Available online at www.sciencedirect.com
ScienceDirect
Neuroscience and Biobehavioral Reviews 30 (2006) 808–822
NEUROSCIENCE AND BIOBEHAVIORAL REVIEWS
www.elsevier.com/locate/neubiorev

Review

Imaging human EEG dynamics using independent component analysis

Julie Onton[a], Marissa Westerfield[b], Jeanne Townsend[b], Scott Makeig[a,*]

[a]Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California San Diego, La Jolla, CA 92093-0961, USA
[b]Department of Neurosciences, University of California San Diego, La Jolla, CA 92093, USA

Abstract

This review discusses the theory and practical application of independent component analysis (ICA) to multi-channel EEG data. We use examples from an audiovisual attention-shifting task performed by young and old subjects to illustrate the power of ICA to resolve subtle differences between evoked responses in the two age groups. Preliminary analysis of these data using ICA suggests a loss of task specificity in independent component (IC) processes in frontal and somatomotor cortex during post-response periods in older as compared to younger subjects, trends not detected during examination of scalp-channel event-related potential (ERP) averages. We discuss possible approaches to component clustering across subjects and new ways to visualize mean and trial-by-trial variations in the data, including ERP-image plots of dynamics within and across trials as well as plots of event-related spectral perturbations in component power, phase locking, and coherence. We believe that widespread application of these and related analysis methods should bring EEG once again to the forefront of brain imaging, merging its high time and frequency resolution with enhanced cm-scale spatial resolution of its cortical sources.

Keywords: EEG; ERP; ICA; Independent component analysis; P300; P3; Aging; Mu; Review

Contents

*Corresponding author.
E-mail address: smakeig@ucsd.edu (S. Makeig).

0149-7634/$ - see front matter © 2006 Published by Elsevier Ltd.
doi:10.1016/j.neubiorev.2006.06.007

1. Imaging human brain dynamics from multi-channel scalp electroencephalographic (EEG) recordings

Even a brief glance at multi-channel EEG data shows that nearby scalp channels record highly correlated signals. Why? Because EEG signals are not produced in the scalp or the brain directly under the recording electrodes. Rather, they are generated by partial synchrony of local field potentials in many distinct cortical domains—each domain being, in the simplest case, a patch of cortex of unknown extent. The radial orientation of pyramidal cells relative to the cortical surface within such a domain allows summation of temporally synchronous extra-neuronal potentials whose summed 'far-field' potentials project to the scalp electrodes near instantly through passive volume conduction. In the absence of such local area synchrony and near parallel orientations of neighboring pyramidal neurons, local field activities would partially or completely cancel each other out, thus preventing far-field potentials of sufficient strength to be detected at scalp electrodes. By the basic laws of electrical conductance, far-field potentials generated within all cortical (and non-brain) domains project to and sum linearly at nearly every scalp electrode. Thus, EEG data recorded at a single electrode are a simple sum (or more technical, a weighted linear mixture) of underlying cortical source signals. The weights of each recorded mixture are determined by the distance of the cortical source domains or patches from the electrode pair ('active' and 'reference'), the orientation of the cortical patch relative to the electrode pair locations, and the electrical properties of intervening tissues (cortex, cerebral-spinal fluid, skull, and skin).

This spatial mixing of EEG source signals by volume conduction produces the strong correlations observed between EEG recordings at nearby electrodes and is the reason why EEG, the first developed and still the most sensitive and dynamic non-invasive brain imaging modality, has long been denigrated as having 'poor spatial resolution.' The term 'spatial resolution' has several meanings, however, and the actual degree of spatial resolution of EEG depends on the intended sense of the term 'resolution'. For any signal modality, three separable meanings of the term 'spatial resolution' are the degree to which the exact location of a single source may be accurately determined; the spatial separation between two sources that is necessary to separate their signals; and the number of such sources that can be separated from the whole data. While the spatial resolution of EEG imaging has in the past been considered to be poor in all three of these aspects, we believe that new techniques for EEG analysis including those discussed in this review significantly improve its spatial resolution by all definirions of the term.

The recovery of the exact cortical distribution of an EEG source region is limited by the undercompleteness of the inverse source localization problem. For example, far-field potentials from two synchronously active but physically opposing cortical source areas—e.g., source areas facing each other on opposite sides of a cortical sulcus—may cancel, and their joint activity will have no effect on the scalp data. If a third area is coherently active, there will be no way to determine from scalp recordings whether the observed activity arises within the third area alone, within all three areas synchronously, or in any other combination of partially self-canceling source areas whose summed activity at the scalp also matches or closely resembles that of the third area alone.

The inverse source localization problem may be greatly simplified by relying on the well-accepted assumptions that EEG signals arise from cortical pyramidal cells oriented perpendicular to the cortical surface and (usually) located within a single contiguous and therefore highly interconnected cortical domain. It is not easy, however, to separately record an EEG scalp distribution generated in only one cortical domain, since many EEG source domains contribute to each recorded EEG signal at nearly all time points. The common method of response averaging, producing event-related potential (ERP) average time courses time-locked to a set of similar stimulus onsets or other events, was originally thought to produce EEG scalp distributions in which only a few source areas—hopefully no more than one—were active at a time. However, in practice such hopes were not realized, since very soon after the earliest sensory signals reach the cortex, sensory information begins to reach and perturb ongoing field potential activities within many brain areas (Hupe et al., 2001; Klopp et al., 2000). A more ideal goal for EEG analysis should be to detect and separate activities in multiple concurrently active EEG source areas, regardless of their relative strengths at different moments.

Recently, a new approach to finding EEG source activities has been developed (Makeig et al., 1996) based on a simple physiological assumption that across sufficient time, the EEG signals arising in different cortical source

domains are near temporally independent of each other. This means that measuring the scalp EEG activity produced in some of the source domains at a given moment allows no inferences about EEG activities in the other source domains at the same instant. As we shall see, this assumption is sufficient to separate signals from both physically distant and adjacent EEG source areas—if their contributions to the scalp EEG are largely independent over time. This insight and the resulting algorithms for signal separation that have emerged in the last decade have created a new field within signal processing in general—known in particular as independent component analysis (ICA) or more generally as blind source separation.

In this paper, we review the use of ICA to decompose recorded EEG data into temporally, functionally, and spatially independent source signals. We propose that the success of ICA in identifying both temporally and functionally independent signals in recorded EEG data should gradually change scientific perceptions about the amount and quality of information available in scalp EEG data. In particular, the success of ICA decomposition shows that high-density EEG data should no longer be considered to have intrinsically low spatial resolution, since under favorable circumstances ICA allows separate filtering and analysis of far-field EEG (or equally of magnetoencephalographic, or MEG) signals simultaneously active in dozens of cortical source domains.

2. EEG sources and source independence

The idea that EEG signals originate from temporally independent or near-independent brain processes is consistent with the long observed fact that cortex is organized into compact regions of specialized function. More particularly, connectivity among pyramidal cells is highly skewed toward short (intra-columnar) connections, principally between inhibitory cells that help sustain oscillatory field activity. (Budd and Kisvarday, 2001). In fact, inhibitory cells not only favor short-range synaptic contacts, but they also communicate via electrical gap junction connections (Gibson et al., 1999). These facts alone suggest that a partially synchronous local field activity pattern, once initiated, should spread through a compact cortical area (of unknown extent), much as observed by Freeman (2004) using small electrode grids placed on the cortex of animals.

Since the density of longer-range cortical connections is so low relative to the density of local connections, a neurobiologically plausible working hypothesis for EEG analysis is that over sufficient time, locally synchronous activities within roughly cm-scale patches of cortex are in fact nearly temporally independent of each other and act as single, distinct, temporally independent sources of EEG activity. Alternatively, locally synchronized field activities in a pair of cortical source patches that are densely connected to each other, as for example via corpus callosum, may become synchronized, forming a single effective EEG source. In either case, EEG scalp signals may be modeled as the sum of distinct, phase-independent, and spatially stationary signals from cortical patches (or coupled patch pairs). A third major category of EEG signal sources are non-brain artifact sources including the eyes, scalp muscles, defective or poorly attached electrodes, and ambient line noise, whose volume-conducted activities are also summed in EEG recordings.

While sufficiently dense multi-scale recordings of macroscopic field activity in cortex are still lacking, the physiological plausibility and heuristic accuracy, at least, of the above EEG source model allows the principled application of a new form of *information-based* signal processing.

3. Independent component analysis

3.1. ICA history

The concept of ICA originated in the engineering field of signal processing around 1990 (Comon, 1994). In the simplest terms, ICA algorithms are a family of related methods for unmixing linearly mixed signals using only recorded time course information, e.g., 'blind' to detailed models of the signal sources as required by earlier signal processing approaches. Three early and relatively effective ICA algorithms were JADE (Cardoso and Laheld, 1996), infomax ICA (Bell and Sejnowski, 1995), and so-called FastICA (Hyvrinen et al., 2001). The original infomax ICA algorithm was soon enhanced by introducing natural gradient normalization and an 'extended' mode capable of learning filters for sources such as sinusoids that have sub-Gaussian value distributions (Lee et al., 1999). Jung et al. (2001) reviewed how these and other methods may all be derived from a common information theoretic framework.

The ICA algorithms above only consider the higher-order statistics of the separate data maps recorded at different time points, with no regard for the time order in which the maps occur. The so-called 'second-order blind identification' (SOBI) approach (Molgedey and Schuster, 1994) considers relationships between multiple time points using an autoregressive model in which sources are assumed to have both differing spatial distributions and stable power spectra. A recent internet search readily found freely available Matlab (The Mathworks, Inc.) platform code for at least 22 methods of ICA decomposition (Makeig and Delorme, 2004). So far, detailed comparisons of the results of different ICA algorithms for EEG decomposition are lacking. Here, we illustrate the results of ICA decomposition using natural-gradient extended infomax ICA (Makeig et al., 1997) which we have found to yield consistently good results applied to adequate amounts and quality of data from 31 or more channels. To help investigators apply ICA and time/frequency analysis to their data, and/or to explore other ways of processing and visualizing their electrophysiological data, our group has created an open source Matlab environment,

EEGLAB, that is freely available online at http://sccn.ucsd.edu/eeglab.

3.2. ICA model assumptions

By simple biophysics, EEG sources project near-instantly to and sum linearly at the scalp electrodes. As outlined above, the relatively sparse interconnection of cortical EEG source areas suggests that their activities may, across sufficient data, indeed be near independent. Thus, EEG may be plausibly modeled as a linear mixture of the activities of multiple brain and non-brain sources with (near) independent time courses. A further ICA assumption, that the cortical EEG source domains remain spatially fixed for the duration of the input data, requires careful consideration. For example, invasive optical recordings at the mm scale or below appear to contain moving potential gradients (Arieli et al., 1996), as do some EEG phenomena such as epileptic seizures, very slow spreading depressions associated with migraine headaches (Lauritzen, 1994), and sleep spindles (Massimini et al., 2004). However, the spatiotemporal dynamics of these and other EEG phenomena have not been well characterized, either at the scalp or cortical levels. The relatively stable results of ICA decomposition, both within and between subjects, provide an opportunity to begin such a characterization (Makeig et al., 2002, 2004a, b; Onton et al., 2005).

3.3. The ICA model

The data submitted to ICA are simply the recorded EEG channel data arranged in a matrix of n channels (rows) by t time points (columns). No channel location information at all is used in the analysis. ICA performs a blind separation of the data matrix (X) based only on the criterion that resulting source time courses (U) are maximally independent. Specifically, ICA finds a component 'unmixing' matrix (W) that, when multiplied by the original data (X), yields the matrix (U) of independent component (IC) time courses:

$$U = WX, \tag{1}$$

where X and U are $n \times t$ matrices, and W is $n \times n$. By simple matrix algebra, Eq. (1) implies that

$$X = W^{-1}U. \tag{2}$$

Here, W^{-1} (the inverse of W) is the $n \times n$ component 'mixing' matrix whose columns contain the relative weights with which the component projects to each of the scalp channels, i.e., the IC scalp map. The portion of the original data (X) that forms the ith IC (X_i) is the (outer) product of two vectors, the ith column of W and the ith row of U,

$$X_i = W_i^{-1}U_i, \tag{3}$$

and the whole data (X) are the sum of the (back-projected) ICs (X_i):

$$X = \sum X_i, \quad \text{where} \quad i = 1, 2, \ldots n. \tag{4}$$

Again, each column of the (W^{-1}) mixing matrix represents the relative projection weight at each electrode of a single component source. Mapping these weights to corresponding electrodes on a cartoon head model allows visualization of the scalp projection or scalp map of each source. The source locations of the components are presumed to be stationary for the duration of the training data. That is, the brain source locations and projection maps (W^{-1}) are assumed to be spatially fixed, while their 'activations' (U) reveal their activity time courses throughout the input data. Thus, the IC activations (U), can be regarded as the EEG waveforms of single sources, although obtaining their actual amplitudes at the scalp channels requires multiplication by the inverse of the unmixing matrix (W^{-1}).

The back-projected ICs (X_i) are in the same μV units as the recorded scalp data. However, neither the IC scalp maps nor the IC activations are themselves calibrated. Rather, the original activity units (μV) and polarities (+/−) are distributed between the two IC factors—the IC scalp map and activation time series. For example, reversing the polarities of the activation and inverse weight matrices, then back-projecting the activations through the respective columns of W^{-1} (as in Eq. (3) above) recovers the original component activities in their native μV units. Thus, neither the sign of the scalp maps nor the sign of the activations are meaningful in themselves, but only their product, which determines the sign of the potential accounted for at each scalp channel. However, IC activation magnitudes may be normalized by multiplying each by the root-mean square (RMS) amplitude of the corresponding IC scalp map. The activation units are then RMS μV across the scalp array.

The style of ICA decomposition considered here is termed complete, i.e., a decomposition in which the number of ICA components recovered is the same as the number of channel inputs. Thus, 30-channel data will be decomposed by ICA into 30 ICs, whereas 60-channel data will be decomposed into 60 ICs. Methods for overcomplete ICA decomposition also exist, though these require additional assumptions. An often-posed question is whether there are really 30 or 60 source components in the data, and if not, what are the effects of recording and decomposing different numbers of data channels? Although a full answer to this question is mathematically difficult and perhaps intractable without a complete EEG source model, anatomic considerations suggest the number of near independent sources in the brain may in general be nearly unlimited, although most of them may be very small and thus difficult to resolve from a limited amount of scalp data.

Results of ICA decomposition of high-density (e.g., 128 or more channel) data acquired from normal subjects during performance of cognitive tasks show that some dozens of temporally and dynamically distinct EEG sources are large enough to be separated into components with physiologically interpretable scalp maps and activations. Applying ICA decomposition to fewer data channels must result in some or all of the extracted components summing activity from more than one underlying source. However, in this case, ICA should efficiently arrange for even these mixtures to have minimal common or mutual information. As readily shown in simulations (unpublished observations), larger resulting components will tend to have scalp maps compatible with physiologically simple source distributions and temporally distinct time courses which will outweigh any remaining source signals not temporally and spatially consistent with the large component properties.

3.4. Component source modeling

ICA, applied to data of sufficient length and quality, typically returns many components that each: (1) account for an appreciable amount of variance in the original data and (2) whose scalp projection resembles that of a single equivalent dipole located in the model brain cavity. An equivalent dipole is a purely mathematical concept (imagine an impossibly small battery) located within or more likely just beneath an area of synchronously active cortex that contributes to scalp recordings. By biophysics, coherent activity across a small patch of cortex will have a near-dipolar projection pattern on the scalp. To estimate the location of the equivalent dipole for an IC scalp map, therefore, we can apply standard inverse source modeling methods to the IC map. Using a simple though not completely accurate spherical or standard boundary element head model, we typically find best-fitting equivalent dipoles for 20 or more IC sources from 64-or-more channel data with less than 15% residual variance between the actual IC scalp map and the model projection of the equivalent dipole to the same electrode montage. Given enough and sufficiently clean data from enough scalp channels, the residual variances of simple equivalent dipole models for independent components can often be extremely low ($<2\%$).

This result does *not* by itself imply that the activity included in each dipolar IC must necessarily be generated within one roughly cm^2-sized cortical patch, as envisioned in the model we have outlined above. Rather, the facts of cortical connectivity, specifically the very high relative density of local connections between neurons, particularly inhibitory neurons, make our hypothesis that ICs with dipolar scalp maps represent synchronous activity within a cortical patches the most parsimonious and physiologically plausible explanation for their 'blind' separation by ICA.

Several issues need to be faced by those working with ICA decomposition, however. First, not all ICs have dipolar scalp maps. In particular, ICs accounting for the least signal variance often have 'noisy' appearing scalp maps, and may not be reproduced even in repeated decompositions of the very same data (e.g., if the same data are presented to the algorithm in different orders). This 'unresolvable noise' subspace may represent a portion of the recorded signals that does not fit the ICA model, or may be a result of numerical limitations inherent to finite-precision recordings and computations.

Second, as introduced above, some EEG phenomena do indeed seem to violate the ICA model assumption that the physical sources of EEG recordings are spatially stationary throughout the data. For example, sleep spindle activity seems to flow across the scalp in varying spatial patterns, as does *K*-complex activity during Stage II sleep (Massimini et al., 2004). Slow spreading depressions, associated with migraine headaches, and spreading epileptic seizures are other EEG phenomena that may not be captured in a single IC (though they might be parsimoniously separated from other EEG sources by ICA into a subspace of a few ICs whose summed activity account for them). Direct cortical recordings also suggest that at the millimeter and smaller scales, waves of field activity may flow across small parts of the cortical mantle, though at the cm-scale of EEG recordings, these flow-pattern effects may (or may not) be indistinguishable from activity in static source distributions.

Finally, in some or all cases, the simplest interpretation of dipolar scalp maps as reflecting activity within a single cortical patch may be too simple. Many possible (though not necessarily plausible) distributions of partially synchronous activity over larger portions of cortex could mathematically account for each dipolar IC map. The synchrony of activity within such extended cortical areas, however, must have some biophysical mechanism. Unless such a mechanism was known in advance, only direct evidence from dense electrode grids could make such models plausible. To date, few, if any, experimental recordings have been conducted using both sufficiently dense and spatially extended cortical electrode grids. Typical estimates of the spatial coherence scale of cortical field potentials are on the order of a cm or less (Bullock et al., 1995).

Recently, however, a nice albeit still indirect piece of evidence supporting the interpretation of dipolar ICs as activity projecting from a single cortical patch has been published in a simultaneous EEG and fMRI study of Debener et al. (2005). ICA decomposition of 32-channel EEG data (after minimization of unique fMRI-induced artifacts) produced, for nearly every subject, an IC with a near vertex-centered scalp map. Correlation of the single-trial error-related activity of these ICs with the hemodynamic BOLD activity forming the fMRI signal revealed a significant relationship only in a small area of medial cortex located directly under the mean IC scalp map maximum, and near the corresponding mean equivalent dipole. We have shown a nearly identical cluster of IC sources to produce theta burst activity following error

feedback in a continuous performance (two-back) task (Fig. 5, Onton and Makeig, 2006).

The scalp maps of some ICs are consistent not with a single-dipole source model, but with dual dipoles. In our experience, these typically can be well modeled by two dipoles bilaterally symmetric in location if not orientation. These sources may arise in two ways—either as distinct synchronous responses in left and right visual cortices in response to central visual stimuli, and/or as ongoing EEG activity synchronized by bi-directional coupling through dense callosal connections. If this is the case, then it might be possible to observe ICs generated in other pairs of cortical areas bi-directionally coupled by dense white matter tracts, for example the arcuate fasciculus that connects frontal cortex to the ipsilateral temporoparietal junction.

Interestingly, we have observed that, in general, the ICs that tend to account for the largest part of cognitive ERP and event-related spectral perturbation (ERSP) features are more nearly dipolar, whereas noisy-appearing, non-dipolar components typically tend to contribute less to overall EEG dynamic changes time-locked to significant task events. This is consistent with a working assumption that dipolar ICs are those generated in one (or two) patch(es) of cortex, whereas non-dipolar ICs, if not accounting for non-brain artifacts, may account for mixtures of small source processes or aspects of processes not fitting the spatial source stationarity assumption. Modeling ICs as dipole sources also gives a convenient way to assess the distribution of EEG source locations in a single subject, or the spatial homogeneity of a set of ICs from a group of subjects. When available, more advanced methods of inverse source modeling, particularly those incorporating structural information from subject magnetic resonance images, should be more adequate and give more detailed information about cortical EEG dynamics.

3.5. Practical considerations

Two important considerations dictate the quality of ICA decomposition for a given dataset. First, the number of time points of n-channel data used in the decomposition must be sufficient to learn the n^2 weights of the unmixing matrix. If the numbers of electrodes and independent cortical sources are large, as in typical EEG data, the number of data points used in ICA decomposition should be at least some multiple, k, of n^2. To decompose large numbers of channels (e.g., 256), k may need to be 20 or larger, meaning (for 256-channel data and $k = 20$), ICA decomposition will require $256^2 \times 20 = 1,310,720$ or more data points. At a 256-Hz sampling rate, this would require decomposing at least 85 min of data.

For smaller numbers of channels, the amount of data required is generally much smaller. For example, for one quarter the number of channels (64), only one-sixteenth the amount of data (less than 6 min) would be required to give the same k. However, ICA decompositions of still more data ($k >> 20$) tend to be more regular, with more dipolar component maps. Thus, the general rule for ICA decomposition is that more data are better—so long as one assumes that the EEG source locations have not changed. For example, jointly decomposing data from awake and sleeping conditions might not be optimal, if the EEG source locations (not their activities) differ in these portions of the data. Further research on this issue is appropriate and ongoing.

3.6. ICA versus PCA

When not enough data are available for complete decomposition, as when separate decompositions of different data conditions are to be attempted for some purpose, there are two choices available. One approach is to ignore some of the data channels; another is to reduce the dimensionality of the data to a smaller number of its largest principal components—in mathematical terms, to a principal component subspace. This possibility raises issues for many researchers as to the relation between principal components and independent components. Although both ICA and principal component analysis (PCA) are linear decompositions obeying Eqs. (1)–(4) above, their mathematical objectives—and thus their biological interpretations—are quite different.

The goal of PCA (for which it is optimal) is to find temporally orthogonal directions in the joint channel data space (corresponding to necessarily orthogonal scalp maps) that each successively explain as much of the remaining data variance as possible. (Note: It is also possible to apply PCA with the roles of time and space exchanged.) The goal of ICA, on the other hand, is to find directions in the joint data (corresponding, typically, to *non*-orthogonal scalp maps) whose activities are as distinct from one another as possible, meaning that their signals have the least possible mutual information. Minimizing mutual information implies not only decorrelating the component signals, but also eliminating or reducing their higher-order joint statistics. This stronger statistical goal allows ICA to relax the orthogonality of the component scalp maps, a physiologically implausible constraint, and to separate phase-uncoupled (or nearly always uncoupled) activities generated in spatially fixed cortical domains (or non-brain artifact sources).

If the scalp maps associated with activities in such domains are not orthogonal (as is nearly always the case), PCA will combine portions of their activities into one or more principal components, rather than separating them into different components as in ICA. Thus, if the recorded data are in fact the sum of (nearly) independent signals from spatially fixed and distinguishable sources, PCA will *lump*, and ICA will *split* the source activities across resultant signal components. PCA attempts to *lump* together activity with as much variance as possible into each successive component, irrespective of whether this activity comes from one or many temporally independent

(or near independent) sources. ICA, on the other hand, insofar as possible *splits* the activities of such sources into distinct components.

PCA reduction may be efficient for data compression, but the flexibility allowed ICA decomposition to find components with spatially overlapping scalp maps is crucial for finding *physiologically* distinct EEG components. In practice, when not enough data are available to perform full-rank ICA decomposition, it might in some cases actually be preferable to reduce the number of channels in the dataset rather than reducing the data dimensionality by retaining a principal subspace of the data using PCA. This is particularly true if any of the channels exhibit periods of spurious noise due to a poor connection with the scalp. If these unreliable channels are weighted with a non-zero value in any IC scalp map, then the associated activations during the periods of noise will be inaccurate.

3.7. Two classes of EEG artifacts

The universal rule of signal processing, '*Garbage in, garbage out*' (GIGO), applies to ICA decomposition as well. What constitutes 'garbage' data for ICA? For this question, it is important to distinguish two types of EEG artifacts. Typical artifacts arising from eye movements, eye blinks and muscle tension have stereotyped scalp projections (since eyes and muscles remain at the same head locations), although a repertoire of eye movements in different directions, or blinking each eye separately, etc., can add more than one stereotyped scalp map pattern to the data and therefore require more than one IC to separate from the data. Another class of artifacts, however, is more problematic for ICA. These include movements of the electrodes on the scalp arising from large muscle movements or external sources (e.g., tugs on the electrode cables, etc.). Such artifacts may quickly introduce dozens or hundreds of unique scalp maps in the recorded data, each of these maps being technically independent of any other data source and thus requiring a separate IC, leaving few ICs available for capturing actual brain sources.

The solution to this problem is to carefully prune the data of non-stereotyped artifacts before decomposition. In practice, one may adopt a multi-stage approach, first pruning the data, then applying ICA, then pruning the data again by looking for and removing moments when the IC activations move together (contra the ICA objective), then applying ICA again. In the future, this approach might be automated, possibly allowing for quicker and cleaner ICA decompositions.

4. Sample application to the study of normal aging

4.1. Audiovisual attention-shifting

The advantages of applying ICA to EEG data may be illustrated with preliminary results from an experiment conducted by two of the authors (JT and MW) to assess the changes in EEG activity underlying aging-related changes in cued attention switching. Thirty-one subjects aged 18–85 were asked to press a button every time a target blue square or high tone was presented in a currently attended modality (visual or auditory). Only one of these sensory modalities was to be attended at a time. During presentation of interleaved auditory and visual stimulus streams, the modality to be attended was switched repeatedly by irregularly spaced audiovisual word cues ('look' or 'hear'). Closely related (distracter) stimuli in both modalities made target detection in each modality difficult, requiring close selective attention to the currently attended modality for correct identification of presented targets. For the data analysis, subjects were divided into two groups of 15 younger (mean $\pm$ st.d.: 26.2 ± 6.37 years) and 16 older (aged 65–85, mean $\pm$ st.d.: 70.9 ± 5.87 years) subjects. Older adults performed this task as well as younger adults. There were no significant differences in accuracy or response time in any condition. Therefore, age-related differences in patterns of electrophysiological activity do not reflect performance-related factors such as number of correct trials.

4.2. Target-evoked response differences in younger and older adults

Studies comparing brain activity of younger and elder adults have revealed various differences in cerebral blood flow suggesting, for example, deficits in stimulus response (Ross et al., 1997) and target detection (Madden et al., 1999) as well as altered motor area activation (Mattay et al., 2002) in elderly subjects. In our study, the primary measure used to assess age group differences was the ERP time-locked to target stimuli cueing a button press. Figure panels 1A and B show the grand average ERPs for both groups at all channels (colored traces) plus the ERP distribution on the scalp (in μV) at several trial latencies. As these plots show, the target ERP of younger subjects exhibited larger mean negative and positive peaks near 180 and 400 ms, respectively. In addition, the mean scalp distribution of the broad 'P300' positivity in the grand mean ERP waveform for the older subject group was more spatially diffuse, encompassing the frontal as well as the posterior scalp electrodes.

Figure panels 1C and D show the data trials that were averaged to create the ERPs for an electrode site (Pz) located over central parietal cortex. Here, instead of averaging all the trials to form a single trace, the scalp potential time series in each trial was converted into single color-coded horizontal traces. The trial traces were then stacked vertically in order of subject response time in each trial, and smoothed across neighboring trials using a moving average, forming an 'ERP image' plot (Jung et al., 1999; Makeig et al., 1999). ERP image plots can be used to visualize features common to some or all of the single trials. The two ERP-image plots in Fig. 1 show that the

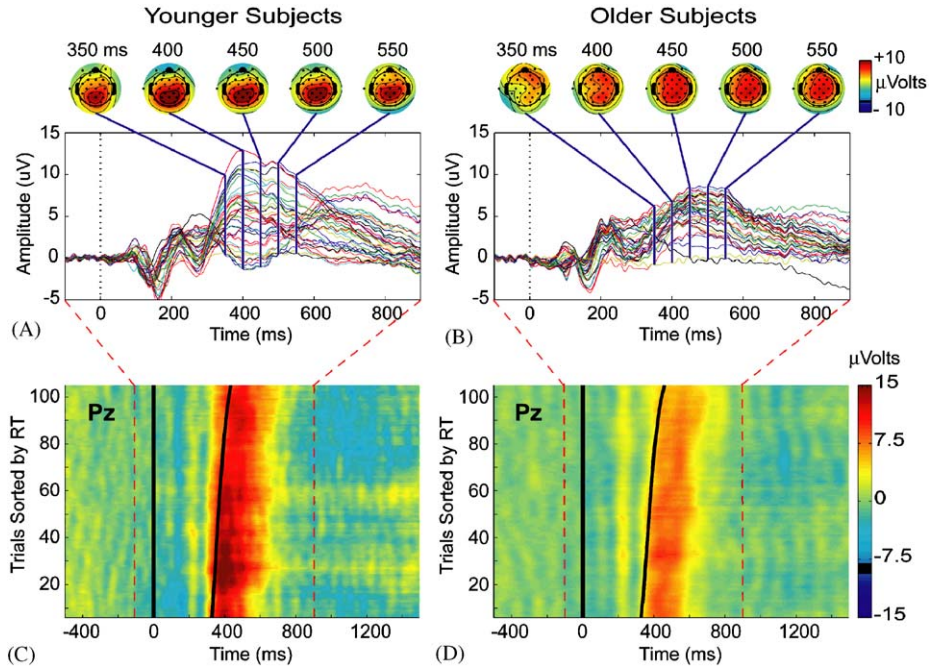

Fig. 1. ERP results of an attention switching study. (A) Grand-mean target evoked response in the younger adult subject group (16 Ss). Each colored trace gives the ERP waveform at one of the 36 scalp channels. Scalp maps show the scalp topography of the ERP at the indicated latencies. The halos around the cartoon heads display estimated potential distribution below the ears (i.e., the bottom half of the head 'sphere'). The scalp maps reflect the typical young adult 'P300' peak topography peaking in this experiment near 400 ms. (B) The same grand mean response for the old adult subject group (17 Ss). Note the anterior spread of the 'P300' positivity. (C) ERP-image plot of single-trial time courses at central parietal site Pz, trials sorted by the latency of the subject button press (curving black trace) following stimulus presentation (vertical black trace) with shortest and longest response time trials excluded. (D) A similar ERP-image for the older subject group. Note the uniformly smaller positivity for all response time ranges in the older subject group.

'P300' positivity in single trials at this central parietal channel was consistently larger in the younger adults, following both shorter (bottom) and longer (top) response times.

ERP-image plotting is a simple but powerful technique for examining trial-to-trial similarities and differences in EEG dynamics that are not revealed by average ERP features. Sorting the trials by response time (RT) reveals that the positive-going brain activity evoked at site Pz about 300 ms after target stimulus onsets is not strictly time-locked to the stimulus, but also is time-locked, in part, (most clearly in older subjects) to the subsequent subject button press.

4.3. ICA decomposition

Separately applying ICA decomposition to the unaveraged, artifact-pruned 36-channel EEG data from each subject returned 36 independent components (ICs) per subject, each IC composed of a scalp distribution or map plus an activation or activity time course throughout the input data. All 36 IC scalp maps from a typical decomposition of data from a young adult subject are shown in Fig. 2. In general, the IC scalp maps may be said to fall into one of four categories: (a) cortical brain sources (here including ICs 1, 3, 4, 7, 11, 12, 14, and 15); (b) physiological artifacts including eye movements and eye blinks (IC2), muscle activity (IC31), cardiac pulse artifacts (IC10); (c) external artifacts including line noise; and (d) spatially irregular components of unknown origin (here including ICs 17, 25, 30, 33–36), the most ambiguous of these categories. The signal strength of most components in this latter category are small and possibly represent mixtures of multiple source areas whose activities were not stereotyped and/or stationary enough to be resolved as single sources given the number of recorded data channels.

We offer these heuristic categorizations from considerable experience with studying ICA decompositions. To categorize independent components in this way, the component maps, equivalent dipole, or other types of source location estimates (if available), as well as time courses, power spectra, and (when relevant) ERPs should be examined carefully.

In Fig. 3, the thick black traces show the envelope (maximum and minimum channel values) of the mean target ERP for the same subject. The shaded region indicates the envelope of the portion of the total scalp data accounted for by the four components whose scalp maps and activity envelopes contribute the most to the P300 portion of the ERP (colored traces). As the figure shows, four ICs make sizeable contributions to the P300 peak near 0.4 s. Three of these (ICs 4, 11, and 12) project most strongly to the posterior scalp (left, center, right), while the maximum projection of the fourth component (IC1) is

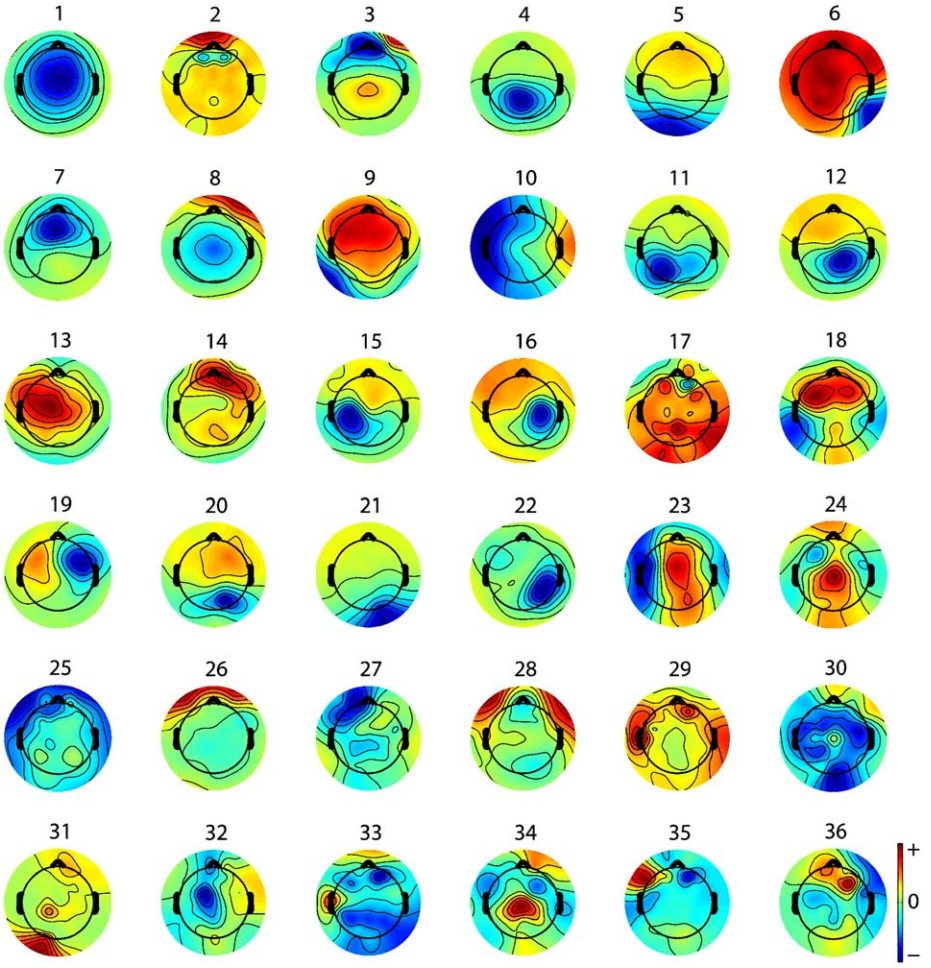

Fig. 2. Interpolated scalp maps of the 36 independent components (ICs) obtained by extended infomax ICA decomposition of the unaveraged, concatenated 36-channel target trials from one younger adult subject. The maps have been individually scaled to their maximum absolute value (green is 0). Components (ICs) sorted by reverse variance accounted for. For example, IC2 accounts for eye blinks, IC15 and IC16 for left and right mu rhythms, ICs 34-36 for physiologically unresolvable noise.

anterior to the vertex. Such 'envelope' plots may appear confusing on first viewing, but to experienced eyes they give a quick overview of the major ICs contributing to an average ERP.

4.4. Component clustering

It is a simple procedure to choose the largest component contributors to an ERP for a single subject (as in Fig. 3), but to conduct normative research, we need to identify sets of *similar* components across many subjects. The issue of component clustering across subjects represents one of the major differences in methodology from traditional EEG analysis. Typically, EEG data are compared across subjects by equating scalp locations, whether or not a given scalp location (say Pz) receives the same combination of source activities in every subject. In general, this will depend on the consistent presence, relative strengths, and net cortical orientations of the EEG sources projecting to Pz in each subject, thus making the signal at Pz implicitly variable across subjects. In contrast, the scalp maps and activations of ICA components are explicitly variable across subjects, providing the challenge of grouping together only truly similar source activities across subjects, using carefully considered clustering strategies.

While component clustering may seem a cumbersome and difficult step, it can yield more homogeneous collections of activity across subjects than simply equating single-channel data by scalp location. Because every brain is uniquely folded, two equivalent EEG sources in different subjects may project to the same electrode location with variable strengths. For example, a cortical source in one subject may be located on a cortical gyrus, while another subject's functionally equivalent source is located at the edge of a nearby sulcus. Despite the close proximity of these two sources, their scalp projections must be quite different.

As an illustration, consider IC7 in Fig. 2, which projects most strongly to the frontal midline scalp and whose equivalent dipole in medial frontal cortex is oriented radial

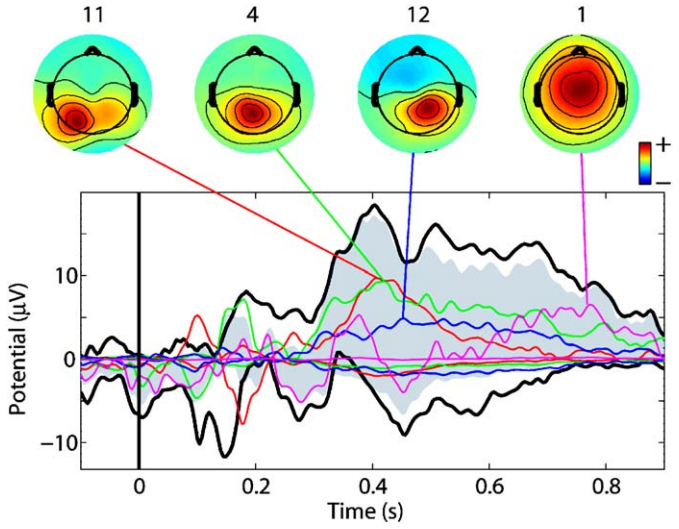

Fig. 3. Independent component contributions to the mean target ERP for the subject whose IC scalp maps were shown in Fig. 2. The vertical black trace at time 0 indicates moment of stimulus onset. The thick black traces show the envelope (most positive and negative channel values) of the whole 36-channel ERP for this subject. Colored traces show the envelopes of the contributions of the four independent components accounting for most of the P300; their scalp maps are shown above the data. The maps are each connected to their respective data envelopes by a colored line attached to their time point of maximum projection to the ERP. The polarity of the individually scaled scalp maps are reversed from those shown in Fig. 2, reflecting the momentary polarity of their mean scalp projections at the latencies shown.

to the scalp surface. In contrast, IC19 has a quite similar equivalent dipole location but an orthogonal orientation parallel or tangential to the skull surface, and therefore a distinctly different projection. Data collected at a frontal midline scalp electrode would most certainly *not,* in this case, record the projections of all sources just below the electrode, since any number of sources (e.g., IC19 in Fig. 2) may be tangential and thus project with zero strength to an overlying electrode. Thus, given inevitable intersubject differences in cortical topography, an electrode at the very *same* scalp location in several subjects will likely detect *different* mixtures of EEG sources. On the other hand, if ICA components from different subjects can be matched accurately, it may be assumed with much greater certainty that these two sources represent analogous and comparable activities.

Because of intersubject variability in the size and/or orientation of the numbers and types of EEG sources large enough or distinct enough to be detected at the scalp and isolated by ICA, all subjects may not contribute a component to every cluster. (This difficulty, though rarely acknowledged, is also present in electrode-level analysis.) Even if all of the sources contributing to a given scalp location are *analogous* across subjects, the relative strengths of the source projections may not be comparable, thus skewing the ERP size and features in unreliable ways. Given the mixing of many signals at any scalp channel, it is impossible to determine *which* sources are absent, reduced, or enhanced in a single subject by examining single-electrode data alone. Without decomposing the recorded signals into independent spatial-temporal sources, it simply cannot be known what sources compose an ERP, and therefore scalp-channel ERPs recorded from several subjects cannot be assumed to be comparable.

4.5. ERPs and component clusters

One straightforward way to cluster components across subjects is to choose components contributing the most energy to an ERP feature of interest (Makeig et al., 1997, 2002). This may be an effective approach for quite stereotyped evoked activity when the relevant time window is known and the ERP is composed mainly of a single ICA component (Debener et al., 2005). Another way to cluster components across subjects is to group components with similar scalp maps, assuming that similar components have comparable scalp projections. This selection can be executed by visual inspection or by testing for scalp map correlations. For example, Fig. 4 displays scalp maps and event-related activities of two component clusters from the aging study, as selected by visual inspection. The component scalp maps for these clusters project maximally to frontal-central and mid-parietal regions, respectively.

Fig. 4A and B show mean scalp maps, grand mean ERP contributions, and trial-by-trial ERP images for two clusters of frontocentral components from the younger (A) and older (B) subject groups. The filled blue portions represent the envelope of the back-projections for each IC cluster. While the cluster mean scalp maps are similar, and the ERP envelopes of the two clusters seem analogous, the frontocentral cluster contributed $\geqslant 3\,\mu V$ to the ERP only in the older subjects' data. The ERP images below show that this contribution in older adults followed the button presses. The equivalent cluster ERP image for the younger subject group contains, instead, a brief negativity following button presses. In contrast, the mid-parietal clusters (Fig. 4C,D) exhibit an opposing effect, the response-locked ERP contribution in younger subjects being considerably stronger than that in older subjects. Without decomposing the data into independent component sources, it would be difficult or impossible to determine the cortical origin of these apparent group ERP differences.

Two other clusters of left and right somatomotor area ICs, respectively, shown in Fig. 5, reveal further possible differences between responses of younger and older subjects which would likely be hidden in scalp-channel analysis. In the younger subject group, a left somatomotor (mu rhythm) cluster, produced a stronger positive contribution to the ERP immediately following the motor response (Fig. 5A,B), followed by a broad negativity (Fig. 5A).

During the same post-motor response period, right somatomotor area component clusters (Fig. 5C,D) produced near-equal ERP contributions in both the young and

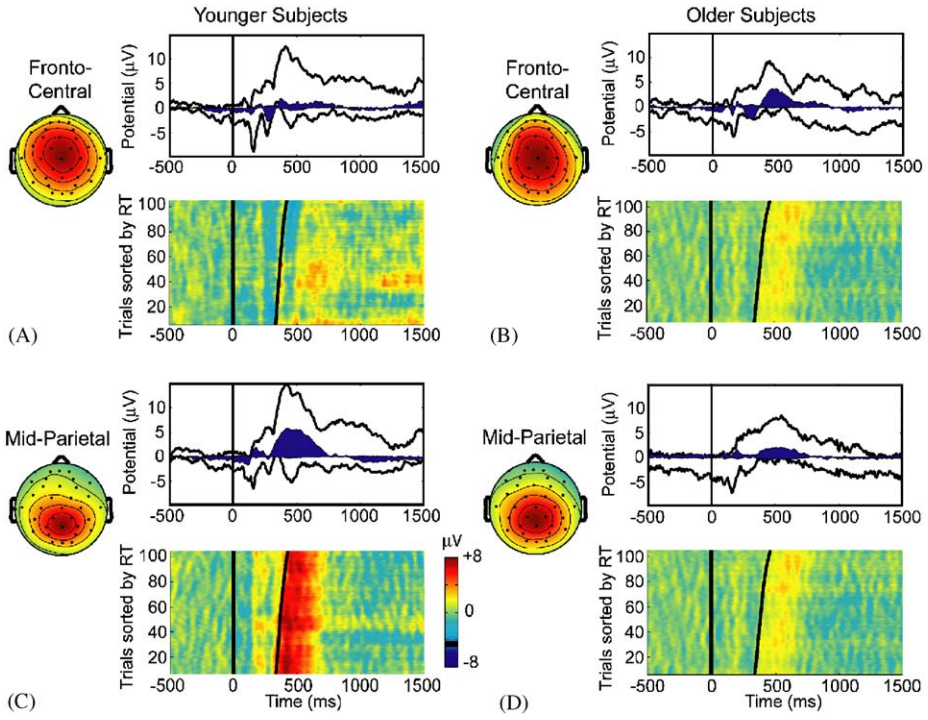

Fig. 4. Mean scalp maps, grand-mean ERP contributions, and RT-sorted ERP images (lower panels) for clusters of fronto-central and parietal-midline components from younger and older subjects, respectively. (A) Fronto-central cluster for younger subjects. (B) Fronto-central cluster for older subjects. (C) Mid-parietal cluster for younger subjects. (D) Mid-parietal cluster for older subjects. The ERP plots (upper right) in each panel show the envelopes of the grand-mean channel ERP for the total data (black traces) and the back-projections of the respective IC clusters (filled blue regions).

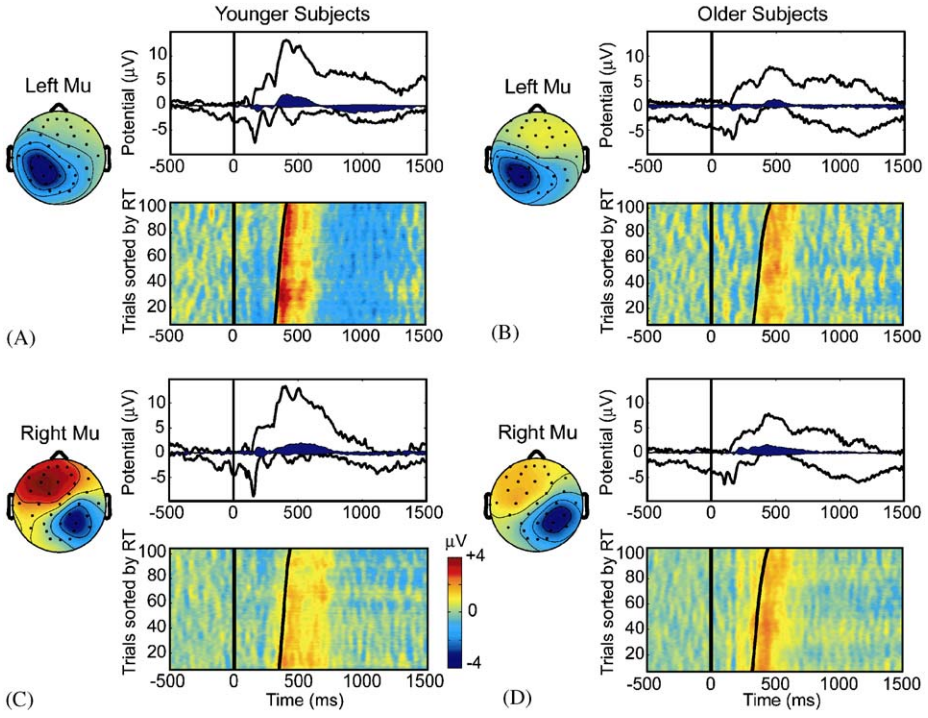

Fig. 5. Mean scalp maps, grand mean ERP contributions, and RT-sorted ERP images for two clusters of components projecting most strongly to left or right central-parietal scalp, respectively, and exhibiting strong 10 and 20 Hz mu rhythm activity (visible as randomly hatched patterns in the ERP image pre-stimulus baselines) for the two subject groups.

old subject groups. Thus, the relative ERP contributions of left and right somatomotor sources appear more equal in older subjects, whereas younger subjects showed a left-dominant pattern (Fig. 5A,C). A model suggested by Cabeza (2002) proposes age-related reductions in brain asymmetry during cognitive processing. In line with

Cabeza's model, these plots show nearly equal bilateral response in older adults compared to a strongly lateralized response in younger adults. In the scalp-channel ERPs (Fig. 1), the prominent left somatomotor area negativity in younger subjects (Fig. 5A) is hardly apparent, and the observed differences in bilateral activation between the subject groups would have likely been overlooked before the independent cortical signals were separated from the scalp channel data by ICA.

These plots are also consistent with age-related decreases in attention modulated neural specificity. This functional 'dedifferentiation' or loss of specialized processes has been observed in a number of studies and likely reflects structural and neurochemical brain changes associated with normal aging (Chen et al., 2002; Grady et al., 1994; Li and Lindenberger, 2002; Li et al., 2001). The increased positivity over frontal regions in the older adults is commonly reported (Friedman et al., 1997; Pfefferbaum et al., 1984; Polich, 1996, 1997) and can be seen in both the averaged ERP and the IC back-projections plotted in Fig. 4. However, the underlying loss of processing specificity demonstrated in the IC clusters of both Figs. 4 and 5, cannot be observed in the averaged ERP. An fMRI study of these same subjects in this same task also suggests age-related loss of processing specificity (Townsend et al., in press), but does not provide important temporal information available in the component analysis. However, as always, before the preliminary observations above may be accepted as fact, the statistical reliability of the apparent group differences need to be evaluated.

4.6. Event-related rhythmicity

In Fig. 5A and C (and confirmed by further examination of single trials for left and right somatomotor (mu) components), the post-response activation in younger subjects can be seen to include a two-cycle ~5-Hz theta burst (two positive peaks separated by about 200 ms), riding on a broader positivity. Interestingly, the theta-band burst feature is not apparent in the equivalent component clusters for older subjects (Fig. 5B,D), an observation we plan to explore further. The young-adult response-locked theta burst pattern was discovered and analyzed in detail for a different experiment (Makeig et al., 2004a). More detailed analysis of age-related changes in this feature, if confirmed, might be relevant for understanding age-related changes in behavior. Again, the apparent absence of the post-response theta burst in older adults would likely have been missed during analysis of the scalp channel data alone.

4.7. Statistics on component clusters

Naturally, the reality of apparent group differences suggested by differences in cluster mean measures, such as Figs. 4 and 5, cannot be assumed without statistical testing, either parametric (e.g., *t*-tests) or non-parametric (e.g., permutation-based tests). The power of the statistical testing should in turn be governed by the accuracy of component cluster selection. Here, preliminary clustering was performed by eye, based on similarities in scalp maps and ERP contributions. However, fairly similar scalp maps may be produced by spatially coherent field activity in well-separated cortical patches and may well express highly disparate time-courses. In our experience, including various measures, including time/frequency information as well as dipole location, into the clustering algorithm helps to identify homogenous groups of components across subjects. For this reason, we have recently released component clustering software in EEGLAB to encourage researchers to explore use of different feature combinations for clustering ICs, and to assist researchers in applying statistical testing to the resulting component clusters (EEGLAB v5, http://sccn.ucsd.edu/eeglab/).

5. Further challenges

Though the preliminary results presented above demonstrate some of the basic capabilities of ICA decomposition of EEG data, they only scratch the surface of the possible questions and measures that can be explored using the ICA approach. Below we will discuss some issues relevant to ICA-based analysis.

5.1. Component clustering methods

Determining the best way to identify clusters of equivalent components across subjects is a difficult problem, and there may not be one best method for all situations. However, it is worth discussing some alternatives to the visual selection of components by their scalp maps or ERP contributions (as in Figs. 4 and 5 above). One possibility is a multiple-measure approach including any number of these IC measures: scalp maps, dipole model locations, ERP contributions in one or more conditions, ERSPs or phase-locking, component cross-coherence, etc. Using many measures may help ensure that the resulting component clusters are accurately matched. Ideally, clustering by activity alone, rather than by spatial location, should yield clusters whose modeled locations would all lie within a constrained cortical region, thus corroborating the cluster homogeneity. Using non-spatial measures instead of a purely or partly spatial approach allows components in the vicinity of a component cluster to be excluded from the cluster if the IC does not behave similarly to the other clustered ICs.

Clustering on multiple measures requires construction of a common distance measure between ICs based on the measures used, plus a choice of clustering method. In practice, standard PCA reduction and *k*-means clustering methods can give satisfactory results. However, the optimal choice of clustering method and which ICs measures are best to include are not uniquely defined and may depend on the task, the homogeneity of the subject population, and

the particular research question. For example, if two ICs were to act similarly during 'correct response' trials, but differently during 'wrong response' trials, the two ICs likely represent functionally different processes. In this case, including only 'correct', only 'wrong', or both trial types in the clustering measure could potentially result in three different group assignments for these ICs.

The goal of IC clustering is to group highly similar activity together from as many subjects as actually express the relevant IC(s) and their characteristic activity. Using this strategy, some or all clusters may not include contributions from all subjects. This could reflect inefficiency in the decompositions, for example if too few channels or too few data points were used, or else ways in which the subjects are actually physiologically or behaviorally dissimilar. Defining multiple clusters that represent relatively common EEG activities across subjects, then examining the nature of subject differences within and between clusters, should allow detailed exploration of subject EEG differences, whose possible genetic linkages might now also be investigated.

5.2. Time/frequency modeling

Decomposing the normalized activations of ICA components into frequency-domain components by short-time Fourier or wavelet transforms provides an alternative method for visualizing event-related EEG dynamics that can highlight aspects of the dynamics not available in ERPs, since ERPs capture only that portion of the channel or component data that is *phase-consistent* at latencies relative to the time-locking events. Because EEG power is higher at lower frequencies, visualization of event-related perturbations in spectral power is best accomplished by dividing mean power in each time/frequency bin, relative to the time-locking events, by the mean baseline power spectrum, giving an event-related spectral power or perturbation (ERSP) image (Makeig, 1993).

ERSP measures, however, consider only spectral power and ignore trial-to-trial variations or consistencies in spectral phase. Two other time/frequency measures are useful for quantifying these: inter-trial coherence (ITC), measuring the degree of consistent phase-locking in short time-frequency windows time-locked to events, and event-related coherence (ERC) between two component processes. The latter measure may seem contradictory, since independent components are by definition not coherent. However, in reality, infomax ICA finds components that are *maximally* independent, meaning that components whose activities partially collapse in only a small percentage of the training data can still be separated into maximally independent data components. ERC between maximally independent components can therefore be measured, and may be of interest (Makeig et al., 2004a). For more discussion of time/frequency measures applied to independent component activations, see Makeig et al. (2004b).

5.3. Trial-to-trial variability

While the mean time/frequency measures mentioned in the preceding section can reveal information about event-related EEG dynamics that ERP measures neglect, they are also averages and thus ignore trial-to-trial activity differences. The idea behind averaging is that event-relevant brain dynamics that are consistently time-locked to a class of events will be recovered by response averaging, while other processes unaffected by the same events will be filtered out by phase cancellation. The response averaging approach, applied blindly to a set of single-trial data, does tend to reveal dominant, time- and phase-locked activity (with respect to the time locking trial events) that is consistent across trials, but it ignores the possible *relevance* of inter-trial variability due to trial-to-trial variations in cognitive processing. As an example, we have recently reported on a frontal midline component cluster that shows several distinctly different patterns of time/frequency activity (Onton et al., 2005). One of these was specifically linked to the number of letters held in memory, while other modulatory factors were presumably associated with other conditions common to each trial subset. In the future, it may be possible to link replicable time/frequency patterns across subjects to specific behavioral demands.

5.4. Component statistics

As in all areas of biological science, testing the statistical significance of observed differences in data measures employed is of utmost importance. In the field of traditional EEG/ERP research, parametric statistics have usually been employed to determine whether an ERP peak is consistently above or below baseline (zero), or consistently different from an equivalent portion of another ERP trace. The same techniques can be applied to analysis of ICA activations. However, we prefer to use permutation-based techniques to create a sample set of simulated values from a large number of derived or 'surrogate' data distributions (Blair and Karniski, 1993). To construct these surrogate distributions, most often some feature of the original data is perturbed or shuffled across trials and/or subjects, so that the resulting surrogate data distribution will have all features of the original data except the effect to be tested.

For example, to test whether or not an ERP trial average at a given latency is consistently positive, one could create surrogate ERP measures, each averaging the same trial data after randomly flipping or permuting the signs of the individual trial values. The 95th or 99th percentile, for example, of the surrogate measure distribution could then be used as a significance threshold to determine whether the observed value in the average of the original trials can be considered unexpected under the given statistical hypothesis and therefore significant. Similar strategies may be employed for time/frequency measures. Their advantage is that they are non-parametric, not relying on

an a priori model of the data distribution but adapting to the actual distribution of data values, whatever its form or cause.

6. Functional electromagnetic brain imaging

This chapter is only an introductory review of a powerful and still new approach to EEG data analysis that may in the future help extricate the field of EEG research from its reputation of having 'poor spatial resolution' relative to other brain imaging methods. ICA provides a method that retains all the temporal resolution of EEG recording while adding more spatial specificity by separately identifying the activities and scalp projections of up to dozens of concurrently active and temporally distinct EEG sources. ICA methods are not only effective for removing artifacts from EEG data (Jung et al., 2000a,b), but also for direct analysis of distinct EEG components and, arguably in many cases, cortical source activities. Given the rich amount and complexity of the data produced by high-density EEG recordings, and the rapidly evolving power of readily available computational and methodological resources, performing true dynamic brain imaging using EEG data, recorded either separately from or simultaneously with MEG data, is beginning to be regarded as a worthy goal and promising possibility by the cognitive neuroscience community. The unique advantages of EEG recording are its true temporal resolution, high dimensionality, ability to record simultaneous activities in all parts of cortex, and the behavioral flexibility it affords subjects, who may be recorded in nearly any body position without rigidly fixing the head. These qualities should in the near future bring EEG again into the forefront of dynamic brain imaging methods, both in basic and applied research as well as in clinical and workplace applications.

Acknowledgments

This report was supported by the Swartz Foundation (Old Field, NY) and by funding awarded to J.T. from the US National Institutes of Health, National Institute on Aging, 5RO1-AG18030. The authors gratefully acknowledge the contributions to this research from long-term collaborations with Terrence Sejnowski, Tzyy-Ping Jung, Arnaud Delorme, and Eric Courchesne.

References

Arieli, A., Sterkin, A., Grinvald, A., Aertsen, A., 1996. Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273, 1868–1871.

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

Blair, R.C., Karniski, W., 1993. An alternative method for significance testing of waveform difference potentials. Psychophysiology 30, 518–524.

Budd, J.M., Kisvarday, Z.F., 2001. Local lateral connectivity of inhibitory clutch cells in layer 4 of cat visual cortex (area 17). Experimental Brain Research 140, 245–250.

Bullock, T.H., McClune, M.C., Achimowicz, J.Z., Iragui-Madoz, V.J., Duckrow, R.B., Spencer, S.S., 1995. EEG coherence has structure in the millimeter domain: subdural and hippocampal recordings from epileptic patients. Electroencephalography and Clinical Neurophysiology 95, 161–177.

Cabeza, R., 2002. Hemispheric asymmetry reduction in older adults: the harold model. Psychology and Aging 17, 85–100.

Cardoso, J.-F., Laheld, B., 1996. Equivariant adaptive source separation. IEEE Transactions on Signal Processing 44, 3017–3030.

Chen, J., Myerson, J., Hale, S., 2002. Age-related dedifferentiation of visuospatial abilities. Neuropsychologia 40, 2050–2056.

Comon, P., 1994. Independent component analysis, a new concept. Signal Processing 36, 287–314.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D.Y., Engel, A.K., 2005. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. Journal of Neuroscience 25, 11730–11737.

Freeman, W.J., 2004. Origin, structure, and role of background eeg activity. Part 1. Analytic amplitude. Clinical Neurophysiology 115, 2077–2088.

Friedman, D., Kazmerski, V., Fabiani, M., 1997. An overview of age-related changes in the scalp distribution of *p*3b. Electroencephalography and Clinical Neurophysiology 104, 498–513.

Gibson, J.R., Beierlein, M., Connors, B.W., 1999. Two networks of electrically coupled inhibitory neurons in neocortex. Nature 402, 75–79.

Grady, C.L., Maisog, J.M., Horwitz, B., Ungerleider, L.G., Mentis, M.J., Salerno, J.A., Pietrini, P., Wagner, E., Haxby, J.V., 1994. Age-related changes in cortical blood flow activation during visual processing of faces and location. Journal of Neuroscience 14, 1450–1462.

Hupe, J.M., James, A.C., Girard, P., Lomber, S.G., Payne, B.R., Bullier, J., 2001. Feedback connections act on the early part of the responses in monkey visual cortex. Journal of Neurophysiology 85, 134–145.

Hyvrinen, A., Karhunen, J., Oja, E., 2001. Independent component analysis.

Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 1999. Analyzing and visualizing single-trial event-related potentials. Advances in Neural Information Processing Systems 11, 118–124.

Jung, T.P., Makeig, S., Humphries, C., Lee, T.W., McKeown, M.J., Iragui, V., Sejnowski, T.J., 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 subjects. Clinical Neurophysiology 111, 1745–1758.

Jung, T.-P., Makeig, S., McKeown, M.J., Bell, A.J., Lee, T.-W., Sejnowski, T.J., 2001. Imaging brain dynamics using independent component analysis. Proceedings of the IEEE 89, 1107–1122.

Klopp, J., Marinkovic, K., Chauvel, P., Nenov, V., Halgren, E., 2000. Early widespread cortical distribution of coherent fusiform face selective activity. Human Brain Mapping 11, 286–293.

Lauritzen, M., 1994. Pathophysiology of the migraine aura. The spreading depression theory. Brain 117, 199–210.

Lee, T.W., Girolami, M., Sejnowski, T.J., 1999. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. Neural Computation 11, 417–441.

Li, K.Z., Lindenberger, U., 2002. Relations between aging sensory/sensorimotor and cognitive functions. Neuroscience and Biobehavioral Reviews 26, 777–783.

Li, S.C., Lindenberger, U., Sikstrom, S., 2001. Aging cognition: from neuromodulation to representation. Trends in Cognitive Sciences 5, 479–486.

Madden, D.J., Gottlob, L.R., Allen, P.A., 1999. Adult age differences in visual search accuracy: attentional guidance and target detectability. Psychology and Aging 14, 683–694.

Makeig, S., Delorme, A., 2004. Physiological plausibility and stability of independent component analysis. Cognitive Neuroscience Society Abstracts.

Makeig, S., 1993. Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. Electroencephalography and Clinical Neurophysiology 86, 283–293.

Makeig, S., Bell, A.J., Jung, T.P., Sejnowski, T.J., 1996. Independent component analysis of electroencephalographic data. Advances in Neural Information Processing Systems 8, 145–151.

Makeig, S., Jung, T.P., Bell, A.J., Ghahremani, D., Sejnowski, T.J., 1997. Blind separation of auditory event-related brain responses into independent components. Proceedings of the National Academy of Science of the United States of America 94, 10979–10984.

Makeig, S., Westerfield, M., Jung, T.P., Covington, J., Townsend, J., Sejnowski, T.J., Courchesne, E., 1999. Functionally independent components of the late positive event-related potential during visual spatial attention. Journal of Neuroscience 19, 2665–2680.

Makeig, S., Westerfield, M., Jung, T.P., Enghoff, S., Townsend, J., Courchesne, E., Sejnowski, T.J., 2002. Dynamic brain sources of visual evoked responses. Science 295, 690–694.

Makeig, S., Delorme, A., Westerfield, M., Jung, T.P., Townsend, J., Courchesne, E., Sejnowski, T.J., 2004a. Electroencephalographic brain dynamics following manually responded visual targets. PLoS Biology 2, E176.

Makeig, S., Debener, S., Onton, J., Delorme, A., 2004b. Mining event-related brain dynamics. Trends in Cognitive Sciences 8, 204–210.

Massimini, M., Huber, R., Ferrarelli, F., Hill, S., Tononi, G., 2004. The sleep slow oscillation as a traveling wave. Journal of Neuroscience 24, 6862–6870.

Mattay, V.S., Fera, F., Tessitore, A., Hariri, A.R., Das, S., Callicott, J.H., Weinberger, D.R., 2002. Neurophysiological correlates of age-related changes in human motor function. Neurology 58, 630–635.

Molgedey, L., Schuster, H.G., 1994. Separation of a mixture of independent signals using time delayed correlations. Physical Review Letters 72, 3634–3637.

Onton, J., Makeig, S. (Eds.), 2006. Information-based Modelling of Event-related Brain Dynamics, vol. 159: Event-related Dynamics of Brain Oscillations. Elsevier, Salzburg, Austria.

Onton, J., Delorme, A., Makeig, S., 2005. Frontal midline EEG dynamics during working memory. Neuroimage 27, 341–356.

Pfefferbaum, A., Ford, J.M., Wenegrat, B.G., Roth, W.T., Kopell, B.S., 1984. Clinical application of the *p*3 component of event-related potentials. I. Normal aging. Electroencephalography and Clinical Neurophysiology 59, 85–103.

Polich, J., 1996. Meta-analysis of *p*300 normative aging studies. Psychophysiology 33, 334–353.

Polich, J., 1997. EEG and ERP assessment of normal aging. Electroencephalography and Clinical Neurophysiology 104, 244–256.

Ross, M.H., Yurgelun-Todd, D.A., Renshaw, P.F., Maas, L.C., Mendelson, J.H., Mello, N.K., Cohen, B.M., Levin, J.M., 1997. Age-related reduction in functional mri response to photic stimulation. Neurology 48, 173–176.

Townsend, J., Adamo, M., Haist, F., 2006. Changing channels: an fMRI study of aging and cross-modal attention shifts. Neuroimage 31, 1682–1692.