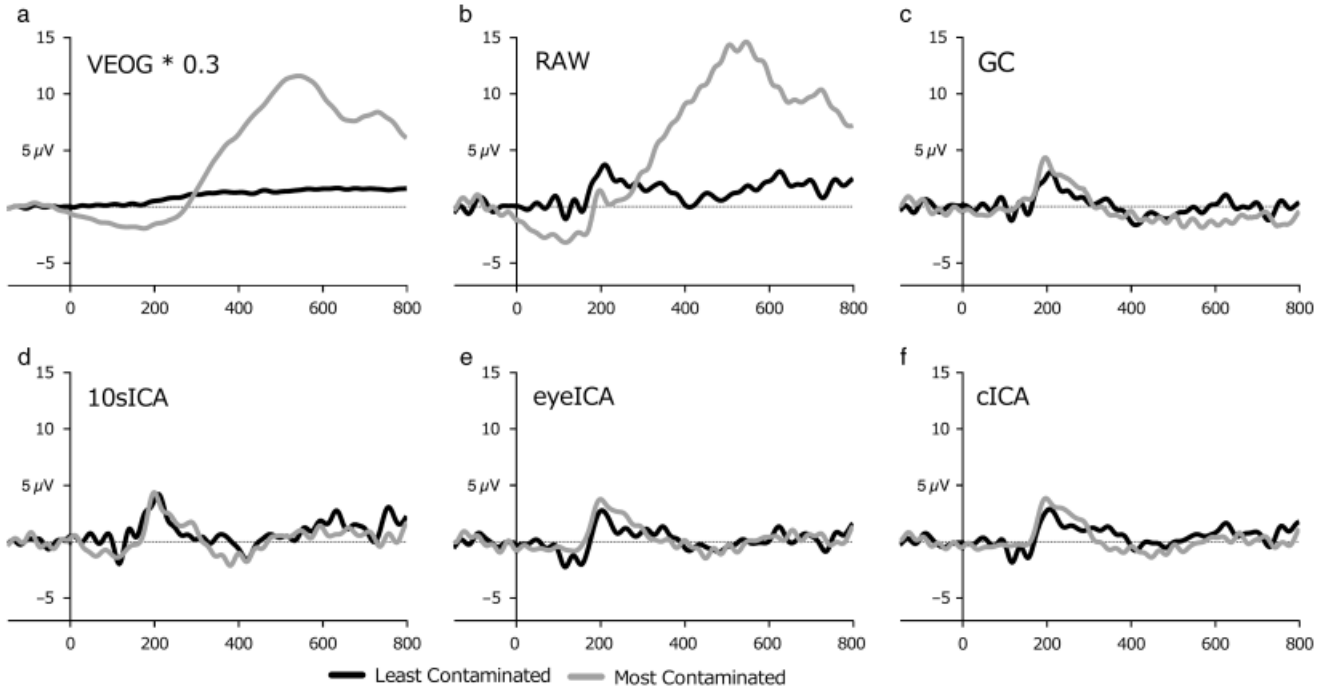

Figure 3. Grand-average ERP waveforms at electrode Fp (calculated as the mean of Fp1 and Fp2) for the valid trials of the cued attention paradigm classified as least (black line) or most (gray line) contaminated by eye movement artifacts. (a) Uncorrected signal measured at the VEOG channel. (b) Trials not corrected for eye movement artifacts (RAW). (c–f) Trials after the removal of the eye movement artifacts with the respective procedures (GC, 10sICA, eyeICA, cICA).

siderable changes in epochs where no eye movement artifacts were detected.

Standard Deviation

A repeated measures analysis of variance (ANOVA) with removal method (RAW, REJECT, GC, 10sICA, eyeICA and cICA) as the within-participants factor on the standard deviation measured across trials and time points in the 1000-ms interval after stimulus presentation yielded a significant effect of methods (Greenhouse–Geisser corrected $p < .005$, $\varepsilon = .29$) at Fp (Figure 5). Post hoc comparisons using the Tukey procedure indicated that there was no difference in standard deviation between the REJECT, GC, 10sICA, eyeICA, and cICA methods, and these methods all had a smaller standard deviation compared to the RAW data ($p < .001$ for each comparison). This effect gradually decreased toward the posterior electrodes (see Figure 5; Fz: $p < .003$; Cz: $p = .07$; Pz: $p = .2$). This result indicates that all eye artifact removal methods equivalently reduced the variance associated with eye movement artifacts.

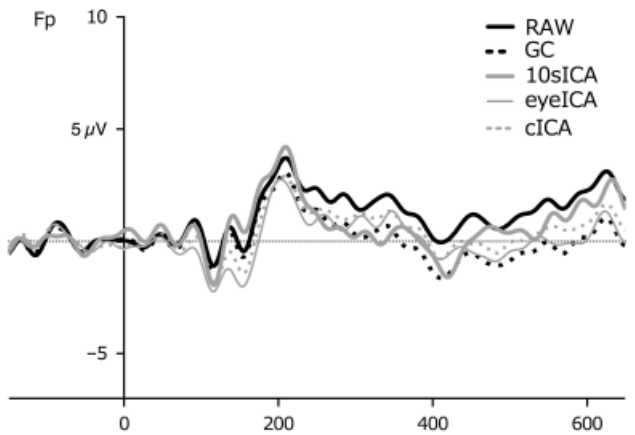

Figure 4. Grand-average waveforms at electrode Fp (calculated as the mean of Fp1 and Fp2) for the valid trials of the cued attention paradigm classified as least contaminated by eye movements. ERPs are shown uncorrected for eye movement artifacts (RAW) and corrected for these artifacts according to the different removal procedures (GC, 10sICA, eyeICA, cICA).

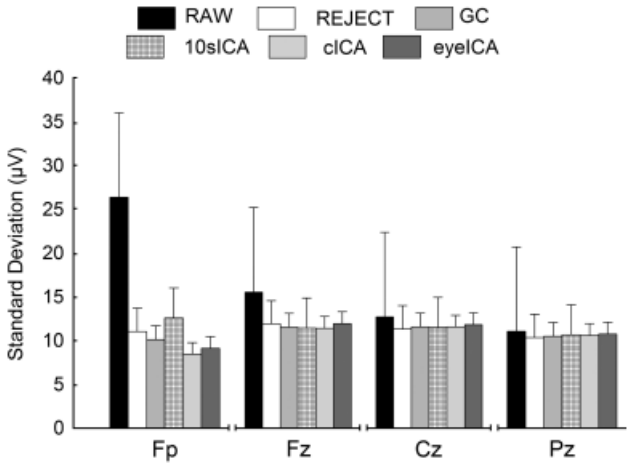

Figure 5. Mean standard deviation across participants for all time points in the 0–1000-ms interval following stimulus presentation and all validly cued trials of the cued attention paradigm. Values are shown for electrodes Fp (calculated as the mean of Fp1 and Fp2), Fz, Cz, and Pz. At electrode Fp, RAW was significantly higher than each of the other procedures ($p < .001$ for each procedure). This effect gradually decreased toward the posterior electrodes (Fz: $p < .003$; Cz: $p = .07$; Pz: $p = .2$).

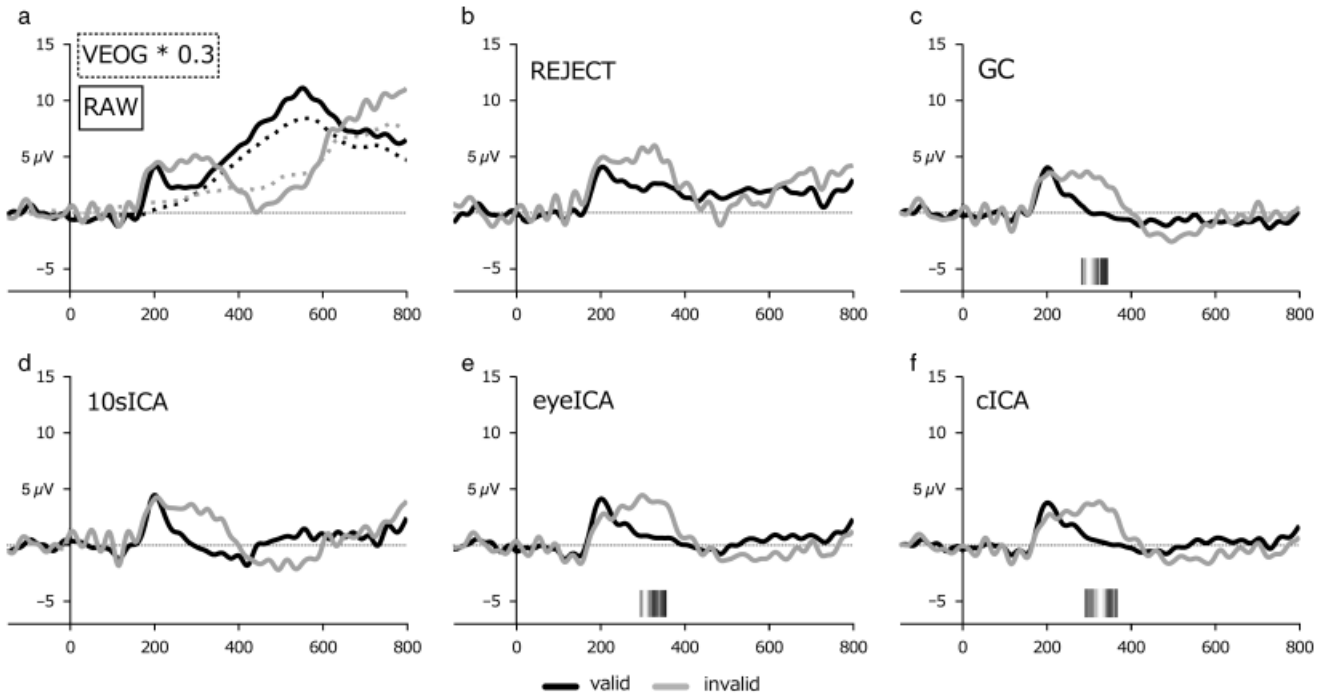

Figure 6. Grand-average ERP waveforms at electrode Fp (calculated as the mean of Fp1 and Fp2) of the valid (black lines) and invalid (gray lines) trials of the cued attention paradigm. ERPs for the six different removal procedures are shown. (a) RAW (lines) and the uncorrected signal as measured at the VEOG channel (multiplied by 0.3; dotted lines). (b) REJECT. (c) GC. (d) 10sICA. (e) eyeICA. (f) cICA method. Gray bars in the c, e, and f graphs indicate significant intervals as found in the SPM valid < invalid t contrast. Lighter colors indicate more significant t values at $p < .05$, corrected for multiple comparisons.

Valid versus Invalid Cue Trials

Figure 6 shows the grand average ERP waveforms for the valid and invalid trials of the cued attention paradigm at electrode Fp (calculated as the mean of Fp1 and Fp2) for the different artifact removal methods. The largest condition-related difference could be observed in the P300 peak, around 300 ms after stimulus onset. Invalid trials were associated with a larger P300 peak amplitude compared to the valid trials, on both the frontal and posterior electrodes. This effect was already evident in the RAW trials. However, in these trials, the P300 effect at the frontal electrodes was partially obscured by the eye movement artifact that could be observed in the ERP from the P300 onward (see Figure 6a). Accordingly, before the removal of the eye movement artifacts, the SPM analysis revealed significant differences between valid and invalid trials for the P300 peak but only at posterior electrodes (not shown in Figure 6). Over the posterior scalp, an extensive significant difference was observed, with a local maximum at electrode P3 ($t = 5.57$, corrected $p < .001$). Over the frontal scalp, a small activation area comprising only electrode F3 was found ($t = 4.49$, corrected $p < .001$).

Figure 6b–f indicates that the different rejection and removal methods each seemed to adequately eliminate the eye movement artifacts from the ERP. However, the effects of the artifact removal for the statistical analysis differed as a function of the removal method. Table 2 gives an overview of the number of significant voxels and clusters in each analysis. Remarkably, the numbers in Table 2 suggest that rejecting trials based on extreme values yielded less significant differences compared to using the RAW data. The same local maxima as for the RAW data were found, but the number of significant voxels was greatly reduced (see Table 2). Here, the gain in statistical power because of the smaller standard deviation is undone by the loss of power because of the smaller number of trials.

The other four methods (GC, 10sICA, eyeICA, and cICA) all yielded more significant voxels compared to the analysis for the raw data, immediately showing the usefulness of these methods. Within these four methods the analysis for the 10sICA method yielded the least significant voxels, whereas the number of significant voxels for the GC, eyeICA, and cICA was similar.

Discussion

The results presented here indicate that removal of eye movement artifacts by ICA-based methods achieved similar performance compared to the more frequently used Gratton and Coles method. Removing the eye movement artifacts by rejecting trials based on extreme values was the least favorable option. Except

Table 2. *Results of the SPM Analyses Comparing the Valid and Invalid Cue Trials of the Cued Attention Paradigm*

	REJECT	RAW	10sICA	eyeICA	GC	cICA
N significant voxels	128	342	434	736	844	945
*N*clusters	2	1	2	1	1	2
Clusters < 20 voxels	1	0	1	0	0	1
N local maxima	2	2	4	3	3	4
Highest t value	4.84	5.57	5.47	5.72	6.08	6.00
Critical t value	3.83	3.86	3.87	3.84	3.83	3.83

Note: For each removal method a valid < invalid t contrast was defined. Results were assessed at a $p < .05$ alpha level, corrected for multiple comparisons.

for the REJECT method, all eye movement artifact removal methods improved statistical sensitivity at the frontal channels without affecting data at channels that were not as strongly influenced by the eye movements. Selecting only eye movement epochs for ICA compilation facilitated the component identification process, thus limiting subjective user input to a minimum.

In the current study, rejecting epochs based on extreme values as a means of removing eye movement artifacts seemed worse than applying no correction at all. The loss of statistical power due to the reduction in the number of trials outweighed the gain in power achieved by reducing the standard deviation and easily outweighed the ease of use of this technique. Here the rejection-based method for removing the eye movement artifacts resulted in a loss of 30% of trials. It should be noted that the participants in the current study were adults instructed to maintain central fixation and to blink at a minimum. It is likely that testing clinical populations or children would result in even higher numbers of lost trials. In contrast, if only few eye movements are present in a data set, removing epochs containing an eye movement artifact might be the method of choice, thus excluding "data manipulation" as present in the subtraction-based artifact removal methods. It should be noted that other factors, such as the specific task used in an experiment, could also influence the frequency of eye movements. This should be taken into account when opting for the REJECT method.

Selecting only eye movement epochs for ICA compilation facilitated the component identification process. The selection of data to use for ICA is important, as the algorithm will calculate sources from this data. The better a source is represented in the raw data, the higher the likelihood that its activity will be captured in one single component and the easier it becomes to identify this source or component as being of interest. Here we compared three data selections for ICA. First we performed ICA on only 10 s of EEG recording, as suggested by others (Jung et al., 1998a, 1998b, 2000a). This eliminates the need to compile large amounts of EEG data with ICA, which is time-consuming. The current results suggest that, although the 10sICA method is better than using raw data, it seems not fully satisfactory. Although this method provided the fastest computation times, the selection of components to remove was less straightforward, with components apparently representing eye movement artifacts not reaching our cutoff of 15% explained variance on the EOG channels.

In contrast, component selection was more straightforward when performing ICA on the complete EEG recording (cICA), indicating a good representation of eye movement artifact in the data entered. The drawback to this procedure is that it may take a considerable amount of time for the ICA to compile, especially with large data sets. This drawback largely disappears when performing ICA only on epochs containing the eye movement artifacts (eyeICA). Indeed, fewer data will need to be compiled while the straightforward identification of components is retained. Our current results indicate that the performances of the cICA and eyeICA were almost similar. The main difference was that the eyeICA yielded a smaller number of significant "voxels" in the task-relevant ERP comparison.

In light of these findings, eyeICA provided the best opportunity for making eye movement artifact removal based on ICA more straightforward. Quantitative identification of components to remove provides an advantage over user-driven, subjective selection of components (Joyce et al., 2004; Romero, Mananas, & Barbanoj, 2008). It increases objectivity of data preprocessing, making results more comparable among research groups. Of course, evaluation of the components selected by the algorithm is warranted at all times. In the current study, selecting data for ICA computation based on the amount of explained variance in the EOG channels facilitated the identification of components representing artifacts. Further research is needed to assess whether the selection criteria used in the current study are usable across studies.

Here, we applied ICA at an early stage in data processing and subsequently "filtered" the data by subtracting components representing eye movement artifacts, similar to the GC method. Another approach would be to calculate ERPs of interest without first correcting the data for eye movement artifacts and subsequently applying ICA to remove components in the ERP waveforms that are identified as eye movement artifacts. Further research is needed to assess the difference in results between these two approaches, as it might be more difficult to disentangle eye movement artifacts from cognition in ERPs that primarily represent cognition as compared to a complete EEG data set that contains all sorts of variance (Ille et al., 2002). A second consideration entails the number of components that are used by ICA and the GC method to describe the eye movement artifacts. The GC method first estimates a correction factor for vertical eye movements, which is dominated by blinks. Subsequently, after subtraction of this "blink factor," a second factor is calculated for the remaining saccades (Gratton et al., 1983). ICA is unconstrained in the number of components identified as eye movement artifacts. In general and evident in the example described in the Methods section, ICA will identify separate components for vertical and horizontal eye movement artifacts. In our example, a second component representing a part of the vertical eye movement artifacts was selected. Together with the first vertical component they possibly account for more of the eye movement artifact-related variance than the vertical eye movement artifact component resulting from the GC method. However, the component solution provided by ICA entirely depends on the data, possibly resulting in a different number of eye artifact components for different participants. The number of components to remove from the data is an important issue when using blind source separation methods such as ICA (Ille et al., 2002; Jung et al., 2000a).

Special caution is needed when interpreting the raw eye movement artifact average (Figure 2). Intuitively one would interpret a smaller eye movement-related ERP with better correction of the eye movement artifact, implying a flat ERP for optimal correction. However, this disregards that eye movements could interfere with perception and are related to cognition (e.g., fewer blinks in a difficult task), which might cause the brain to produce an eye movement-related ERP (Berg & Davies, 1988). More research is needed to fully address the effect of eye movements on cognition and to assess whether the current methods and ICA in particular are sufficiently successful in separating eye movements artifacts from cognition associated with eye movements.

The number of participants and trials in the current study were limited. However removal of eye movement artifacts should be completed for each subject separately because artifacts will differ among subjects (Iwasaki et al., 2005). Second, it remains difficult to evaluate the different eye artifact removal methods because we do not know what a corrected waveform should look like (Croft, Chandler, Barry, Cooper, & Clarke, 2005). Here, we tried to counteract this limitation by evaluating the removal

methods on different aspects of the data. Across these evaluations we showed that ICA is a valid technique for the removal of eye movement artifacts. Current software allows for an easy implementation of blind source separation methods, which can also achieve good results in the absence of EOG channels (Romero et al., 2008). Selecting only eye movement epochs for ICA compilation facilitated identification of components representing eye movement artifacts.

REFERENCES

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, *7*, 1129–1159.

Berg, P., & Davies, M. B. (1988). Eyeblink-related potentials. *Electroencephalography and Clinical Neurophysiology*, *69*, 1–5.

Berg, P., & Scherg, M. (1994). A multiple source approach to the correction of eye artifacts. *Electroencephalography and Clinical Neurophysiology*, *90*, 229–241.

Corbetta, M., Kincade, J. M., & Shulman, G. L. (2002). Neural systems for visual orienting and their relationships to spatial working memory. *Journal of Cognitive Neuroscience*, *14*, 508–523.

Croft, R. J., & Barry, R. J. (2000). EOG correction of blinks with saccade coefficients: A test and revision of the aligned-artifact average solution. *Clinical Neurophysiology*, *3*, 444–455.

Croft, R. J., Chandler, J. S., Barry, R. J., Cooper, N. R., & Clarke, A. R. (2005). EOG correction: A comparison of four methods. *Psychophysiology*, *42*, 16–24.

Debener, S., Hine, J., Bleeck, S., & Eyles, J. (2008). Source localization of auditory evoked potentials after cochlear implantation. *Psychophysiology*, *45*, 20–24.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics. *Journal of Neuroscience Methods*, *134*, 9–21.

Esposito, F., Seifritz, E., Formisano, E., Morrone, R., Scarabino, T., Tedeschi, G., et al. (2003). Real-time independent component analysis of fMRI time-series. *NeuroImage*, *20*, 2209–2224.

Flexer, A., Bauer, H., Pripfl, J., & Dorffner, G. (2005). Using ICA for removal of ocular artifacts in EEG recorded from blind participants. *Neural Networks*, *18*, 998–1005.

Ford, M. R., Sands, S., & Lew, H. L. (2004). Overview of artifact reduction and removal in evoked potential and event-related potential recordings. *Physical Medicine and Rehabilitation Clinics of North America*, *15*, 1–17.

Frank, R. M., & Frishkoff, G. A. (2007). Automated protocol for evaluation of electromagnetic component separation (APECS): Application of a framework for evaluating statistical methods of blink extraction from multichannel EEG. *Clinical Neurophysiology*, *118*, 80–97.

Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. J. (1991). Comparing functional (PET) images: The assessment of significant change. *Journal of Cerebal Blood Flow and Metabolsim*, *11*, 690–699.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484.

Huang, R. S., Jung, T.-P., Delorme, A., & Makeig, S. (2008). Tonic and phasic electroencephalographic dynamics during continuous compensatory tracking. *NeuroImage*, *39*, 1896–1909.

Iidaka, T., Matsumoto, A., Nogawa, J., Yamamoto, Y., & Sadato, N. (2006). Frontoparietal network involved in successful retrieval from episodic memory. Spatial and temporal analyses using fMRI and ERP. *Cerebal Cortex*, *16*, 1349–1360.

Ille, N., Berg, P., & Scherg, M. (2002). Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies. *Journal of Clinical Neurophysiology*, *19*, 113–124.

Iwasaki, M., Kellinghaus, C., Alexopoulos, A. V., Burgess, R. C., Kumar, A. N., Han, Y. H., et al. (2005). Effects of eyelid closure, blinks, and eye movements on the electroencephalogram. *Clinical Neurophysiology*, *116*, 878–885.

Joyce, C. A., Gorodnitsky, I. F., & Kutas, M. (2004). Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. *Psychophysiology*, *41*, 313–325.

Jung, T.-P., Humphries, C., Lee, T.-W., Makeig, S., McKeown, M. J., Iragui, V., et al. (1998a). Extended ICA removes artifacts from electroencephalographic data. *Advances in Neural Information Processing Systems*, *10*, 894–900.

Jung, T.-P., Humphries, C., Lee, T.-W., Makeig, S., McKeown, M. J., Iragui, V., et al. (1998b). Removing electroencephalographic artifacts: Comparison between ICA and PCA. *Neural Networks Signal Processing*, *VIII*, 63–72.

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M. J., Iragui, V., et al. (2000a). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, *37*, 163–178.

Jung, T.-P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000b). Removal of eye activity artifacts from visual event-related potentials in normal and clinical participants. *Clinical Neurophysiology*, *111*, 1745–1758.

Jung, T.-P., Makeig, S., Westerfield, W., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2001). Analysis and visualization of single-trial event-related potentials. *Human Brain Mapping*, *14*, 166–185.

Kansaku, K., Muraki, S., Umeyama, S., Nishimori, Y., Kochiyama, T., Yamane, S., et al. (2005). Cortical activity in multiple motor areas during sequential finger movements: An application of independent component analysis. *NeuroImage*, *28*, 669–681.

Makeig, S., Delorme, A., Westerfield, M., Jung, T.-P., Townsend, J., Courchesne, E., et al. (2004). Electroencephalographic brain dynamics following manually responded visual targets. *PLOS Biology*, *2*, 747–762.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, *32*, 3–25.

Romero, S., Mananas, M. A., & Barbanoj, M. J. (2008). A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: A simulation case. *Computers in Biology and Medicine*, *38*, 348–360.

Slagter, H. A., Lutz, A., Greischar, L. L., Francis, A. D., Nieuwenhuis, S., Davis, J. M., et al. (2007). Mental training affects distribution of limited brain resources. *PLOS Biology*, *5*(6), e138.

Stone, J. V. (2002). Independent component analysis: An introduction. *Trends in Cognitive Sciences*, *6*, 59–64.

Verleger, R., Gasser, T., & Möcks, J. (1982). Correction of EOG artifacts in event-related potentials of the EEG: Aspects of reliability and validity. *Psychophysiology*, *19*, 472–480.

Vigario, R., Sarela, J., Jousmaki, V., Hamalainen, M., & Oja, E. (2000). Independent component approach to the analysis of EEG and MEG recordings. *IEEE Transactions on Biomedical Engineering*, *47*, 589–593.

Vigario, R. N. (1997). Extraction of ocular artefacts from EEG using independent component analysis. *Electroencephalography and Clinical Neurophysiology*, *103*, 395–404.

Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). A three-dimensional statistical analysis for rCBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, *12*, 900–918.

Zeki, S., Perry, R. J., & Bartels, A. (2003). The processing of kinetic contours in the brain. *Cerebal Cortex*, *13*, 189–202.

(Received October 9, 2008; Accepted November 12, 2009)