Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals

Article *in* Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology · January 2011

DOI: 10.1016/j.clinph.2010.04.036 · Source: PubMed

CITATIONS
37

READS
408

4 authors, including:

Takako Fujioka
Stanford University
51 PUBLICATIONS **2,346** CITATIONS

SEE PROFILE

Nasser Mourad
College of Enginering Shaqra University
27 PUBLICATIONS **200** CITATIONS

SEE PROFILE

Laurel Trainor
McMaster University
224 PUBLICATIONS **13,724** CITATIONS

SEE PROFILE

Clinical Neurophysiology xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

Clinical Neurophysiology

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

Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals

Takako Fujioka [a,b,1], Nasser Mourad [c,1], Chao He [a], Laurel J. Trainor [a,b,*]

[a] Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON, Canada L8S 4K1
[b] Rotman Research Institute, Baycrest, University of Toronto, 3560 Bathurst st., Toronto, ON, Canada M6A 2E1
[c] Department of Electrical and Computer Engineering, McMaster University, 1280 Main Street West, Hamilton, ON, Canada L8S 4K1

ARTICLE INFO

Article history:
Accepted 18 April 2010
Available online xxxx

Keywords:
Development
Infant
Artifact correction
High-density EEG recording
Continuous EEG
Simulation

ABSTRACT

Objective: EEG recording is useful for neurological and cognitive assessment, but acquiring reliable data in infants and special populations has the challenges of limited recording time, high-amplitude background activity, and movement-related artifacts. This study objectively evaluated our previously proposed ERP analysis techniques.
Methods: We compared three artifact removal techniques: *Conventional Trial Rejection* (CTR), *Independent Channel Rejection* (ICR; He et al., 2007), and *Artifact Blocking* (AB; Mourad et al., 2007). We embedded a synthesized auditory ERP signal into real EEG activity recorded from 4-month-old infants. We then compared the ability of the three techniques to extract that signal from the noise.
Results: Examination of correlation coefficients, variance in the gain across sensors, and residual power revealed that ICR and AB were significantly more successful than CTR at accurately extracting the signal. Overall performance of ICR and AB was comparable, although the AB algorithm introduced less spatial distortion than ICR.
Conclusions: ICR and AB are improvements over CTR in cases where the signal-to-noise ratio is low.
Significance: Both ICR and AB are improvements over standard techniques. AB can be applied to both continuous and epoched EEG.

1. Introduction

The brain's response to events such as the presentation of sounds, speech, and music can be examined with event-related potentials (ERPs) extracted from EEG recordings. These ERPs can be measured in preverbal infants and other groups for whom verbal responses and behavioural methods can be difficult (Trainor, 2008). Thus they are particularly useful for investigating learning and maturation during development as well as for the objective assessment of neurological, perceptual and cognitive status in special populations (Steinschneider and Dunn, 2003). ERPs track the stages of information processing over time, from sensory to perceptual to cognitive, such that the points can be identified at which differences are evident across age or between groups of subjects. Furthermore, recently-developed high-density EEG recordings allow up to several hundred places on the scalp to be sampled concurrently, which enables good estimation of the sources of activation in the brain.

In addition to neural processing of the presented stimulus events, EEG signals contain two types of noise: background brain activity that is largely unrelated to processing the event, and movement artifacts due to activity such as eye blinks and head movement. In contrast to the unrelated brain activity, the ERP signal is phase locked to the stimulus onset. Accordingly, the standard procedure for estimating an ERP is to average over a large number of trials. Movement artifacts are typically an order of magnitude larger than the ERP signal, and they are typically eliminated before averaging is done. There are several approaches for eliminating high-amplitude artifacts from EEG data. In a common approach, called *Conventional Trial Rejection* (CTR), high-amplitude artifacts are identified in individual channels on individual trials by their large amplitude; the data across all electrodes are then eliminated for trials containing artifact in any channel. This approach works successfully for most adult data because there are many trials and few contain such artifact, leaving a sufficient number of trials for averaging. In a second approach, EEG responses to eye blinks and movements are measured in each subject, the sources of this

* Corresponding author at: Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON, Canada L8S 4K1. Tel.: +1 905 525 9140x23007; fax: +1 905 529 6225.
E-mail address: ljt@mcmaster.ca (L.J. Trainor).
[1] Takako Fujioka and Nasser Mourad contributed equally to this manuscript and should be considered as co-first authors.

1388-2457/$36.00
doi:10.1016/j.clinph.2010.04.036

Please cite this article in press as: Fujioka T et al. Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. Clin Neurophysiol (2010), doi:10.1016/j.clinph.2010.04.036

activity are modeled, and these sources are subtracted from the EEG responses obtained during the study (Berg and Scherg, 1994; Gratton et al., 1983; Lins et al., 1993b). This approach works well for adult data because eye movement and eye blinks give rise to consistent ERP responses and can be modeled with a small number of sources. A third approach is based on modeling the measured EEG data as a linear combination of a set of independent components. This *Independent Component Analysis* (ICA) technique can be utilized for separating the signal from the artifact because the few independent components with largest amplitude contain most of the artifacts. These independent components can then be removed from the EEG data before averaging (Jung et al., 2001).

Unfortunately, these standard methods do not work optimally with data from infants or other populations for whom being still for long periods of time is problematic. In this paper we compare and evaluate different methods for artifact elimination and averaging with infant data, and we present two new methods that yield much improved results. There are a number of challenges involved in recording and analyzing EEG in young infants and those with neuro-developmental disorders. First, in infants and young children, background EEG activity is relatively high in amplitude and slow in frequency (Bell and Wolfe, 2008). This is particularly problematic for ERP analysis, because the amplitude of ERP components, typically less than 5 μV, can be ten times smaller than the background EEG. Second, the recording time that infants can tolerate is severely limited, resulting in far fewer trials than with normal adults. Because the effectiveness of averaging in reducing background noise depends on the number of trials, it is generally less effective for infant than for adult data. Furthermore, in order to obtain high density recordings that allow examination of activation sources, a system that can tolerate electrical impedances of up to about 50 kΩ must be used in practice, because it can be applied to the head in less than 5 min compared to more than 30 min in the case where impedances must be kept below 5 kΩ. Unfortunately, high impedance recordings often are more subject to extrinsic noise at the sensors and along the sensor wires.

Third, it is difficult to precisely identify and remove artifacts arising from eye blinks and eye movements in infants because the EEG components resulting from such movements are not as systematic and temporally confined as in adults (Bell and Wolfe, 2008). This means that identification of eye-related artifacts in infant EEG is problematic by itself. Thus, although it is well known that approaches like ICA are quite effective in removing typical ocular artifacts in adult data, they are often ineffective when applied to infant data. Furthermore, because of the limited amount of time that an infant will remain cooperative, it is not practical to record EEG responses to eye movements in infants for later template matching correction using regression, principal component analysis (PCA) or source models (Berg and Scherg, 1994; Gratton et al., 1983; Lagerlund et al., 1997; Lins et al., 1993a,b). Finally, infants tend to move abruptly and often, which introduces high amplitude artifact into the EEG signal. Such abrupt movements can cause the temporary loss of good contact between particular sensors and the scalp in high-impedance systems. These movement artifacts often contaminate only a few electrodes on any one trial, and different movements affect different trials at different times, as illustrated in the example EEG in Fig. 1A. Accordingly, they cannot be modeled as sources across trials and, as a consequence, they cannot be removed using ICA. To illustrate this, we applied ICA on the infant EEG dataset shown in Fig. 1A. The EEG dataset was collected using 124 electrodes in a geodesic net (HydroCel GSN, Electrical Geodesics, Inc., Eugene, OR) with C_z as the reference electrode. For presentation purposes only, the data is rearranged such that the channels presented in the figure are the only channels with noticeable artifact in the shaded interval. We utilized ICA to remove the artifacts in the shaded interval. The potential artifact sources identified by ICA are shown in Fig. 1B. The first problem is that the ICA algorithm spread the artifacts across 14 different sources. It can be seen that these independent sources also contain clean brain signals at different time points in addition to the artifacts. Consequently, removing the first 14 sources will also remove valuable EEG data as well. Fig. 1C presents the EEG data after removing the first 14 sources. As shown in this figure, the artifacts were partially, but not totally, removed from the EEG data in the shaded interval. Also, it is clear that the artifacts outside the shaded interval still remain. In sum, ICA does not work well with infant data of this type because even if some of the noise is modeled and removed by ICA, noises in other time windows remain because they have a different source from those eliminated (Fig. 1AB).

In general, the *Conventional Trial Rejection* (CTR) method used with adults is not optimal with infants because the number of trials is small, infants move a great deal, and there are many, varying sources of artifact. In this context, it should be noted that the oddball paradigm, which is commonly used with infants, is particularly problematic. In this procedure, ERPs are measured to occasional changes (deviants) in an ongoing stream of sound events. Because deviants occur rarely, there are very few trials to go into the average response to deviants.

From this discussion it is clear that there is a need to develop better strategies for estimating ERP signals in infant EEG data. In this paper we systematically compare two alternative methods against the (CTR) procedure. The first is *Independent Channel Rejection* (ICR). We developed this procedure previously (He et al., 2007) to take advantage of the fact, discussed above, that artifacts in infant data are often limited to one or a few electrode sites on any particular trial. In ICR, if a trial contains artifact at one electrode site, data from that site is eliminated for that trial, but data from the rest of the channels contributes to the average. Thus a different number of trials are averaged for each channel. While our previous studies suggest that this method appears to work well (He et al., 2007, 2009a,b; He and Trainor, 2009), the use of different numbers of trials for each electrode site might potentially lead to spatial distortions. In our present comparison of methods we include an evaluation of spatial distortion.

The second method we propose is the *Artifact Blocking* (AB) algorithm developed by members of our group (Mourad et al., 2007). The AB algorithm is performed in two steps. In the first step a reference matrix is constructed from the EEG data matrix by setting to zero all the samples of the EEG data matrix with absolute amplitude exceeding a pre-specified threshold θ. If the value of θ is chosen wisely, the clipped samples will correspond to the high-amplitude artifacts. As a result, the reference matrix does not contain any information about the high-amplitude artifacts. This step is equivalent to the ICR procedure. However, in the second step, the AB algorithm goes one step further than the ICR algorithm by utilizing the EEG data matrix and the reference matrix for estimating a smoothing matrix. The smoothing matrix is estimated such that multiplying the original EEG data matrix by the smoothing matrix produces a new "clean" EEG data matrix. As described in the Supplementary Material, Appendix B, even though the smoothing matrix is applied onto the original EEG data matrix, it has the effect of "projecting" the reference matrix onto the range of the original EEG data matrix, i.e., the new EEG data matrix is the closest matrix (in the range of the original data matrix) to the reference matrix. Since the reference matrix does not have any information about the high-amplitude artifacts, the new EEG data matrix will be clean and does not have any high-amplitude artifacts. Accordingly, the smoothing matrix has the effect of "blocking" the high-amplitude artifacts from the EEG data matrix, hence the name (see the Supplementary Material, Appendix B). Fig. 1D shows the application of AB to the same EEG data as illustrated with the ICA algorithm. It can be seen that, unlike ICA, AB successfully elim-

Please cite this article in press as: Fujioka T et al. Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. Clin Neurophysiol (2010), doi:10.1016/j.clinph.2010.04.036

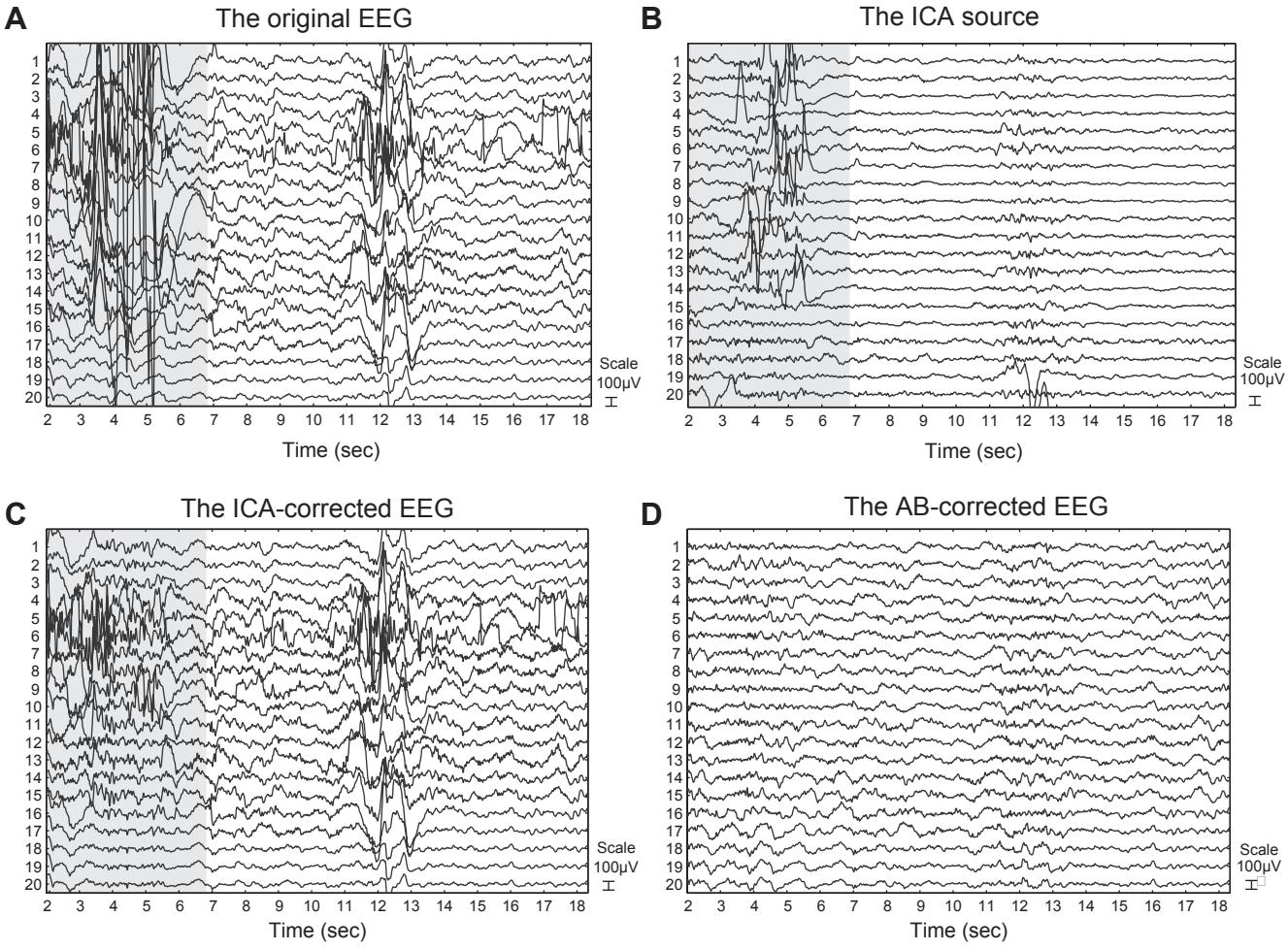

Fig. 1. Comparison between *Independent Component Analysis* (ICA) and *Artifact Blocking* (AB) algorithms in removing high-amplitude artifacts. (A) the EEG data containing artifacts and high-amplitude background activity, (B) the sources obtained by ICA, (C) the corrected data using ICA, and (D) the corrected data using the AB algorithm.

inates the artifact. It can also be seen that channels with high-amplitude artifact are not reduced to zero because AB is a linear transformation of the original data. In a sense, this method resembles interpolation methods because it tries to recover the EEG samples corresponding to the zero samples in the reference data matrix by projecting the reference data matrix onto the range of the original EEG data matrix. However, in contrast to other interpolation techniques that are usually used in the field of EEG data analysis, AB has the advantage that it requires no prior knowledge of the volume-conductor models (Perrin et al., 1987, 1989) or the three-dimensional scalp surfaces (Law et al., 1993; Srinivasan et al., 1996). Thus, the computational demands of AB are much lower. The nature of the AB algorithm makes it particularly suitable for data from infants and atypical populations where structural MRI scans are often not available and there is considerable individual anatomical variation, including the presence or absence of holes in the skull which can have a large affect on electrical volume conduction (Chauveau et al., 2004). The AB algorithm also has advantages over ICA in that it is data-driven and has no assumptions regarding the number of components and the statistical independence between the components. As with ICA, conventional averaging is performed once the AB algorithm has been applied.

Here we present a systematic examination of the three methods, *Conventional Trial Rejection* (CTR), *Independent Channel Rejection* (ICR), and *Artifact Blocking* (AB). In order to evaluate the effectiveness of each algorithm, it was necessary to know the exact ERP signal to be extracted. Thus, we synthesized the ERP signal as the scalp manifestation of an auditory source located bilaterally in the temporal lobes. This signal was then embedded in real EEG data recorded in silence from 4-month-old infants. Each of the three algorithms was applied, and the derived ERP estimates were compared to the known embedded ERP signal. The methods were compared by examining correlations between embedded and derived ERPs as well as by analyzing the gain (amplitude ratio of derived signal ERP to the original signal) and residuals in signal power across the scalp. We also examined how these parameters varied over different numbers of trials for averaging by using randomly resampled trials.

2. Methods

2.1. Subjects

Twelve healthy, full-term 4-month-old infants (5 F, 17 M; mean = 4.7 months, SD = 0.19) with no known hearing deficits participated in the study. Written consent was obtained, and a questionnaire on musical background was completed.

2.2. Background EEG recording procedure

Three episodes of two-minute background EEG activity were recorded from the subjects while no specific auditory stimulation was given. Only this part of the data during the "no-auditory" time windows was used in the present paper. However, these three epi-

Please cite this article in press as: Fujioka T et al. Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. Clin Neurophysiol (2010), doi:10.1016/j.clinph.2010.04.036

sodes alternated with one in which a single piano tone repeated every 450 ms and one in which two piano tones (one on 80% and the other 20% of repetitions) played in random order. Note that although infants did not receive auditory stimulation, the standard procedure was followed where they watched a silent movie and puppet show in order to minimize movements. Some eye movement artifact might have been induced by this, although it is likely less than what would have been present without this visual focus. EEG was recorded using 124-sensor HydroCel GSN nets (Electrical Geodesics, Inc., Eugene, OR) referenced to C_z in a sound-treated room with background noise level less than 29 dB(A). The sampling rate was 1000 Hz. Impedance of the electrodes was kept under 50 kΩ when measured at the beginning of EEG recording.

2.3. Synthesizing the ERP signal

The ERP signal was synthesized using BESA's dipole simulator (MEGIS Software GmbH Gräfelfing, Germany) which, given dipole locations and orientations of brain activity, calculates the ERP pattern at the scalp across the electrode sites in the HydroCel GSN nets that we used to record the background EEG in the infants. The auditory evoked responses were simulated with a pair of sequential downward and upward dipoles co-located in the temporal lobe approximately in the primary auditory area (Talairach coordinate: *x* ±0.66, *y* 0.02, *z* 0.20) in each hemisphere. The source waveform was designed to create, at the surface of the head, a frontal negativity at 0–300 ms, and a frontal positivity at 240–480 ms with peak amplitudes 12 nAm and 20 nAm, respectively. No latency jitter or amplitude fluctuation was used. Note that this forward solution provided 128-channel data, of which four electrodes on the forehead were omitted as these are not used in the nets for infants.

The real background EEG data for each of the 12 infants were off-line filtered between 0.5 and 20 Hz, down sampled to 215 Hz and the synthesized ERP signal was embedded (added) every 700 ms. Thus, epochs were 700 ms, including prestimulus and poststimulus periods of 100 ms and 600 ms, respectively. The resulting number of trials was thus 514, except for one subject whose total number of trials was 385 because of a shorter recorded EEG episode than for the other subjects. Both the synthesized ERP and the background EEG data used a common average reference at C_z.

2.4. Estimation of the ERP signal using CTR, ICR and AB

For each infant, the embedded ERP was estimated using the CTR, ICR (see the Supplementary material, Appendix A) and AB (see the Supplementary material, Appendix B) methods. In applying the AB algorithm, a single smoothing matrix was used for each data set as the results did not improve when different smoothing matrices were used for different data segments (see the Supplementary material, Appendix B). The threshold θ used within the AB algorithm was empirically selected as ±50 μV as it was the lowest value for which the output EEG through the AB algorithm was not over-smoothed (see the Supplementary material, Appendix B). For all three methods, the threshold β, for assuming the presence of artifact, was set to ±100 μV. Note that for the AB procedure, β was applied to the output of the AB algorithm. While 514 trials per infant is a reasonable number, in more than 80% of trials the outer ring of channels met the criterion for the presence of artifact. Thus, these 18 channels were removed from further analysis. The mean number of trials per infant obtained during analysis for each method was 131, 441.7, and 492.3 for CTR, ICR, and AB, respectively. Note that the number of trials was much reduced in CTR compared to ICR and AB. Also the range and variance across individuals in the number of trials varied widely across the methods: CTR (min: 13, max: 238, SD: 83.0), ICR (min: 46, max: 504, SD: 67.8, mean across electrodes), and AB (min: 376, max: 504, SD: 36.65).

2.5. Evaluation of the estimated ERP

To compare the estimated ERP and the synthesized ERP signal in each individual data set, the following simple linear model was considered. For the *k*th EEG dataset, let the estimated ERP signal at the *i*th electrode be expressed as

$$\mathbf{y}_i^k = a_i^k \mathbf{s}_i + \mathbf{n}_i^k, \quad i = 1, \ldots, N. \quad k = 1, \ldots, N_{eeg} \tag{1}$$

where $\mathbf{s}_i \in \mathscr{R}^{(To\times 1)}$ is the known embedded ERP signal at the *i*th electrode, a_i^k is an unknown gain/attenuation parameter, $\mathbf{n}_i^k \in \mathscr{R}^{(To\times 1)}$ is the residual background noise in the estimated ERP signal, and N_{eeg} is the number of EEG data sets (N_{eeg} = 12). Based on this model, three indices were derived: correlation coefficient between the estimated and original ERP waves, gain, and power of residual noise.

2.5.1. Correlation coefficient between the embedded and estimated ERP signals

The correlation coefficient between the estimated ERP signal, $\mathbf{y}_i^k$, and the known embedded ERP signal $\mathbf{s}_i$ quantifies how well the waveform of the embedded ERP signal is preserved in the estimated ERP signal. A higher correlation coefficient indicates better performance.

Specifically, for the *k*th infant, let $R_{y_i s_i}[k]$ denote the correlation coefficient between $\mathbf{y}_i^k$ and $\boldsymbol{s}_i$. Then the average correlation coefficient at the *i*th electrode is calculated as

$$R_i = \frac{1}{N_{eeg}} \sum_{k=1}^{N_{eeg}} R_{y_i s_i}[k], \quad i = 1, \ldots, N \tag{2}$$

For each of the three methods, the correlation was calculated at each electrode site for each of the 12 infants. The correlations were then averaged across the 12 infants and visualized across electrode sites in topographic maps. A repeated measures Analysis of Variance (ANOVA) with two within-subjects factors (method: CTR, ICR, AB; electrode group: left front-temporal, right front-temporal, left occipital, right occipital) was conducted to determine whether some procedures produced statistically significantly higher correlations than others. Post-hoc tests were conducted using Fisher's Protected Least Significant Difference.

2.5.2. Gain and spatial distortion

The correlation coefficient is insensitive as to whether the estimated amplitude matches the embedded amplitude of the ERP, and whether the gain is consistent across electrode sites. The closer the gain parameter, a_i^k, is to 1, the better the obtained ERP. More importantly, the more consistent the gain parameter across channels, the less spatial distortion of the ERP signal. The gain parameter was defined as

$$\hat{a}_i^k = \arg \min_a \left\| \mathbf{y}_i^k - a\mathbf{s}_i \right\|_2^2$$

This problem has a closed form solution given by

$$\hat{a}_i^k = \frac{\mathbf{s}_i^T \mathbf{y}_i^k}{\mathbf{S}_i^T \mathbf{s}_i}, \quad i = 1, \ldots, N \tag{3}$$

The resulting estimation of $\hat{a}_i^k$ was averaged across the 12 infants at each electrode site, and expressed in a topographic map of the gain parameter. Using the standard deviation across all the electrodes as an index of spatial distortion in each infant, the three different artifact-rejection methods were compared statistically by one-way repeated measures ANOVA.

Please cite this article in press as: Fujioka T et al. Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. Clin Neurophysiol (2010), doi:10.1016/j.clinph.2010.04.036

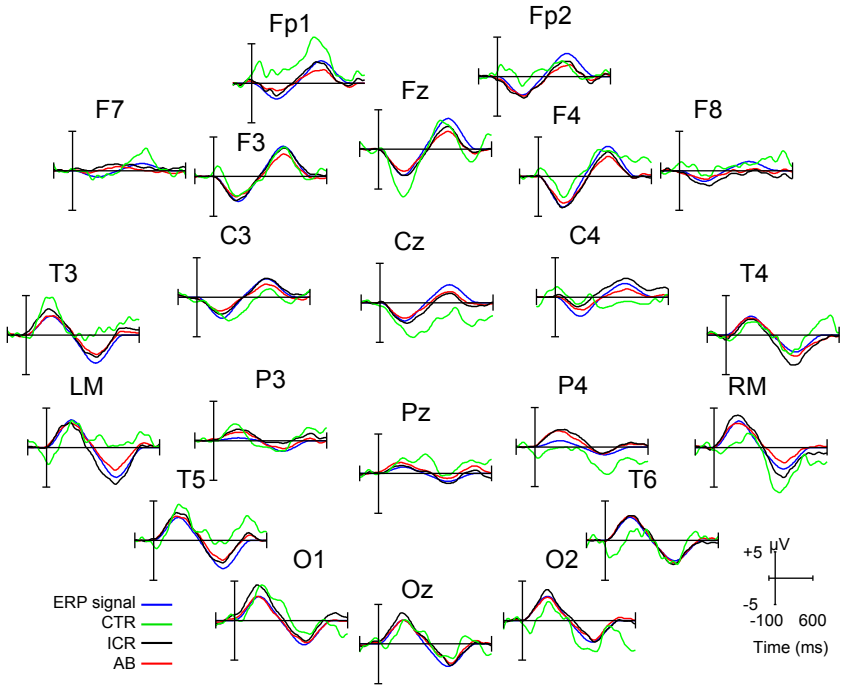

Fig. 2. The original ERP signal and the ERP signals estimated by the *Conventional Trial Rejection* (CTR), *Independent Channel Rejection* (ICR), and *Artifact Blocking* (AB) procedures at selected electrodes across the head.

2.5.3. Residual power

The residual power parameter quantifies the amount of noise left in the estimated ERP signal after applying each artifact-rejection method. Clearly, the best estimator is the one with smallest residual power. Using the estimated gain parameter $\hat{a}_i^k$, the residual activity in the estimated ERP was calculated as

$$\hat{\mathbf{n}}_i^k = \mathbf{y}_i^k - \hat{a}_i^k \mathbf{s}_i \tag{4}$$

at the ith electrode. As with the parameters above, the power of $\hat{\mathbf{n}}_i^k$ was averaged across the infants at each electrode site and mapped out topographically. The different artifact-rejection methods were compared statistically by one-way ANOVA on the averaged power across electrodes in individual subjects.

2.5.4. Effect of number of trials

We examined how the performance of the algorithms varied as a function of different numbers of trials, using the three parameters described in the previous sections. First, 500 trials of the same simulated data (real background EEG and the embedded ERP signal) were generated for each subject (one was omitted because the recorded EEG was less than the length of 400 trials). We then randomly selected 100 trials and calculated the ERP at Fz as well as the three measures of performance, correlation, gain, and residual power. This selection was repeated 50 times with replacement to obtain the best representative estimate of the ERP extracted by each method as if there were 100 trials in the experiment, under the assumption that the background EEG obtained from each subject follows the same normal distribution. We then repeated this procedure for 200, 300, and 400 trials. Mean and standard error of the mean for each of the three parameters for each subject was evaluated by two-way ANOVAs with two within-subject factors, number of trials (100, 200, 300, 400) and method (CTR, ICR, AB). The significance level was set at 0.05. Post-hoc analysis was conducted using Fisher's Protected Least Significant Difference.

3. Results

Fig. 2 shows the original and the estimated ERP at a set of selected electrodes. It can be seen that the ERP estimated using the CTR procedure was noisy compared to those using ICR and AB. In particular, the waveform at some electrodes was drastically different from the embedded signal, with the absence of peaks in some cases and falsely added peaks in others. On the other hand, both ICR and AB successfully estimated the precise morphology of the embedded waveforms at every electrode site. The amplitude of the ERP estimated by ICR and AB was slightly larger or smaller than that of the original ERP signal depending on the electrode location. The first peak around 150 ms was slightly exaggerated by the ICR procedure at some sites, most noticeably at occipital electrodes.

3.1. Correlation coefficient between the embedded and estimated ERP signals

Topographic maps of the correlation coefficients between the original ERP and the estimated ERP signals averaged across the 12 infants for each of the three methods are shown in Fig. 3. The ERP signals estimated by CTR have low correlation coefficients at all electrode sites, whereas the ERP signals estimated by both the AB and ICR techniques have high correlation coefficients at most electrode sites. The few sites with low correlation coefficients for the ICR and AB methods are a consequence of the power distribution of the original embedded ERP signal, which is shown in the upper panel of Fig. 3. Those electrodes at which the embedded signal has low power (represented by the blue color) correspond to those at which the correlation coefficients are low (blue color in the correlation coefficient maps of the bottom row of Fig. 3). This is due to the fact that it is impossible to estimate in noise a signal whose amplitude approaches zero. Consequently, a low correlation coefficient between the embedded and estimated signal is inevitable at these sites. Correlation coefficients were averaged at left and right frontal sites and at left and right occipital sites for each infant and each artifact rejection procedure, and subjected to a two-way

Please cite this article in press as: Fujioka T et al. Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. Clin Neurophysiol (2010), doi:10.1016/j.clinph.2010.04.036

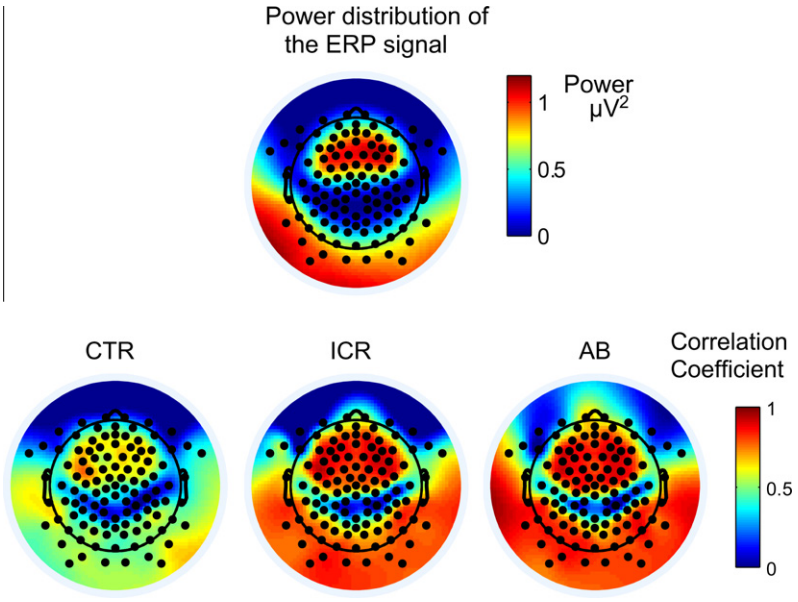

Fig. 3. Correlations between the embedded and estimated ERPs. (Upper panel) Topographic map of the power distribution of the embedded ERP signals. (Lower panel) Correlation coefficients between the embedded ERP signal and the ERP signal estimated using the *Conventional Trial Rejection* (CTR), *Independent Channel Rejection* (ICR), and *Artifact Blocking* (AB) procedures.

ANOVA with method and electrode group as within-subject factors. Results revealed a robust difference between procedures, $F(2, 11) = 13.4$, $p = 0.0002$, and no interaction between methods and electrode groups. Post-hoc tests (Fisher's Protected Least Significant Difference) revealed that there was a significant difference between CTR and ICR ($p < 0.01$) and between CTR and AB ($p = 0.0001$). There was no significant difference between ICR and AB.

In summary, both AB and ICR do very well at estimating the embedded ERP signal in background EEG data from 12 infants, whereas the CTR does poorly. In large part, this result reflects the effect of the number of trials utilized in estimating the ERP signal. While both AB and ICR used most of the trials, the *Conventional Trial Rejection* used only a limited number of trials after artifact rejection.

3.2. Gain and spatial distortion

The upper row of Fig. 4 presents the topographic maps of the grand mean gain parameter, representing the amplitude of the estimated signal compared to the embedded signal, for each method. The lower row presents histogram plots of all the amplitude

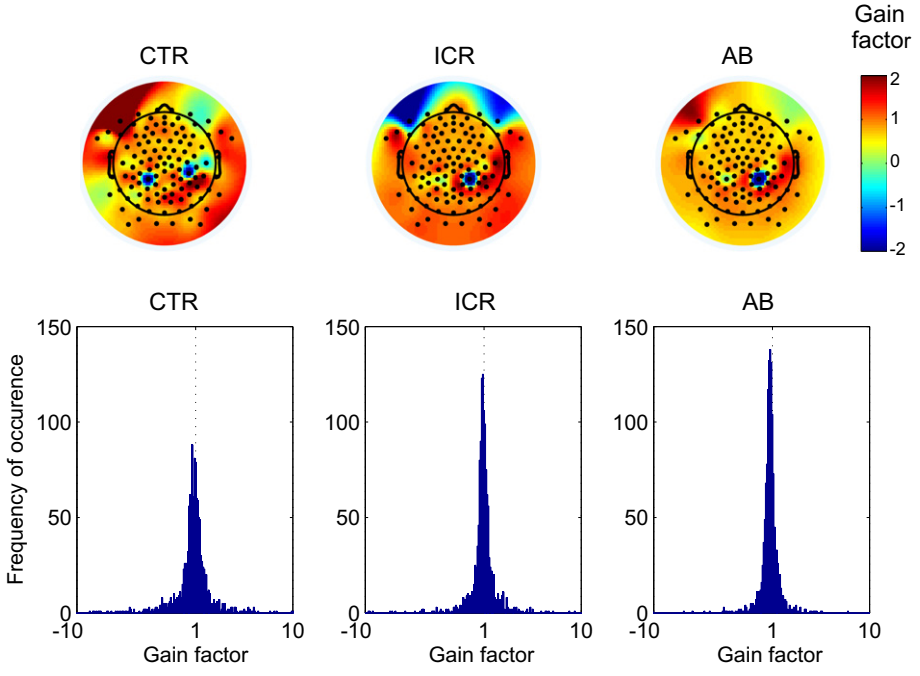

Fig. 4. Gain parameters of the estimated ERP signals. (Upper panel) Topographic map of the average gain parameters for the three procedures. (Lower panel) Histograms of the gain parameters at all electrodes across all infants.

Please cite this article in press as: Fujioka T et al. Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. Clin Neurophysiol (2010), doi:10.1016/j.clinph.2010.04.036

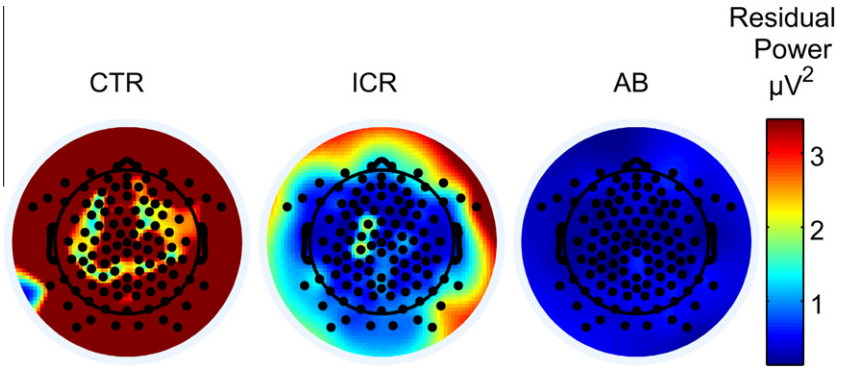

Fig. 5. Residual power. Topographic maps of the distribution of the residual background noise in the ERP signal estimated using *Conventional Trial Rejection* (CTR) (left), *Independent Channel Rejection* (ICR) (middle), and *Artifact Blocking* (AB) (right).

parameters estimated at all electrodes for all infants (i.e. each plot is a histogram of $N \times N_{eeg}$ values).

Ideally, the amplitude parameter should equal 1. Of most importance, when there is little spatial distortion, the gain parameter will be relatively constant across the head. As shown in the upper panel of Fig. 4, the CTR procedure produced large spatial distortion, that is, a wide variation in the gain parameter across sites, whereas the variation produced by ICR and AB techniques is much less. This observation is confirmed by the histogram plots shown in the lower panel of Fig. 4. The variances of the gain parameters when the ERP signal is estimated using CTR, ICR, and AB are 15.08, 4.20, 2.63, respectively. The variance in the gain parameters were calculated for each infant, and converted to standard deviation. The three procedures were compared by a one-way repeated measures ANOVA. Results showed the methods differed significantly, $F(2, 11) = 8.08$, $p = 0.002$. Post-hoc tests showed that the standard deviation was smaller in AB than in CTR ($p < 0.01$) and ICR ($p < 0.05$). The difference between ICR and CTR approached significance ($p = 0.08$). Clearly, the AB method has the smallest variance among the three methods, the ICR has somewhat larger variance, and the CTR has the largest variance, and hence worst performance. Thus, the AB method produces the least spatial distortion, the ICR method next least, and the CTR method the most spatial distortion.

3.3. Residual power

The topographic maps of the average residual power are shown in Fig. 5. As shown in this figure, the AB procedure produced the lowest residual power at most of the electrodes, while the CTR procedure produced the high residual power at almost of the electrodes. While ICR shows lower residual powers than CTR, the residual powers associated with ICR vary somewhat from electrode to electrode, which is likely a direct result of utilizing different numbers of trials for estimating the ERP signal at different electrodes. Statistically, the average residual power at the average of left and right frontal and occipital sites differed across procedures ($F(2,11) = 4.46$, $p = 0.02$), as examined by a one-way ANOVA. Post-hoc tests revealed that the average residual power was larger in CTR than both ICR ($p < 0.05$) and AB ($p < 0.05$), whereas there was no significant difference between ICR and AB.

3.4. Effect of number of trials

Because fewer than 400 trials are often obtained from individual infants, we examined the effect of the number of recorded trials on the performance of the three algorithms using the three parameters described in the previous section at a single electrode, Fz, where the signal was large. We calculated the three parameters through the averaged ERP data obtained from resampled trials with a designated number of trials as described in the methods section. The results are plotted in Fig. 6. Although performance improves with an increasing number of trials for all three algorithms, it can be seen that across all numbers of trials, the AB algorithm has the best performance while the CTR algorithm has the worst performance. Specifically, the AB algorithm has the highest correlation coefficient, the lowest residual power, and an almost constant gain factor. The ANOVA on correlation coefficient (Fig. 6A) revealed significant effects of number of trials, $F(3, 30) = 131.5$, $p < 0.0001$, and method, $F(2, 20) = 4.18$, $p = 0.03$), and no interaction. Post-hoc comparison showed that all possible pairs with different numbers of trials were significantly different ($p < 0.0001$), and that AB was better than CTR overall ($p < 0.01$), and for each number of trials (100, 200, 300: $p < 0.01$, 400: $p < 0.05$). In contrast, the ANOVA for the gain parameter (Fig. 6B) showed no systematic differences between the three methods across number of trials. Fi-

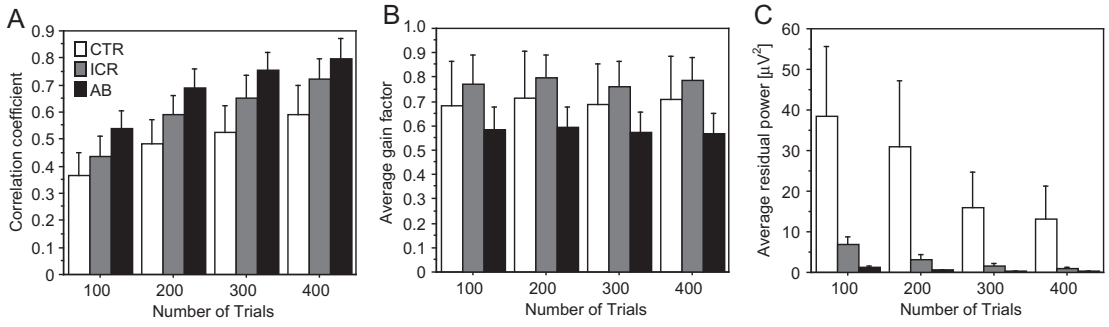

Fig. 6. Effect of the number of trials on the three performance measures. (A) Correlation coefficient, (B) gain parameter, and (C) residual power. For each subject, a designated number of trials (100, 200, 300, and 400) was randomly selected from data at Fz electrode repeatedly 50 times to estimate the individual ERP signal with CTR, ICR, and AB methods. The error bars indicate the standard error of the mean (SEM).

Please cite this article in press as: Fujioka T et al. Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. Clin Neurophysiol (2010), doi:10.1016/j.clinph.2010.04.036

nally, for residual power (Fig. 6C), number of trials, $F(3, 30) = 4.38$, $p < 0.01$, method $F(2, 20) = 3.83$, $p = 0.04$), and their interaction, $F(6, 60) = 2.90$, $p = 0.02$), were all significant. Number of trials made a significant difference particularly between 100 and 300 ($p < 0.01$), and between 100 and 400 ($p < 0.01$). The residual power was significantly larger in CTR than ICR and AB (both: $p < 0.05$). For 100 trials, the effect of method was significant ($p < 0.05$) such that both ICR and AB were better than CTR ($p < 0.05$). For 200 trials, the methods differed significantly ($p < 0.05$) because ICR was marginally better than CTR ($p = 0.05$), while AB was significantly better than CTR ($p < 0.05$). With more trials, the effect of method on the residual power was not significant at the levels of 300 and 400 trials, although the former did approach significance (ns, $p = 0.07$). Thus, overall, the fewer the number of trials, the greater the superiority of the AB algorithm over the other two. The results also indicate that ICR works considerably better than the CTR algorithm.

4. Discussion

Both *Independent Channel Rejection* (ICR) and *Artifact Blocking* (AB) procedures were far better than the *Conventional Trial Rejection* (CTR) procedure at estimating an ERP signal embedded in infant EEG, as indicated by all three evaluation measures examined.

The poor performance of CTR is likely related to the lower number of remaining trials compared to the other two procedures. On average, AB and ICR resulted in three times the number of trials compared to CTR. In contrast, the ICR and AB procedures utilize techniques to eliminate noise without throwing away the signal at the same time. As shown in the results (Fig. 6), AB and ICR work significantly better than does CTR, even at 100 trials, the lowest level of recorded trials tested. In particular, AB gives significantly better correlation and much lower residual power compared to CTR. With increasing numbers of trials, the correlation and residual power improve regardless of method. Consequently, the detailed results (Figs. 3–5) using all the available trials generalize to smaller numbers of trials. In a typical infant ERP recording, the number of trials is usually less than several hundred. It is noteworthy that the number of trials in the CTR procedure is also greatly affected by the choice of channels, such that when those close to the eye or ear or the edge of the net are included, many trials would be rejected by the CTR method. In the present paper we have excluded these electrodes before applying the algorithms. Their inclusion would likely lead to even larger differences between the methods.

The waveform morphology and peak latencies were well estimated by both ICR and AB, but not by CTR. For ICR and AB, there were slight amplitude differences at some electrode sites compared to the embedded signal. In contrast, the morphology produced by CTR was quite inaccurate at many electrodes, making the amplitude and latency of peaks difficult to estimate. Perhaps the CTR data could have benefited from further low-pass filtering to retrieve peak information, but this may well have distorted the peaks further. Some research on infants and young children has suggested that morphology and peak latency have more consistent gradual change over the course of maturation than do the amplitude measures of each peak (Morr et al., 2002;Ponton et al., 2000). The reason for this, however, might be partly because ERP estimates using the standard CTR procedure produce data that is more variable in amplitude than latency across sessions and individuals. The use of the ICR or AB procedures might render peak amplitude estimations sufficiently robust as to be useful, as well as give more accurate estimates of morphology and peak latency.

The results showed that the AB and ICR procedures performed similarly in terms of correlations between the estimated and embedded ERP signals (Fig. 3). However, the AB procedure had a slight advantage over ICR in producing less spatial distortion across the head, that is, the estimated amplitude was more consistently related to the amplitude of the embedded signal across electrodes (Fig. 4). This is particularly important if source analysis is to be performed to estimate where in the brain particular components originated. AB also excelled over ICR in residual power (Fig. 5), probably because AB uses more information from the recorded data than does ICR. Although ICR does not exclude an entire trial if there is noise at some electrodes, it does exclude those electrodes for that trial. By contrast, AB infers the signal on artifact-contaminated data segments from available artifact-free data segments.

One obvious advantage of AB over ICR is that AB is applicable to continuous data whereas ICR is not. Thus, AB can be used to remove high-amplitude artifacts from EEG data that will be analyzed in the frequency domain, for example, to measure steady state responses and power in frequency bands such as alpha, beta and gamma. AB also has advantages over other interpolation procedures for dealing with missing data because of its light computational load and its lack of theoretical assumptions in contrast to source modeling approaches. Although we only examined ERP estimation in the present study, these features of AB suggest that it will be particularly useful for clinical applications where it is necessary to obtain data from difficult subjects in a short period of time and analyze it quickly.

5. Conclusion

Our analysis revealed that the *Artifact Blocking* (AB) and *Independent Channel Rejection* (ICR) methods perform much better than *Conventional Trial Rejection* (CTR) at extracting ERP signals in noisy infant EEG data, as measured by correlations between estimated and original (embedded) ERP signals, variance in estimated compared to embedded amplitude across the scalp (spatial distortion) and residual variance in power. Furthermore, the AB method had significantly lower spatial distortion than ICR, making it a better choice for analysis of the sources of activity in the brain. The AB method has advantages over other interpolation methods in that it has few assumptions and is fast to calculate. The AB method has the additional advantage that it can be applied to continuous data and is therefore a suitable method when the goal is to examine steady state activity or activity in different frequency bands such as alpha, beta, and gamma.

Acknowledgements

The authors thank Elaine Whiskin for assistance in recruiting and testing the infants. This research has been supported by grants to L.J.T. from the Natural Science and Engineering Research Council of Canada and the Canadians Institutes of Health Research.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.clinph.2010.04.036.

References

Bell MA, Wolfe CD. The use of the electroencephalogram in research on cognitive development. In: Schmidt LA, Segalowitz SJ, editors. Developmental psychophysiology: theory, systems, and methods. New York: Cambridge University Press; 2008. p. 150–70.

Berg P, Scherg BM. A multiple source approach to the correction of eye artifacts. Electroencephalogr Clin Neurophysiol 1994;90:29–241.

Chauveau N, Franceries X, Doyon B, Rigaud B, Morucci JP, Celsis P. Effects of skull thickness, anisotropy, and inhomogeneity on forward EEG/ERP computations using a spherical three-dimensional resistor mesh model. Hum Brain Mapp 2004;21:86–97.

Gratton G, Coles MG, Donchin E. A new method for off-line removal of ocular artifact. Electroencephalogr Clin Neurophysiol 1983;55:468–84.

He C, Hotson L, Trainor LJ. Mismatch responses to pitch changes in early infancy. J Cogn Neurosci 2007;19:878–92.

He C, Hotson L, Trainor LJ. Development of infant mismatch responses to auditory pattern changes between 2 and 4 months old. Eur J Neurosci 2009a;29:861–7.

He C, Hotson L, Trainor LJ. Maturation of cortical mismatch responses to occasional pitch change in early infancy: effects of presentation rate and magnitude of change. Neuropsychologia 2009b;47:218–29.

He C, Trainor LJ. Finding the pitch of the missing fundamental in infants. J Neurosci 2009;29:7718–8822.

Jung T-P, Makeig S, Westerfield M, Townsend J, Courchesne E, Sejnowski TJ. Analysis and visualization of single-trial event-related potentials. Hum Brain Mapp 2001;14:166–85.

Lagerlund TD, Sharbrough FW, Busacker NE. Spatial filtering of multichannel electroencephalographic recordings through principal component analysis by singular value decomposition. J Clin Neurophysiol 1997;14:73–82.

Law SK, Nunez PL, Wijesinghe RS. High-resolution EEG using spline generated surface laplacians on spherical and ellipsoidal surfaces. IEEE Trans Biomed Eng 1993;40:145–53.

Lins OG, Picton TW, Berg P, Scherg M. Ocular artifacts in EEG and event-related potentials I: scalp topography. Brain Topogr 1993a;6:51–63.

Lins OG, Picton TW, Berg P, Scherg M. Ocular artifacts in recording EEGs and event-related potentials II: source dipoles and source components. Brain Topogr 1993b;6(1):65–78.

Morr ML, Shafer VL, Kreuzer JA, Kurtzberg D. Maturation of mismatch negativity in typically developing infants and preschool children. Ear Hear 2002;23:118–36.

Mourad N, Reilly JP, De Bruin H, Hasey G, MacCrimmon D. A simple and fast algorithm for automatic suppression of high-amplitude artifacts in EEG data. In: ICASSP, IEEE international conference on acoustics, speech and signal processing – proceedings, 1. Honolulu, HI: 2007.

Perrin F, Bertrand O, Pernier J. Scalp current density mapping: value and estimation from potential data. IEEE Trans Biomed Eng 1987;34:283–8.

Perrin F, Pernier J, Bertrand O, Echallier JF. Spherical splines for scalp potential and current density mapping. Electroencephalogr Clin Neurophysiol 1989;72:184–7.

Ponton CW, Eggermont JJ, Kwong B, Don M. Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. Clin Neurophysiol 2000;111:220–36.

Srinivasan R, Nunez PL, Tucker DM, Silberstein RB, Cadusch PJ. Spatial sampling and filtering of EEG with spline laplacians to estimate cortical potentials. Brain Topogr 1996;8:355–66.

Steinschneider M, Dunn M. Electrophysiology in developmental neuropsychology. In: Segalowitz SJ, Rapin I, editors. Handbook of neuropsychology. Child neuropsychology. 2nd ed. 2003:91–146.

Trainor LJ. Event related potential measures in auditory developmental research. In: Schmidt LA, Segalowitz SJ, editors. Developmental psychophysiology: theory, systems, and methods. New York: Cambridge University Press; 2008. p. 69–102.

Please cite this article in press as: Fujioka T et al. Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. Clin Neurophysiol (2010), doi:10.1016/j.clinph.2010.04.036