◆ Human Brain Mapping 29:711–725 (2008) ◆

Ranking and Averaging Independent Component Analysis by Reproducibility (RAICAR)

Zhi Yang,[1,2,3] Stephen LaConte,[2] Xuchu Weng,[1] and Xiaoping Hu[2]*

[1]*Laboratory for Higher Brain Function, The Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, People's Republic of China*
[2]*Biomedical Imaging Technology Center, The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, Georgia 30322*
[3]*College of the Humanities and Social Sciences, Graduate University of Chinese Academy of Sciences, Beijing 100101, People's Republic of China*

Abstract: Independent component analysis (ICA) is a data-driven approach that has exhibited great utility for functional magnetic resonance imaging (fMRI). Standard ICA implementations, however, do not provide the number and relative importance of the resulting components. In addition, ICA algorithms utilizing gradient-based optimization give decompositions that are dependent on initialization values, which can lead to dramatically different results. In this work, a new method, RAICAR (Ranking and Averaging Independent Component Analysis by Reproducibility), is introduced to address these issues for spatial ICA applied to fMRI. RAICAR utilizes repeated ICA realizations and relies on the reproducibility between them to rank and select components. Different realizations are aligned based on correlations, leading to aligned components. Each component is ranked and thresholded based on between-realization correlations. Furthermore, different realizations of each aligned component are selectively averaged to generate the final estimate of the given component. Reliability and accuracy of this method are demonstrated with both simulated and experimental fMRI data. *Hum Brain Mapp 29:711–725, 2008.*

Key words: fMRI; independent component analysis; data analysis

INTRODUCTION

Independent Component Analysis (ICA) is a data-driven approach that is widely used in functional magnetic resonance imaging (fMRI) [Calhoun et al., 2006; McKeown et al., 1998; Van de Ven et al., 2004]. The commonly used spatial ICA (sICA) considers the fMRI dataset as a linear mixture of spatially independent components that are "mixed" by their respective time courses. ICA has the advantage of being able to detect spatially distributed networks and temporal dynamics in the brain without assuming a known response [De Luca et al., 2006; McKeown et al., 1998]. It is thus suitable for exploratory analysis of fMRI data, where the general linear model (GLM) might be hampered by the lack of an appropriate *a priori* response model.

Contract grant sponsor: National Institutes of Health; Contract grant number: RO1EB00200; Contract grant sponsor: National Science Foundation of China; Contract grant numbers: 30425008, 30670674; Contract grant sponsor: Chinese Ministry of Science and Technology; Contract grant number: 2003CB515400.

*Correspondence to: Xiaoping Hu; Wallace H. Coulter Department of Biomedical Engineering, Emory University, Woodruff Memorial Research Building, 101 Woodruff Circle, Suite 2001, Atlanta, GA 30322. E-mail: xhu@bme.gatech.edu

Received for publication 11 October 2006; Revised 27 April 2007; Accepted 1 May 2007

DOI: 10.1002/hbm.20432
Published online 27 June 2007 in Wiley InterScience (www.interscience.wiley.com).

Despite its advantages, ICA has several theoretical and practical limitations, including its inability to determine the number of components and to order these components [Hyvarinen et al., 2001]. In standard ICA, the number of components is assumed to be equal to the rank of the voxel × time data matrix (this is usually the number of time points collected since it is the smaller of the two dimensions). This assumption is generally not appropriate in fMRI, where the number of signal sources is usually much less than the length of the fMRI time series [Beckmann and Smith et al., 2004]. Even though some ICA decomposition algorithms permit fewer components, the number of components generally needs to be specified before the decomposition, making the results user-dependent (e.g. [Kiviniemi et al., 2003; Van de Ven et al., 2004]). Another limitation is that there is no standard approach for ordering components, which may necessitate manual inspection of hundreds of components and make comparison of different ICA results problematic. Algorithmically, ICA decompositions utilizing gradient-based optimization are stochastic and are based on iteratively updating the "unmixing matrix" whose initial values are usually generated randomly [Himberg et al., 2004]. The randomness of the initialization introduces randomness into the ICA decomposition; consequently, a single decomposition is not reliable.

Several groups have attempted to address the above issues. McKeown and Sejnowski [1998] used an observation maximum likelihood method to estimate the number of components. Their method does not extract the actual number of components but can be used as a means to make comparisons across methods [Esposito et al., 2002; Formisano et al., 2004]. Beckmann and Smith [2004] have developed a probabilistic ICA method implemented in the MELODIC package that estimates the number of components using probabilistic PCA. Others [Esposito et al., 2001; Gu et al., 2001] introduced methods for ordering the ICA components—for example by spatial characteristics such as the number of voxels and connection properties. Moritz et al. [2003] proposed ranking the independent components by the relationship between their power spectrum and the stimulus frequency. Lu and Rajapakse [2003] developed a constrained ICA algorithm and ordered components by the kurtosis of their probability density distributions. LaConte et al. [2001] ordered components based upon their reproducibility across different epochs in event-related experiments.

Himberg et al. [2004] proposed an approach for "assessing both the algorithmic and statistical reliability of estimated independent components'', via clustering of repeated ICA realizations and visualization of the clusters. In the present work, a similar philosophy is utilized to develop an approach, RAICAR (Ranking and Averaging Independent Component Analysis by Reproducibility), for ranking the components, determining the number of reliable components, and improving their estimates through averaging. RAICAR is a framework that makes use of spatial reproducibility to evaluate ICA components; it can be used to deal with algorithmic variability arising from the optimization and/or data variability arising from measurement noise. Although not the focus of the present work, data variability can be examined by studying the reproducibility of ICA components with resampled data. In this paper, we focus on algorithmic variability by investigating the reproducibility of the component maps with repeated applications of ICA using different initialization values. To deal with algorithmic variability, RAICAR aligns the components of individual ICAs and ranks and selects the number of components based on their reproducibility. The underlying assumption is that spurious components exhibit greater fluctuation across realizations than stable ones. Therefore, the reproducibility of each component (*i.e.* the resilience against the influence of randomness) reflects its relative reliability, and also allows estimation of how many components are of sufficient stability to be retained. The final estimation of the retained spatial sources is obtained by selectively averaging them across the realizations, improving their quality and decreasing their stochastic nature. Application to simulated and experimental data demonstrates that this approach leads to interpretable and reliable results.

METHODS

RAICAR Algorithm

The basic idea of ICA is to decompose a data matrix into several independent sources that are linearly combined. Based on this idea, the spatial ICA employed in fMRI data analysis treats spatial pattern maps as the independent components, which are mixed together according to their corresponding time courses [McKeown et al., 1998]. This notion can be expressed as $\mathbf{X} = \mathbf{MS}$, where $\mathbf{X}$ is the observed dataset, $\mathbf{M}$ is the mixing matrix of time courses, and $\mathbf{S}$ represents the independent spatial maps. Suppose an observed dataset $\mathbf{X}$ (with T time points and V voxels) consists of C spatial sources. $\mathbf{X}$ is a $T \times V$ matrix; $\mathbf{M}$ is $T \times C$ while $\mathbf{S}$ is $C \times V$. The aim of ICA is to estimate both $\mathbf{M}$ and $\mathbf{S}$ simultaneously.

The independence assumption is utilized in order to perform the simultaneous estimation. That is, $\mathbf{M}$ and $\mathbf{S}$ are determined by maximizing the independence of the components. Nongaussianity can be used to quantitatively represent this independence, and several measures of nongaussianity have been proposed [Hyvarinen and Oja, 2000]. Algorithmically, the ICA model is often rewritten as $\mathbf{S} = \mathbf{WX}$, where $\mathbf{W}$ is a square full-rank unmixing matrix and $\mathbf{M} = \mathbf{W}^{-1}$ [McKeown et al., 1998]. In most ICA algorithms (e.g. FastICA [Hyvarinen, 1999]), $\mathbf{W}$ is iteratively updated until maximum nongaussianity is achieved. With the iterative procedure being a gradient-based optimization, $\mathbf{W}$ is usually initialized with random numbers at the beginning of the iteration, introducing randomness into the decomposition [Himberg et al., 2004]. The implication of having a random initialization is that different initial

conditions can lead to different results in the gradient-based optimization.

RAICAR performs the ICA decomposition K times (K realizations). In our implementation we use the FastICA algorithm with different initial conditions,[1] yielding K mixing matrices, $\mathbf{M}_1$, $\mathbf{M}_2 \cdots \mathbf{M}_K$, and correspondingly K sets of sources, $\mathbf{S}_1$, $\mathbf{S}_2 \cdots \mathbf{S}_K$. The number of components, C, is set to the estimated rank of the data matrix during the decomposition (the default for FastICA is the number of eigenvalues of $\mathbf{X}$ larger than 10^{-7}—for fMRI C often equals the smaller dimension of the data matrix, $\mathbf{X}$) and is the same for all realizations. Therefore, each realization leads to C components, hereafter referred to as a realization-component (RC) and indexed by a component number (ranging from 1 to C) and a realization number (ranging from 1 to K).

To examine the reproducibility of the RCs, the first step is to construct a cross-realization correlation matrix (CRCM). This matrix is $C \cdot K \times C \cdot K$ with the following structure:

$$\begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \mathbf{R}_{13} & \dots & \mathbf{R}_{1K} \\ \mathbf{R}_{21} & \ddots & & & \vdots \\ \mathbf{R}_{31} & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ \mathbf{R}_{K1} & \dots & \dots & \dots & \mathbf{R}_{KK} \end{bmatrix}$$

The $\mathbf{R}_{ij}$'s $(i = 1, 2, \dots, K, j = 1, 2, \dots, K)$ are $C \times C$ submatrices whose elements are the absolute value of the spatial correlation coefficients (SCC) for all pairs of components from realizations i and j. Thus, the SCC is defined as the absolute value of the Pearson's correlation coefficient between component maps. Note that the CRCM is symmetric, and submatrices $\mathbf{R}_{ii}$ are identity matrices and are ignored in the subsequent algorithm.

After construction of the CRCM, the next step is to align the realizations. Since there is no predetermined order for the ICA components, a given component's position may appear in any of the C positions from realization to realization. To align across realizations, the following procedure is repeatedly applied (see Fig. 1). First, the global maximum in the CRCM is identified. Let us assume that the maximum is located at the mth row and the nth column (denoted as $[m, n]$ hereafter) in submatrix $\mathbf{R}_{ab}$ (there is an identical value in $\mathbf{R}_{ba}$ since the CRCM is symmetric). This maximum allows us to establish correspondence between component m in realization a and component n in realization b, defining the starting point of an aligned component, with realization components RC_{ma} and RC_{nb}. In the remaining realizations, the RCs having the maximum correlation with these RCs are identified. Specifically, the mth row of the submatrices $\mathbf{R}_{ai}$ (i refers to all the realizations other than a or b) and the nth column of the submatrices $\mathbf{R}_{ib}$ are searched, for respective maxima. The positions of the maxima are denoted $[m, p_i]$ in $\mathbf{R}_{ai}$ and $[q_i, n]$ in $\mathbf{R}_{ib}$, as shown in Figure 1. In many cases, the two maxima for each realization correspond to the same RC (i.e. $p_i = q_i$ for realization i); in this case, component p_i of realization i is assigned to the aligned component defined by $[m, n]$ in $\mathbf{R}_{ab}$. In the case that $p_i \neq q_i$, the SCCs of components p_i *and* q_i of realization i are compared, and the one with the larger SCC is assigned to the aligned component defined above. The rows and columns that contain the entries and their diagonal reflections are subsequently eliminated from the CRCM before the next repetition starts. The procedure is repeated C times until C aligned components are identified. Each aligned component entails K RCs, one for each realization, and cross-correlation between these aligned RCs produces $K(K-1)/2$ SCCs.

After the alignment, a histogram of the SCCs in the CRCM (upper triangle) is generated.[2] This histogram is bimodal, with modes near 0 and 1, representing a large number of components that are not correlated and a small number of components that are highly correlated, respec-

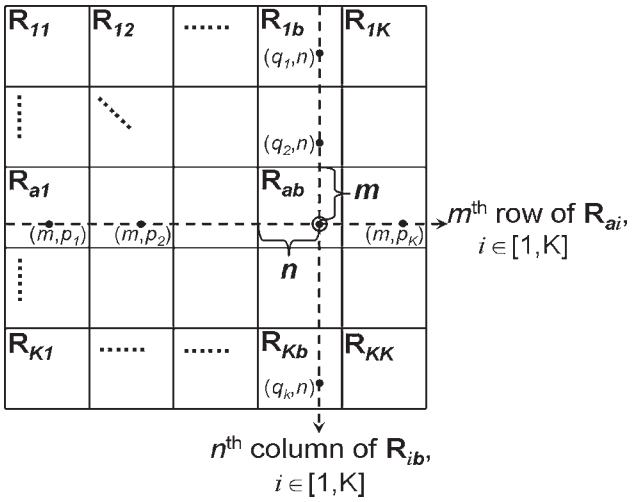

Figure 1.

Each submatrix in the cross-realization correlation matrix (CRCM), $\mathbf{R}_{ij}$, is the spatial cross-correlation matrix between realizations i and j. The circled dot represents the global maximum in the CRCM. After finding this maximum, the mth row in each submatrix $\mathbf{R}_{ai}$ and the nth column in each submatrix $\mathbf{R}_{ib}$ are searched for a corresponding submatrix-specific maximum. These maxima are indicated by dots with their positions given in parentheses.

[1]For the results reported here, the random initial conditions were generated by using the computer's system time to generate a random seed. From this seed, a random matrix was generated, which determined the initial conditions for each realization.

[2]Our implementation used 100 bins with SCC values ranging from 0 to 1 using default histogram function in Matlab™ (Mathworks, Natick, USA).

tively. To eliminate insignificant SCCs, an SCC threshold is set at a point between the two modes. In this report, the SCC threshold is found by smoothing the histogram and searching for the minimum. A reproducibility index for each aligned component is then generated by summing the SCCs, among the aligned component's RCs, that are above the SCC threshold. Importantly, as our results demonstrate, the ranking result is not sensitive to the exact choice of SCC threshold, provided it lies in the valley of the histogram between the two modes.

The reproducibility index is used to rank the aligned components in descending order. If an aligned component is consistent across ICA realizations, its reproducibility index will be high and thus its ranking order will be high. The reproducibility index is a measure of component reliability and can be used to determine the number of reproducible components. In fact, given that the SCCs exhibit a bimodal distribution, most components will have a very low reproducibility index, and a few components will have a relatively high reproducibility index. As we show in our results, the ordered reproducibility index generally drops off sharply, allowing us to estimate the number of components, by keeping those components whose reproducibility index is above a cut-off point. Visual inspection of the ordered reproducibility plot allows the selection of the cut-off point "by eye" and is recommended in practice. The choice of a cut-off is often required in many data reduction techniques (e.g. in dimensionality reduction using a plot of singular values) and is the responsibility of the experimenter. In this work, to be consistent across experiments, the cut-off point is set to 50% of the index's maximum possible value ($\frac{K(K-1)}{2} \times 0.5$). Another possibility includes finding the maximum slope in the reproducibility plot. And, even without selecting a cut-off, the number of components is generally greatly reduced compared to C, making it possible to examine all of these components.

To generate the final components, the spatial maps and the corresponding mixing time courses of the RCs of each aligned component are selectively averaged. That is, only RCs that have at least one SCC higher than the threshold for all of the other the RCs of the given aligned component are included when generating the averaged components.

Simulation 1

A simulation study was conducted to evaluate the performance of RAICAR. Six nonoverlapping spatial sources with equal area were generated and mixed together according to the mixing time courses indicated in Figure 2. The mixing time courses had zero mean and consisted of 162 time points. A slowly varying global baseline was added to all the pixels, and the resultant time course for each pixel was further degraded by adding Gaussian white noise. The SNR for each component was determined by the variance of the mixing time courses, relative to that of the Gaussian white noise. Specifically, the SNR of the six mixing time courses were 0.35, 0.29, 0.24, 0.20, 0.16, and 0.14, respectively. The SNR of the slowly varying global baseline was set to 0.11. With these values, the contrast-to-noise ratio (CNR $\equiv \Delta\mathbf{S}/\sigma_{\text{noise}}$) of the mixing time courses ranged from 0.92 to 3.87, consistent with CNR values reported in the fMRI literature [Esposito et al., 2002; Huettel et al., 2004].

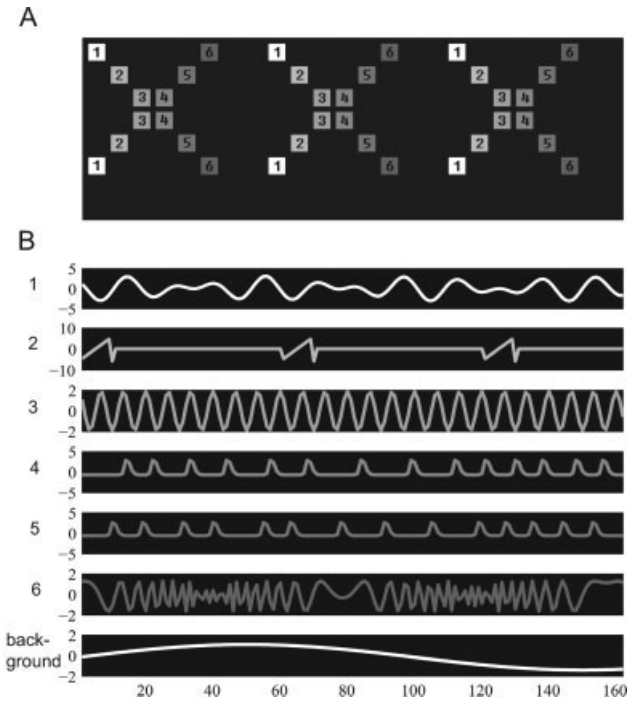

Figure 2.
Simulated spatial sources and their mixing time courses. **A:** Spatial map of the numbered sources. All the sources are equal in area. **B:** The corresponding mixing time courses of the sources. The bottom panel shows the added global baseline.

No pre-processing was applied to the simulated data. FastICA (http://www.cis.hut.fi/projects/) for Matlab™ (Mathworks, Natick, USA) was used to carry out the ICA decompositions. Thirty different ICA decompositions were obtained with random initial conditions ($K = 30$) to perform RAICAR. To examine the relevance of the resultant components, they were matched to the true sources according to their temporal correlations, and a receiver operating characteristics (ROC) analysis was conducted based on the spatial maps. To test reliability, we repeated the above procedure 10 times.

Simulation 2

In Simulation 2, we superimposed the sources from Simulation 1 onto resting state fMRI time series. The resting state data were acquired using a 3T Siemens Trio scanner (Siemens Medical Solutions, Malvern, PA), with TR = 2 s, TE = 34 ms, and flip angle = 90°. The resting state data

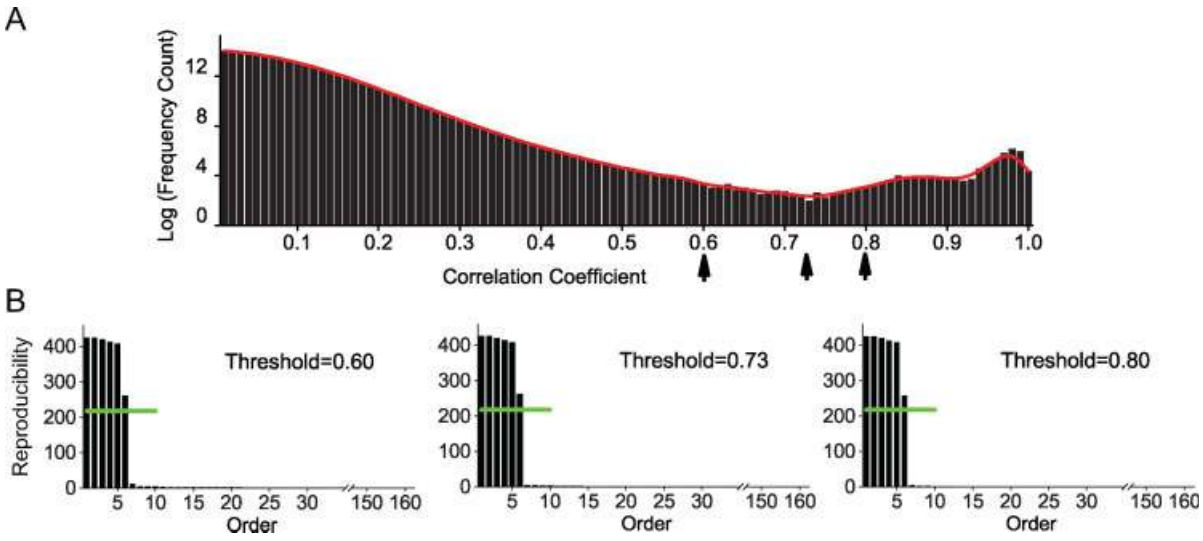

Figure 3.

Reproducibility ranking of the simulated data. **A:** Bimodal distribution of the correlation coefficients. The majority of the correlation coefficients lie in the lower range of 0–0.60, while the remaining fall in the upper range of 0.80–1.00. These two ranges are separated by a broad valley (roughly 0.60–0.80). The solid red curve is the smoothed histogram used to determine the threshold. The arrows indicate the three SCC thresholds used. **B:** The reproducibility index plots generated using the three SCC thresholds. The half-maximum cut-offs are shown with horizontal green lines, indicating six components in each case. The orders of the components derived with all three thresholds are also the same.

were corrected for motion and for physiological noise using the AFNI RETROICOR plugin [Cox, 1996; Glover et al., 2000]. We applied RAICAR to three combinations of data: resting state data only, resting state data with low-contrast simulated sources, and resting state data with high-contrast simulated sources. Contrast was controlled by scaling the magnitude of the simulated sources on a pixel-by-pixel basis. For the low contrast case, the simulated mixing time courses had the same CNRs as in Simulation 1, and for the high contrast case, the CNRs of the simulated mixing time courses were 2.7 times those in Simulation 1.

Delayed Motor Task

The data were acquired on three healthy right-handed participants using single shot T2*-weighted EPI on a GE Signa 1.5T scanner (GE Medical Systems, Milwaukee, WI)

TABLE I. The order of the components corresponds to the descending order of their SNR

Order	Reproducibility index	SNR
1	425.6	0.35
2	425.3	0.29
3	419.8	0.24
4	413.3	0.20
5	407.7	0.16
6	261.8	0.14

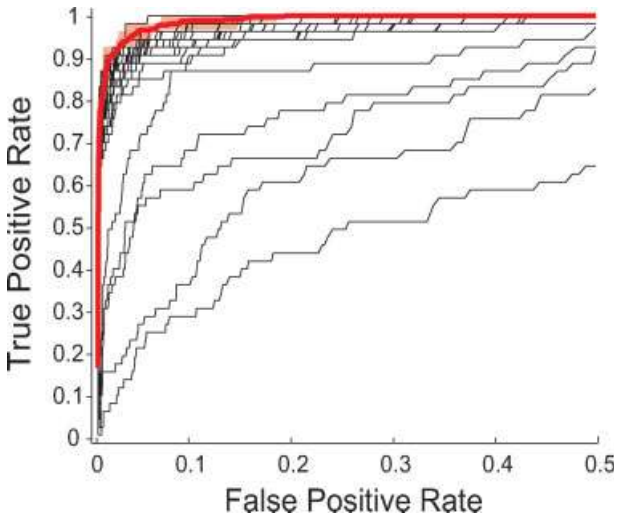

Figure 4.

Detectability of RAICAR and individual ICA realizations using the sixth simulated source (with lowest SNR). ROC curves of the individual ICA realizations (black) and RAICAR results (red) are shown. The light red region shows the spread of 10 RAICAR repetitions and the red curve shows their mean. The individual ICA realizations exhibit variable results, while the repeated RAICAR results are all virtually identical and outperform the majority of individual ICA realizations.

with the following imaging parameters: TR/TE = 2,000/50 ms, flip angle = 90°, FOV = 220 mm, matrix = 64 × 64, five oblique axial slices, starting from the top of the head, slice thickness/gap = 5/0 mm, and 126 volumes. The task consisted of right-handed finger tapping in a delayed movement paradigm. The participants were presented with a visual cue indicating the finger tapping sequence, but did not move their fingers until a "Go" command was given [Catalan et al., 1998; Ogawa et al., 1998]. For each trial, the visual "Cue" and "Go" stimuli lasted for 2 s each and the interval between them was 12 s. The inter-trial interval (ISI) was also 12 s. Each run consisted of nine such trials, each lasting 28 s.

These data were motion corrected, baseline detrended and masked to exclude the voxels outside the brain using AFNI [Cox, 1996]. RAICAR was applied, with 30 repetitions of randomly initialized ICA decompositions ($K = 30$). Identification of the task-related components was achieved by correlating the mixing time course of the resulting components to the stimulus sequence, convolved with an ideal hemodynamic response function (HRF) generated by AFNI. To examine the robustness of the reproducibility method, it was applied three times, each time using three different SCC thresholds.

Constant Force Grip Task

Subjects repeatedly gripped a water-filled bottle with their right hand. The force of the grip was gauged by the water pressure in real time and presented to the subject so that he/she can adjust the force to meet a target level of 50% of his/her maximal voluntary contraction (MVC) level [Liu et al., 2000, 2002], calculated based on the maximal grip force measured at the beginning of the experiment. Subjects performed the gripping by following visual cues (generated by a waveform generator [Wavetek Datron, San Diego, CA]) projected onto the screen above the subjects' eyes in the magnet. Each visual cue was a rectangular pulse that indicated (target amplitude for 50% MVC and desired duration of 3.5 s) the desired contraction. The duration of each contraction was 3.5 s, followed by a 6.5-s resting interval [Peltier et al., 2005]. fMRI data were collected on a 3T Siemens Trio scanner (Siemens Medical Solutions, Malvern, PA), with 30 axial EPI slices (TR/TE = 2,000/30 ms, voxel = 3.4 × 3.4 × 4 mm^3, flip angle = 90°). Motion correction and brain masking were applied as preprocessing procedures, and RAICAR was preformed ($K = 30$). To examine the sensitivity to the SCC threshold for these data, we repeated RAICAR with two additional SCC threshold values.

Investigation of the Impact of *K*

The stability of the rank positions was studied as a function of K for Simulation 1. We also investigated the variability of the RAICAR results as a function of K (the number of individual ICA realizations used by RAICAR). For Simulation 1 and both experimental datasets, RAICAR was applied 20 times at different K values ranging from 5 to 55 in steps of 5. All component maps were matched to the results reported for the above studies (which used $K = 30$). For each K value, the variance for each component in the 20 RAICAR repetitions was calculated and averaged

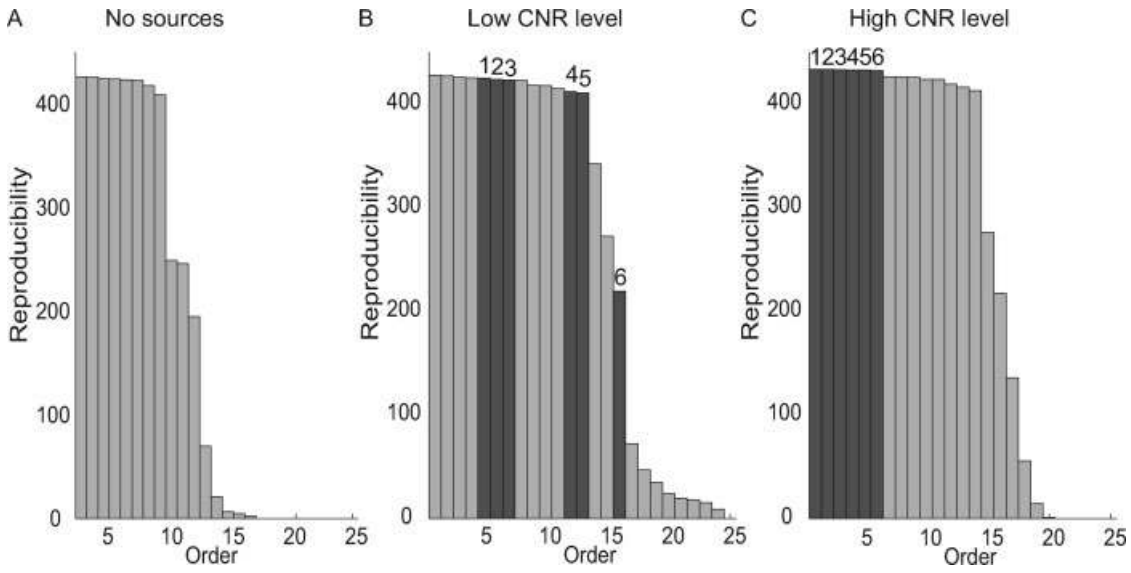

Figure 5.
Reproducibility rankings obtained from simulation 2. **A:** resting-state data only. **B:** resting-state data with low CNR sources. **C:** resting-state data with high CNR sources. The increased CNR level shifts the components towards the left (increasing their rank).

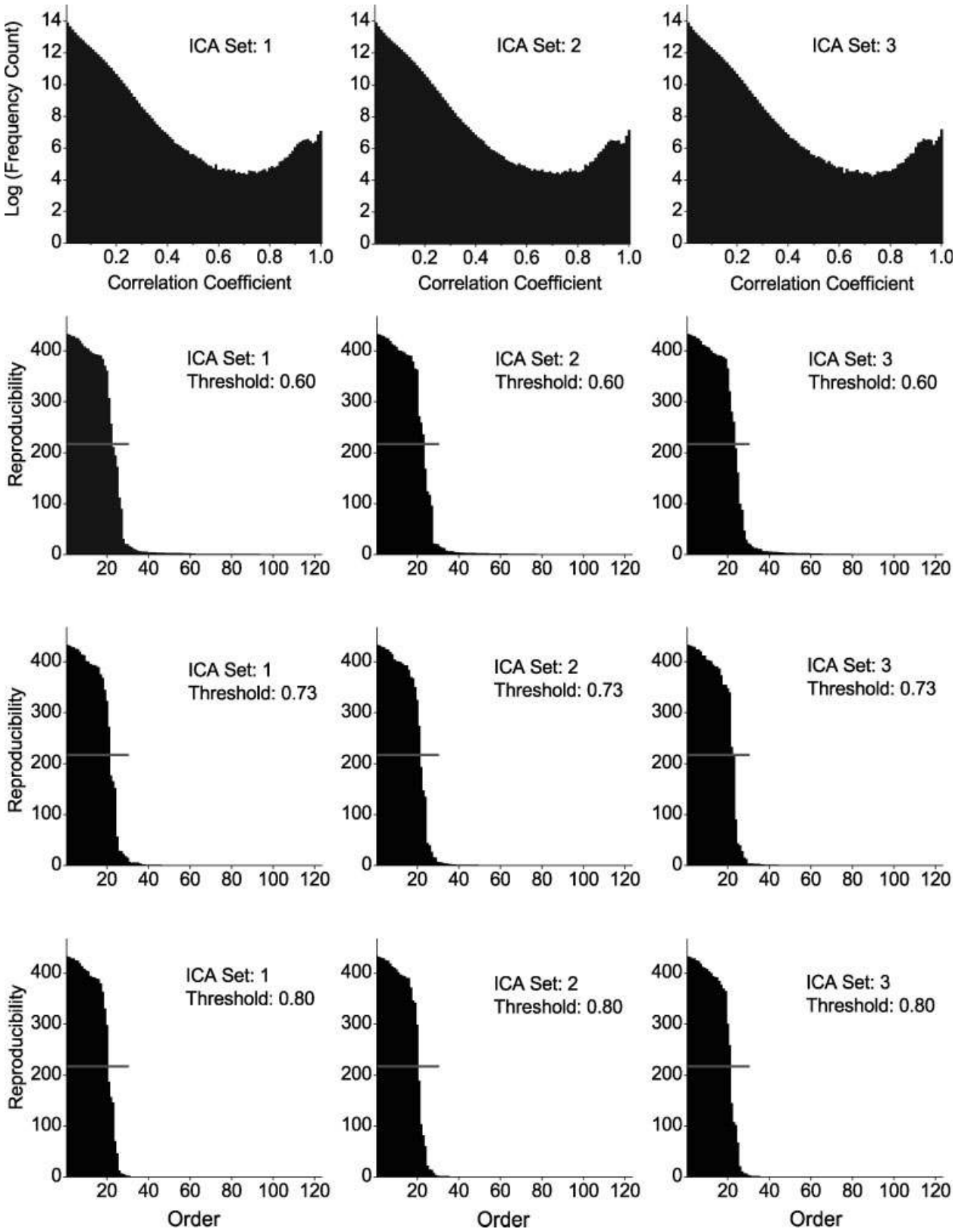

Figure 6.

Reproducibility ranking of the event-related, delayed motor data in one subject. Columns correspond to different sets of ICA realizations. The top row shows the histograms of the SCCs for the three sets of ICA realizations. The correlation coefficients are distributed in two modes, one near zero and the other near one. The bottom three rows are ordered reproducibility index plots for three different thresholds. It can be seen that the number of components passing the cut-off (shown as horizontal gray lines) do not vary significantly with the SCC thresholds and different sets of ICA realizations.

over all voxels and all components. The average variance as a function of K was used to measure convergence.

RESULTS AND DISCUSSION

Simulation 1

The simulation data were decomposed into 162 components by FastICA ($C = 162$). Figure 3 shows the histogram of the SCCs and the plot of the ordered reproducibility index. The SCC histogram (Fig. 3A) follows a bimodal distribution. The smoothed histogram is shown as the red curve in Figure 3A, and has a minimum at 0.73, which was used to threshold the SCCs. In addition to 0.73, two other SCC thresholds, 0.60 and 0.80 (marked with arrows in Fig. 3A), were also used to test the sensitivity of the analysis to the SCC threshold. Out of 162, the number of reliable components determined using the half-maximum cut-off (217.5) was six for all SCC thresholds. Each of these six components uniquely matches one of the original sources in terms of both spatial pattern and mixing time course. As shown in Table I, the order in which they appear in Figure 3 coincides with the order of decreasing SNR of the sources. Therefore, the present method successfully extracted both the number and the order of the components in the simulated data. While the rank-order may not coincide with the SNR order in general, it does reflect the "strength" of the component.

The ordered reproducibility indices generated from the three SCC thresholds are plotted in Figure 3B. In these plots, the first six components (those above the cut-off) exhibit the same ranking and have reproducibility indices that are essentially independent of the SCC threshold. In other words, the SCC threshold has a negligible effect on the ranking of the reproducible components and their number.

The ROC curves generated for the sixth source in the simulation are shown in Figure 4. RAICAR outperforms nearly all the individual ICA realizations. The wide spread in the detectability (area under the curve) among individual ICAs clearly indicates the stochastic nature of FastICA, which can not be neglected. In contrast, the spread of the 10 repetitions of RAICAR is relatively small. The temporal correlation coefficients between the original source and the mixing time course estimated by different methods are 0.89 for the best individual ICA, 0.31 for the worst individual ICA, and 0.91 for RAICAR. Identical results were obtained with two other repetitions of the ranking process.

It is important to note that the sources in our simulation studies are not exactly independent since they are mutually exclusive (e.g. Source 1 does not spatially overlap with any other source). This lack of complete independence may lead to difficulties in ICA decomposition but is probably more reflective of real fMRI data where independence is not guaranteed. Fortunately, our results show that some dependence in the data did not prevent us from obtaining good results.

Simulation 2

Ten reproducible components were obtained in the resting state data (see Fig. 5A), reflecting functional connectivity [Biswal et al., 1995] and structured noise. When the simulated sources were added to the resting state data, they were recovered by RAICAR and ranked in the order of descending SNR. As shown in Fig. 5B,C, the increase of contrast of the simulated sources shifted their position in the rank. At low contrast, the simulated sources were located away from the top ranks and dispersed among the resting-state components, while at high contrast, they became the top components. This may indicate that at the high CNR levels, the simulated components have CNR higher than or comparable to those of the resting components. These observations indicate that RAICAR works well in the presence of fMRI background noise and the ranking provides a useful measure of the strength of the components.

Delayed Motor Task

For the delayed motor data, the number of components in each individual ICA realization was 123 for all three subjects ($C = 123$). The thresholds determined by the SCC histograms were 0.73, 0.75, and 0.71, respectively, for the three subjects. Figure 6 shows RAICAR result of one of the subjects. Similar to the simulated data, the SCCs (top row) are distributed in two modes, one near zero and the other close to one. The bottom three rows show the reproducibility index plots, obtained using three different SCC thresholds, with three different sets of ICA realizations. As Table II shows, the number of components determined by the cut-off index does not vary substantially with either the SCC threshold or the set of ICA realizations. However, there is a slight variation in the number of components determined since the drop-off of the reproducibility index is more gradual in these data than in the simulated data. Table II also shows that "Go" and "Cue" components of the delayed movement task have ranks that are not highly dependent on the SCC threshold or the set of ICA realizations. Similar conclusions hold for the results for the other two subjects. The insensitivity to the SCC threshold indicates that it is a parameter that does not have to be carefully selected, and the insensitivity to the specific set of

TABLE II. The estimated number of components (Num. IC) and the positions of the task-related components (go, cue)

	ICA set 1			ICA set 2			ICA set 3		
Thr.	Go	Cue	Num. IC	Go	Cue	Num. IC	Go	Cue	Num. IC
0.60	19	14	21	20	14	23	19	15	23
0.73	19	13	21	19	14	21	19	15	22
0.80	19	14	21	18	13	20	19	15	22

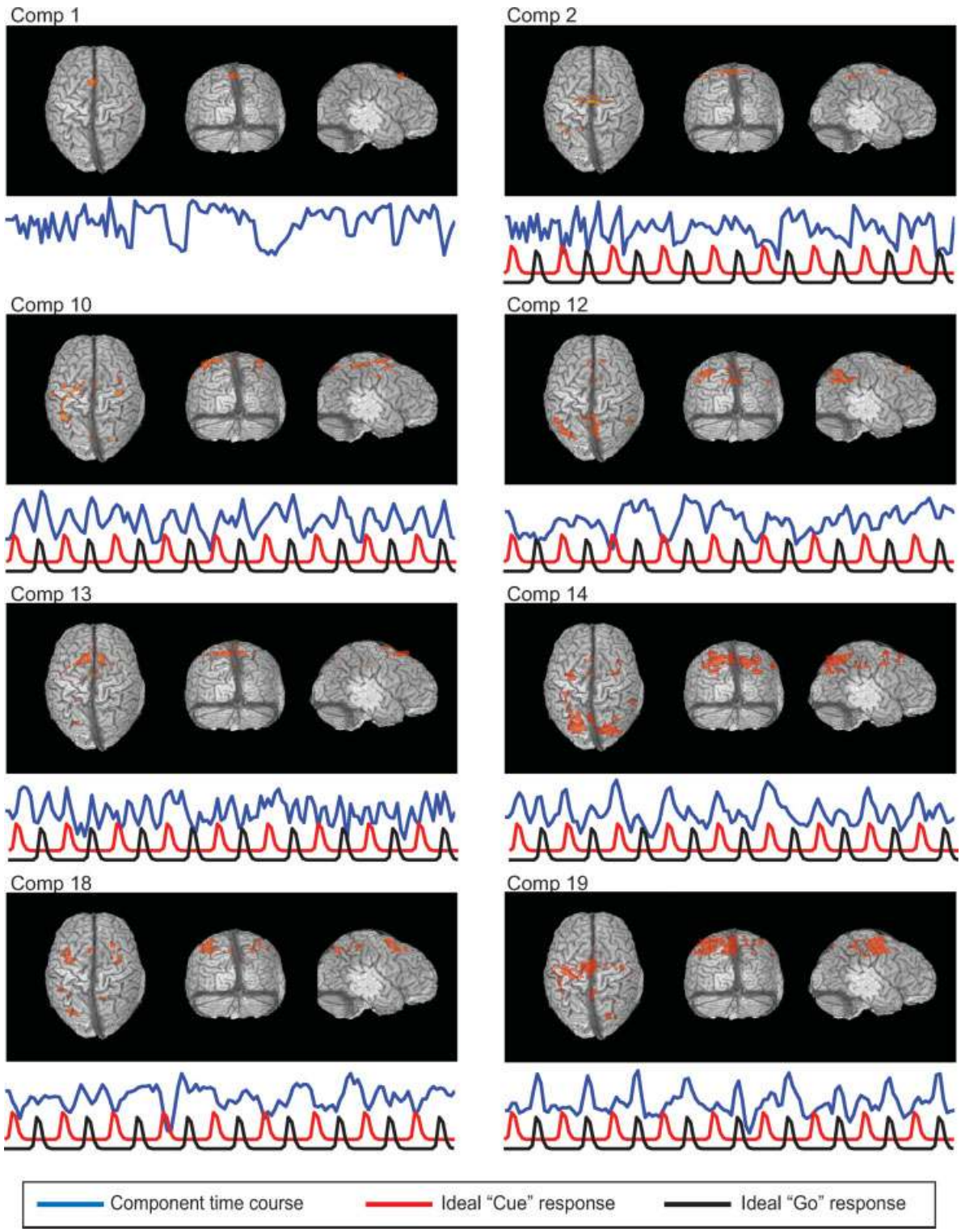

Figure 7.

Several components extracted by RAICAR in the delayed motor dataset. Component maps are displayed on the transparent glass brain with corresponding time courses shown in blue; the red and black curves illustrate the ideal response of "Cue" and "Go" task, respectively. Components 1 and 2 are due to cardiac noise; Component 10 is task-related activation in the sensory and eye field regions; Component 12 shows the activations corresponding to the visual motion control; Component 13 is related to head motion; Component 14 corresponds to the "Cue" task; Component 18 shows the activations in prefrontal cortex; Component 19 is activated by the "Go" task.

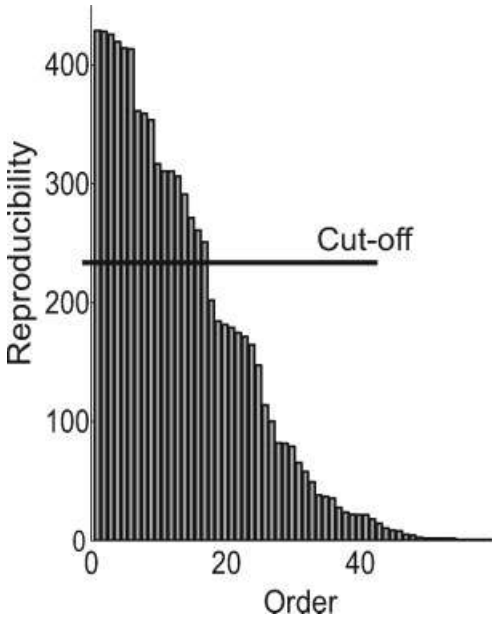

Figure 8.
The reproducibility rank obtained from the constant force grip dataset. 17 components were above the reproducibility cut-off.

ICA realizations minimizes the stochastic nature of an individual ICA.

Figure 7 shows eight components extracted by RAICAR. The two top-ranked components are related to cardiac noise; Component 10 shows the task-related activations on the sensory and the eye field regions; Component 12 shows the inferior parietal regions (BA40/7) and the precuneus, which are related to visual motion control; component 13 clearly reflects artifacts due to head motion; component 14 contains activations in posterior parietal and prefrontal regions, which are related to motor preparation [Hanakawa et al., 2003]; component 18 primarily shows activation in bilateral prefrontal cortex, which is related to motor preparation and spatial imagery; component 19 is related to the motor execution task.

To compare the RAICAR "Go" map (component 19) and the individual ICA maps, we calculated their spatial correlation coefficients with the GLM activation map corresponding to the "Go" task. As expected based on our simulated ROC studies, the correlation of the RAICAR pattern of activation was in the upper range of the individual ICA realizations, even though in this case the range was relatively narrow. Specifically, the 30 individual ICA realizations had a mean correlation of 0.489 and standard deviation of 0.017 while the RAICAR correlation was 0.516.

Constant Force Grip Task

For the constant force grip data, the number of components in each individual ICA realization was 71 ($C = 71$). Using a threshold of 0.80, as determined by the SCC histogram, the reproducibility rank was generated and the number of components was determined as 17. Figure 8 shows the reproducibility rank, and Figure 9 presents 3D views of all components extracted by RAICAR. The first component reflects regions that have been implicated in the default mode network [Raichle et al., 2001]. The second component is task-related activation in the primary motor area. Components 4 and 5 are also task-related, likely corresponding to motor control; interestingly, their mixing time courses follow the behavioral paradigm but are modulated by a slow variation. Component 11 reveals some task-related artifacts at the base of the brain. Other components are either artifacts or unexplained brain activation.

To make a visual comparison between RAICAR and individual ICA maps, Figure 10 shows components from both RAICAR and a randomly selected individual ICA. For the top components in the reproducibility rank, RAICAR maps and individual ICA maps do not exhibit great differences. However, as the order approaches the end of the reproducibility rank, individual ICA maps tend to be noisier than the RAICAR maps.

Table III lists the reproducibility ranks with different SCC thresholds. Although the total number of components varies with different SCC thresholds, the number of reproducible components (above the reproducibility cut-off) dose not change substantially, again showing the insensitivity to the SCC threshold.

Row and Column Equivalence of the CRCM

Our implementation handles the case that $p_i \neq q_i$, by choosing the RC with the larger SCC in the alignment process. To evaluate how frequent this case arises in practice, we tracked its occurrence during the alignment for the simulation 1 and the two experimental data sets. Table IV shows the occurrence rate of $p_i = q_i$ averaged over the reproducible components and 20 repeated applications of RAICAR. For all three datasets, more than 99% of the cases are $p_i = q_i$, and the standard deviations across 20 repetitions are rather small. This means that the $p_i \neq q_i$ cases are rare in the reproducible components.

Impact of *K*

Figure 11 shows the convergence characteristics of RAICAR. Figure 11A gives the positions of the components identified by RAICAR from data in Simulation 1 as a function of K. For K above 20, there is no change in the positions. Figure 11B displays the area under the ROC curve as a function of K, generated from the lowest SNR source in Simulation 1. The area under the ROC curve is constant when K is larger than 20. Panels C, D, and E of Figure 11 show the average variance among 20 RAICAR repetitions for Simulation 1, the delayed motor dataset, and the con-

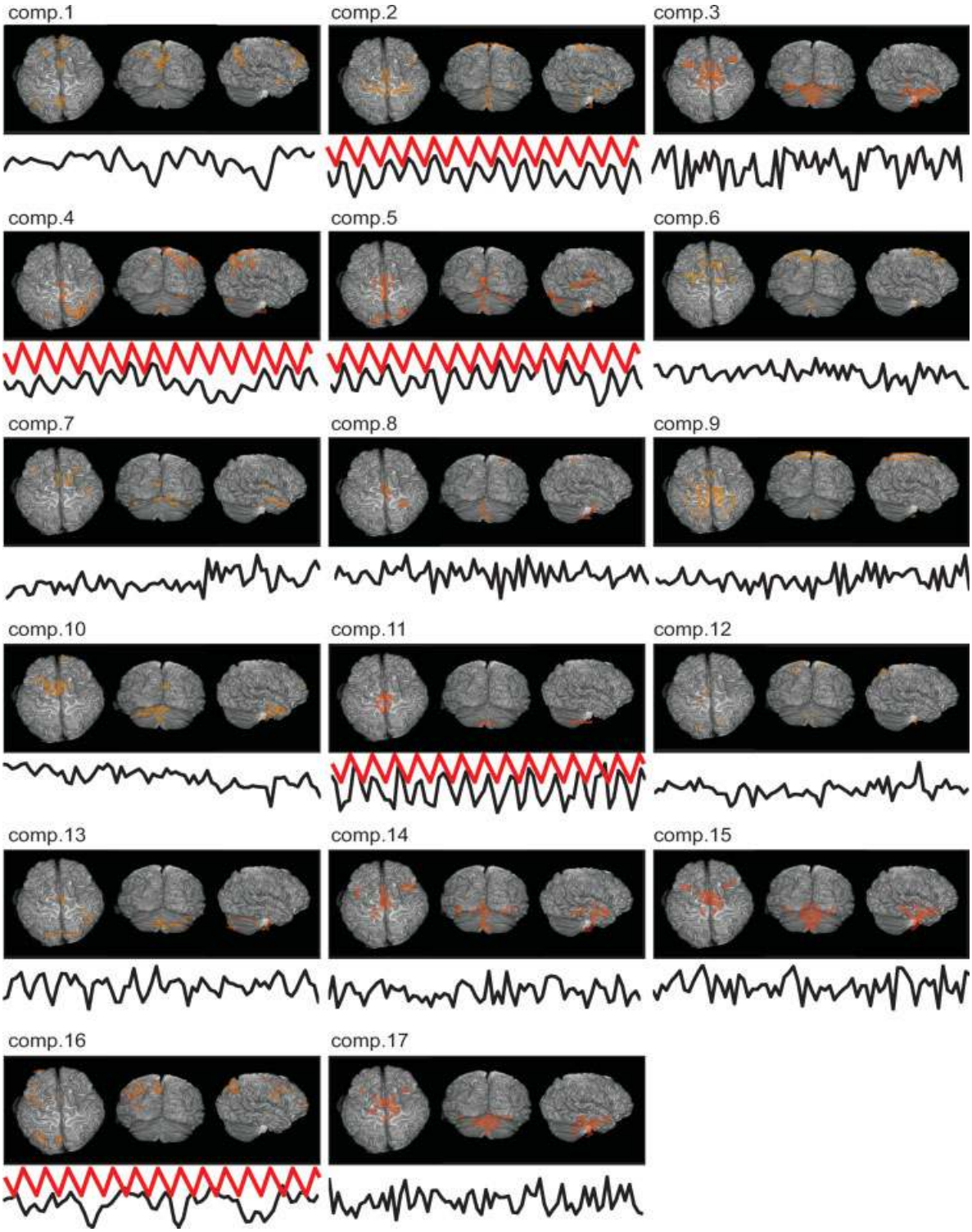

Figure 9.

The 17 components for the constant force grip dataset. The component maps are shown on the transparent glass brain with the corresponding time courses shown below. The ideal task response is shown (red) for components that are highly correlated with it. The first component may reflect the default mode network; Components 2, 4, and 5 arise from task-related activations, which include functional areas for motor control and execution; Component 11 seems to be due to task-related artifacts at the base of the brain. Other components are either artifacts or unexplained brain activation.

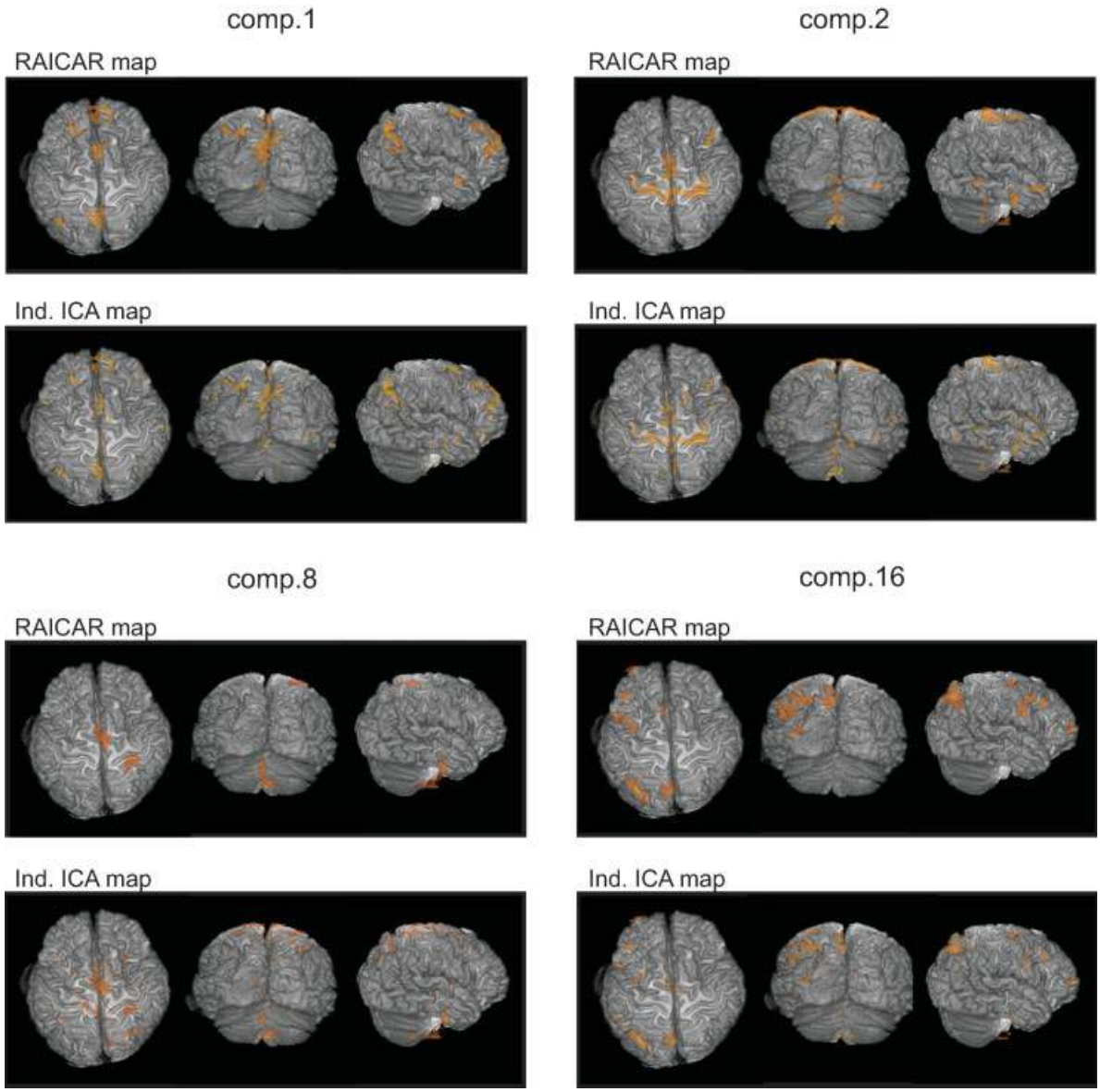

Figure 10.
Comparison of the component maps extracted by RAICAR and a randomly selected individual ICA realization. The top row of each component shows the RAICAR map and the bottom row shows the individual ICA map. For higher ranked components, both results tend to be similar. While for lower ranked components the individual ICA maps tend to be noisier than the corresponding RAICAR maps.

stant force grip dataset as functions of K. When $K = 30$, the average variance is no more than 0.005 in all three datasets. This variance is negligible when compared with the intensity of the component maps (mean = 0, standard deviation = 1), indicating that a K of 30 is sufficient for RAICAR applied to the three data sets examined. In practice, the proper choice of K can be data dependent. The methods described here allow users to choose proper K value for different datasets.

TABLE III. The number of reproducible components (using the half-max cutoff) is not sensitive to different SCC thresholds

SCC threshold	Number total components	Number reproducible components
0.70	71	18
0.75	65	18
0.80	58	17

TABLE IV. Occurrence rate of the $p = q$ cases

Data set	Percentage of $p = q$ (mean ± std)
Simulation1 (6 comps)	(99.88 ± 0.24)%
Delayed motor (21 comps)	(99.93 ± 0.20)%
Constant force grip (17 comps)	(99.26 ± 0.43)%

The means and standard deviations are generated based on 20 RAICAR repetitions.

General Discussion

RAICAR is an extension of AFRICA [LaConte et al., 2001] and is unique compared to other methods for ordering independent components [Esposito et al., 2001; Gu et al., 2001; Lu and Rajapakse, 2003; Moritz et al., 2003], in that it relies upon reproducibility. Selective averaging of RCs is used by RAICAR to generate the final components. Other consensus methods, such as clustering or weighted averaging, could be used to estimated the components from repeated ICAs; similarly, it is possible to use measures other than correlation coefficient (e.g. Fisher transform of the correlation coefficients or R^2) to generate the reproducibility index. Finally, other stochastic algorithms beyond FastICA can also be incorporated. These possibilities remain to be explored.

As mentioned before, RAICAR is demonstrated here only with algorithmic variability. It is also possible to consider data variability. Procedurally, the primary difference would be to repeat ICAs with different data subsets instead of or in addition to using different initialization values.

Our results suggest that the number and order of RAICAR components is important. In particular, the results of Simulation 2 indicate that the number of components is consistent with the ICA model of mixed sources and their

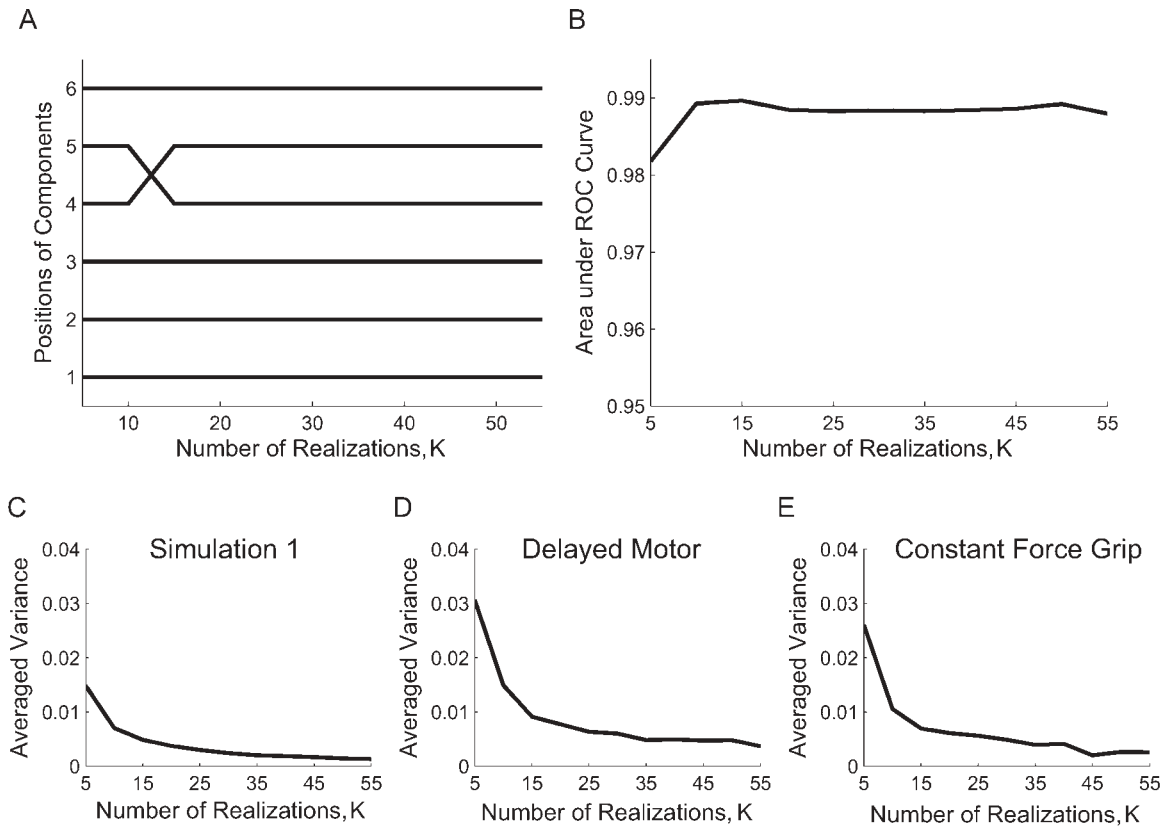

Figure 11.

Convergence of RAICAR. **A:** The positions of the components as a function of *K*, generated from simulation 1. The positions of the components do not change when *K* is larger than 20. **B:** The area under the ROC curves as a function of *K*, generated from the lowest SNR source in Simulation 1. The estimation error of the component maps does not change substantially when *K* is larger than 20. **C**, **D**, and **E:** The average variance of the component maps as a function of the number of ICA realizations (*K*), generated from the six sources in simulation 1, the 21 sources in the delayed motor dataset, and the 17 sources in the constant force grip dataset. For all three datasets, these curves rapidly approach asymptotic levels. When $K = 30$ (as used for our reported results), the average variance is no more than 0.005.

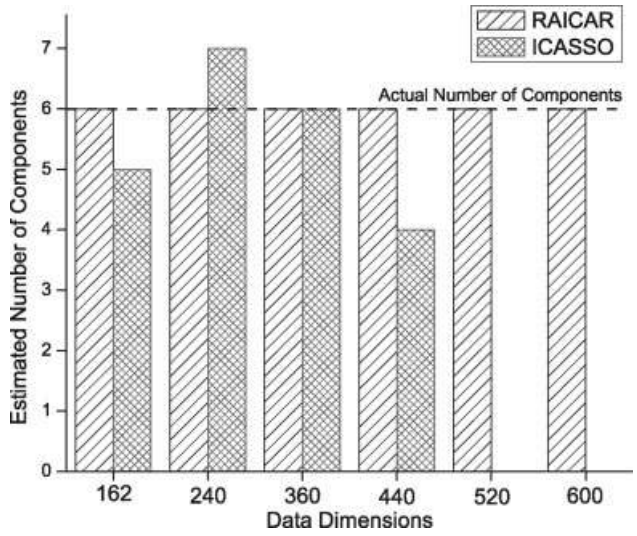

Figure 12.
Comparison of the estimated number of components for RAICAR and ICASSO. The comparison was conducted using the sources from Simulation 1 with different time series lengths. RAICAR estimates the correct number of sources at all lengths, while ICASSO results vary with data length. When the length exceeds 440, ICASSO estimates could not be obtained due to large memory requirements.

positions can be manipulated by changing their CNR. Similar observations can be made from our experimental results. In the delayed motor task, the task-related components were ranked between 10 and 19. This is likely due to the low CNR event-related signals. In contrast, in the constant force grip task, the task-related components resided mainly at the top part of the reproducibility rank, consistent with the high CNR in the continuous hand grip data. These observations further support the notation that the reproducibility rank is meaningful and the reproducibility can be used as a measure of "strength."

The ICASSO method [Himberg et al., 2004] was developed to validate and visualize ICA results by clustering the independent components. Their findings support the premise taken here that randomness exists in gradient-based ICA algorithms and can be reduced to produce better results. Our approach bears similarity to ICASSO in repeating the ICA decomposition and extracting components from the multiple ICA realizations. Importantly, though, there are substantial differences between RAICAR and ICASSO. (1) In RAICAR, the matrix-based alignment is much less demanding in terms of computational requirements than the hierarchical clustering used in ICASSO. In fact, ICASSO requires far more memory and CPU operations, and its practical use often requires data reduction with PCA as a preprocessing step. As pointed out by McKeown et al. [1998], some ICA components of interest may comprise only a few percent of the total variance and be inadvertently eliminated by PCA. Note that the significant reduction in computational burden makes RAICAR practical for large datasets that may arise from long time series or inter-session/subject data. (2), ICASSO identifies the resultant component from each cluster using the centrotype while RAICAR uses selective averaging of the aligned components. This difference is expected to lead to different spatial and temporal results. (3) The metrics for evaluating the quality of components for both methods are different. ICASSO provides two metrics—the quality index and *R*-index, whose interpretation requires expert knowledge. RAICAR uses a spatial reproducibility index which is intuitive to understand and use.

To provide a more direct comparison of the two methods, they were both applied to our simulated sources. In Figure 12, the estimated numbers of components from the two methods are shown for data generated with different time lengths. The RAICAR results give a correct estimation for all data lengths while ICASSO estimation is less accurate, varying across the data lengths. Moreover, as shown in Table V, ICASSO required much longer processing time, and, for data with more than 440 time points, ICASSO could not be successfully run (due to excessive memory requirements) on a workstation with 6 GB RAM.

From a practical point of view, RAICAR does not require difficult user decisions in terms of parameter selection. The parameters needed are (1) the number of ICA realizations, (2) the SCC threshold, and (3) the reproducibility index cut-off. We have demonstrated that RAICAR is not sensitive to the first two parameters, allowing flexibility in their selection, and the cut-off can be chosen by visual inspection of the reproducibility ranking. Therefore, we expect RAICAR to perform well for all users, regardless of experience level.

CONCLUSION

We have introduced an ICA method, RAICAR, based on reproducibility to improve the decomposition and inter-

TABLE V. Comparison of computation time between RAICAR and ICASSO

	162	240	360	440	520	600
Time length						
RAICAR	3 min	5 min 40 s	11 min	15 min 36 s	21 min	26 min 22 s
ICASSO	52 min	2 h 43 min	~9 h	~14 h	—	—

The results are generated on a workstation with RedHat™ Enterprise Linux WS4, 3.6G hyperthreading CPU, and 6 GB RAM.

pretation of fMRI data with ICA. RAICAR effectively minimizes the stochastic nature of individual ICA realizations. As demonstrated with both simulated and experimental data, RAICAR is insensitive to the choice of its parameters and has three primary strengths. First, it estimates the number of components. Second, it provides the order of the components, based on component reproducibility. Third, it leads to improved data decomposition by selectively averaging across ICA realizations.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Shing-Chung Ngan, Dr. Yihong Zhu, Dr. G. Andrew James, Dr. Tiejun Zhao, Christopher Glielmi, Jaemin Shin and Dr. Ying Guo for helpful discussions.

REFERENCES

Beckmann CF, Smith SM (2004): Probabilistic independent component analysis for functional magnetic resonance imaging. IEEE Trans Med Imaging 23:137–152.

Biswal B, Yetkin FZ, Haughton VM, Hyde JS (1995): Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 34:537–541.

Calhoun VD, Adali T, Giuliani NR, Pekar JJ, Kiehl KA, Pearlson GD (2006): Method for multimodal analysis of independent source differences in schizophrenia: Combining gray matter structural and auditory oddball functional data. Hum Brain Mapp 27:47–62.

Catalan MJ, Honda M, Weeks RA, Cohen LG, Hallett M (1998): The functional neuroanatomy of simple and complex sequential finger movements: A PET study. Brain 121:253–264.

Cox RW (1996): AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173.

De Luca M, Beckmann CF, De Stefano N, Matthews PM, Smith SM (2006): fMRI resting state networks define distinct modes of long-distance interactions in the human brain. Neuroimage 29:1359–1367.

Esposito F, Formisano E, Cirillo S, Elefante R, Tedeschi G, Goebel R, Di Salle F (2001): Criteria for the rank ordering of fMRI independent components. NeuroImage 13 (Suppl 1):114.

Esposito F, Formisano E, Seifritz E, Goebel R, Morrone R, Tedeschi G, Di Salle F (2002): Spatial independent component analysis of functional MRI time-series: To what extent do results depend on the algorithm used? Hum Brain Mapp 16:146–157.

Formisano E, Esposito F, Di Salle F, Goebel R (2004): Cortex-based independent component analysis of fMRI time series. Magn Reson Imaging 22:1493–1504.

Glover GH, Li TQ, Ress D (2000): Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. Magn Reson Med 44:162–167.

Gu H, Engelien W, Feng HH, Silbersweig DA, Stern E, Yang YH (2001): Mapping transient, randomly occurring neuropsychological events using independent component analysis. Neuroimage 14:1432–1443.

Hanakawa T, Immisch I, Toma K, Dimyan MA, Van Gelderen P, and Hallett M (2003): Functional properties of brain areas associated with motor execution and imagery. J Neurophysiol 89(2):989-1002.

Himberg J, Hyvarinen A, Esposito F (2004): Validating the independent components of neuroimaging time series via clustering and visualization. Neuroimage 22:1214–1222.

Huettel SA, Song AW, McCarthy G (2004): Functional Magnetic Resonance Imaging. Sunderland: Sinauer Associates. xviii, 492 p.

Hyvarinen A (1999): Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Networks 10:626–634.

Hyvarinen A, Oja E (2000): Independent component analysis: Algorithms and applications. Neural Networks 13(4/5):411–430.

Hyvarinen A, Karhunen J, Oja E (2001): Independent Component Analysis. New York: Wiley. xxi, 481 p.

Kiviniemi V, Kantola JH, Jauhiainen J, Hyvarinen A, Tervonen O. (2003): Independent component analysis of nondeterministic fMRI signal sources. Neuroimage 19 (Part 1):253–260.

LaConte SM, Ngan SC, Hu X (2001): Enhancing functional paradigm specific independent components with the AFRICA technique. In: ISMRM Ninth Scientific Meeting and Exhibition, April 21–27, Glasgow, Scotland, UK, 2001.

Liu JZ, Dai TH, Elster TH, Sahgal V, Brown RW, Yue GH (2000): Simultaneous measurement of human joint force, surface electromyograms, and functional MRI-measured brain activation. J Neurosci Methods 101:49–57.

Liu JZ, Zhang L, Yao B, Yue GH (2002): Accessory hardware for neuromuscular measurements during functional MRI experiments. Magma 13:164–171.

Lu W, Rajapakse JC (2003): Eliminating indeterminacy in ICA. Neurocomputing 50:271–290.

McKeown MJ, Sejnowski TJ (1998): Independent component analysis of fMRI data: Examining the assumptions. Hum Brain Mapp 6(5/6):368–372.

McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ (1998): Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Mapp 6:160–188.

Moritz CH, Rogers BP, Meyerand ME (2003): Power spectrum ranked independent component analysis of a periodic fMRI complex motor paradigm. Hum Brain Mapp 18:111–122.

Ogawa S, Menon RS, Kim SG, Ugurbil K (1998): On the characteristics of functional magnetic resonance imaging of the brain. Annu Rev Biophys Biomol Struct 27:447–474.

Peltier SJ, LaConte SM, Niyazov DM, Liu JZ, Sahgal V, Yue GH, Hu XP (2005): Reductions in interhemispheric motor cortex functional connectivity after muscle fatigue. Brain Res 1057(1/2):10–16.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001): A default mode of brain function. Proc Natl Acad Sci USA 98:676–682.

Van de Ven VG, Formisano E, Prvulovic D, Roeder CH, Linden DE (2004): Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. Hum Brain Mapp 22:165–178.