

A Method for Multitask fMRI Data Fusion Applied to Schizophrenia

Vince D. Calhoun,^{1–3*} Tulay Adali,⁵ Kent A. Kiehl,^{1,2} Robert Astur,^{1,2}
James J. Pekar,^{4,6} and Godfrey D. Pearlson^{1–3}

¹Olin Neuropsychiatry Research Center, Institute of Living, Hartford, Connecticut

²Department of Psychiatry, Yale University, New Haven, Connecticut

³Department of Psychiatry, Johns Hopkins University, Baltimore, Maryland

⁴Department of Radiology, Johns Hopkins University, Baltimore, Maryland

⁵Department of CSEE, University of Maryland Baltimore County, Department of CSEE,
Baltimore, Maryland

⁶FM Kirby Research Center for Functional Brain Imaging, Kennedy Krieger Institute,
Baltimore, Maryland

Abstract: It is becoming common to collect data from multiple functional magnetic resonance imaging (fMRI) paradigms on a single individual. The data from these experiments are typically analyzed separately and sometimes directly subtracted from one another on a voxel-by-voxel basis. These comparative approaches, although useful, do not directly attempt to examine potential commonalities between tasks and between voxels. To remedy this we propose a method to extract maximally spatially independent maps for each task that are “coupled” together by a shared loading parameter. We first compute an activation map for each task and each individual as “features,” which are then used to perform joint independent component analysis (jICA) on the group data. We demonstrate our approach on a data set derived from healthy controls and schizophrenia patients, each of which carried out an auditory oddball task and a Sternberg working memory task. Our analysis approach revealed two interesting findings in the data that were missed with traditional analyses. First, consistent with our hypotheses, schizophrenia patients demonstrate “decreased” connectivity in a joint network including portions of regions implicated in two prevalent models of schizophrenia. A second finding is that for the voxels identified by the jICA analysis, the correlation between the two tasks was significantly higher in patients than in controls. This finding suggests that schizophrenia patients activate “more similarly” for both tasks than do controls. A possible synthesis of both findings is that patients are activating less, but also activating with a less-unique set of regions for these very different tasks. Both of the findings described support the claim that examination of joint activation across multiple tasks can enable new questions to be posed about fMRI data. Our approach can also be applied to data using more than two tasks. It thus provides a way to integrate and probe brain networks using a variety of tasks and may increase our understanding of coordinated brain networks and the impact of pathology upon them. *Hum Brain Mapp* 27:598–610, 2006. © 2005 Wiley-Liss, Inc.

Key words: fMRI; schizophrenia; functional; brain; independent component analysis; ICA

Contract grant sponsor: National Institutes of Health; Contract grant numbers: 1 R01 EB 000840-01, 5 R01 MH43775-08, 5 R01 MH43775-08, and 1 R01 EB 005846-01.

*Correspondence to: Vince Calhoun, PhD, Olin Neuropsychiatry Research Center, Institute of Living, 200 Retreat Ave., Hartford, CT 06106. E-mail: vince.calhoun@yale.edu

Received for publication 27 January 2005; Accepted 20 July 2005

DOI: 10.1002/hbm.20204

Published online 8 December 2005 in Wiley InterScience (www.interscience.wiley.com).

INTRODUCTION

A typical functional magnetic resonance imaging (fMRI) experiment involves scanning participants during the performance of multiple tasks. These data are usually analyzed separately to detect differential activation of the two tasks [Derbyshire et al., 1998; Liu et al., 2004] or to activate robustly a desired set of regions [Hirsch et al., 2000]. Using such approaches, questions can be addressed about differential activation in the same voxels. These designs do not directly examine the joint (shared) information between different tasks and different voxels, however, even though it makes intuitive sense that fMRI activation in different tasks for the same individual will contain some shared information. Because the brain is a highly interconnected organ, it is reasonable to expect that functional changes in one area may result in or be related to modulations of brain activity in distant regions [Mesulam, 1998]. Computing such interrelationships is difficult in practice, however, due to the need to examine the relationship between tens of thousands of voxels.

There are multiple applications that may benefit from tools that enable the examination of joint information between tasks. For example, it may be useful to examine tasks (or conditions) that are theoretically related but probe slightly different aspects of a particular functional domain (such as cognition or sensory processing). Such work may reveal, for example, different regions in the two tasks that are subsets of the main effect of each task but are activated by the tasks similarly. This would then provide new evidence of a link between these regions, which could be probed further by additional experiments. Another approach, one that is demonstrated in this article, is to use tasks that probe very different functional domains, each of which may be modulated by, e.g., an underlying disorder such as schizophrenia. We expect patients with schizophrenia to share a common, widespread deficit resulting in activation differences that transcend individual cognitive domains. This viewpoint is supported by the schizophrenia literature showing differences in almost every type of cognitive or sensory task studied [Bullmore et al., 1999; Calhoun et al., 2004b; Laurens et al., 2003; Manoach et al., 2000; Schroder et al., 1999]. In this context, the use of tasks that probe different functional domains is advantageous, and joint analyses of these data may even help us to unify better the diversity of findings present in previous work.

Existing tools for examining joint information include region-based approaches such as structural equation modeling or dynamic causal modeling [Friston et al., 2003; McIntosh and Gonzalez-Lima, 1994]. For example, although it is most common to use structural equation modeling (SEM) to examine the relationship between different regions, these types of approaches can also be used to look at the correlational structure between regions activated by different tasks [Rajah and McIntosh, 2005] or between functional and structural variables [Meyer-Lindenberg et al., 2004]. However, these approaches do not provide an examination of the full set of brain voxels. A natural set of tools that avoid this

problem include those that transform data matrices into a smaller set of modes or components. Such approaches include those based upon singular value decomposition [Friston et al., 1996; McIntosh et al., 1996] as well as more recently, independent component analysis [ICA; McKeown et al., 1998].

ICA is a statistical and computational technique for revealing hidden factors that underlie sets of random variables, measurements, or signals [Hyvarinen et al., 2001]. The ICA model assumes that the data are linear mixtures of statistically independent sources and attempts to decompose the data into maximally independent components and their mixing coefficients (also called loading parameters). The ICA method is being increasingly utilized for fMRI data analysis to reveal hidden structure in the spatial and temporal dimensions of these data [Calhoun et al., 2003; McKeown et al., 2003]. In contrast to a first-level ICA approach (i.e., analysis of the preprocessing fMRI time series data for each task and each subject separately), we instead introduce the idea of a second-level (group), *feature-based* analysis of the fMRI activation maps (the “features”) generated from a first-level analysis. We propose a method that enables the decomposition of activation maps generated from two tasks into joint, maximally spatially independent components.

As a simple example, the approach we propose is to decompose the features into a set of “independent” building blocks or basis functions that are contributed to by the individual participants in a common manner. As an example of this application, consider the example of a cardboard cutout of a smiley face presented in Figure 1a in which each participant is associated with a light bulb of a certain wattage that illuminates various parts of a face (the parts that are illuminated [yellow] are determined by the “task”). The tasks result in the illumination of the right eye, the nose, and the smile (Task 1) or the left eye and the smile (Task 2) with an intensity determined by the participant-specific light bulb. The light bulbs have different wattages (so, e.g., the smiley faces corresponding to Participant 1 are illuminated by a 60-watt bulb and the smiley faces corresponding to Participant N are illuminated by a 40-watt bulb). If we carry out a joint ICA of these images by stacking the two task images side-by-side to form a row for each participant, then stacking the participant rows on top of one another and extracting spatially independent components (each of which has a Task 1 part and a Task 2 part), we would likely find one of the joint components as shown in Figure 1b (the Task 1 part is shown on the left and the Task 2 part is shown on the right), which has all the illuminated parts of the face represented in the component and the wattages reflected in the loadings (mixing coefficients). This then enables us to extract (potentially) different regions in two tasks, whose amplitude is “coupled.” If, on the other hand, we carried out ICA separately on the two tasks, we would have two sets of sources and loadings but it would not be clear how the two analyses related to one another.

Although this is an overly simplified example, it provides some intuition about the approach. Specifically, the corre-

sponding analogy for the brain is that the different parts of the face (right [R] eye + nose + smile or left [L] eye + smile) reflect (potentially) different neural circuits activated by the two tasks and the light bulb wattage (used in both tasks) reflects the common or shared contribution for a given participant (which can also be used to examine a systematic difference between two groups). In our implementation (presented below), smiley faces correspond to statistical parametric mapping (SPM) activation maps produced for a given task. Another aspect of the ICA decomposition is that joint components are by definition correlated to one another (due to the linear mixing requirement). The value of the mixing coefficient thus provides a relative measure of functional connectivity between detected regions at the group level.

We demonstrate our method on a data set consisting of two participant groups (schizophrenia patients and healthy controls) matched for age and gender, all of whom were scanned on two fMRI paradigms (the auditory oddball task [Kiehl et al., 2001] and the Sternberg working memory task [Manoach et al., 1997]). Both tasks have been found to show activation differences in schizophrenia patients. The auditory oddball evoked response is one of the most robust findings in electrophysiology and the associated fMRI involvement reveals robust differences in patients with schizophrenia by activating regions involved in target detection or attentional orienting [Kiehl and Liddle, 2001]. Working memory is also compromised in schizophrenia, and fMRI activation differences have been identified for the Sternberg working memory task [Manoach et al., 1999, 2000].

The proposed method offers a technique to examine (potentially different) regions in the two tasks, which co-vary between participants. We take a different approach to our previous work, which focused on using temporal lobe differences in spatially independent synchronous activity to classify schizophrenia patients and controls using a single fMRI task [Calhoun et al., 2004b]. The goal of the current study is not classification of patient groups, but rather the identification of a common network, jointly revealed by the data from two tasks, which is differentially involved in patients versus controls.

Schizophrenia is likely to be associated with disconnection [Andreasen et al., 1999; Lim et al., 1999; Pearlson, 1997; Stevens et al., 1998] and involves heteromodal association cortical regions, different portions of which are activated by the auditory oddball and Sternberg tasks. It is reasonable to expect that if schizophrenia is impacting a diffuse set of regions that are coordinated in the healthy brain and uncoordinated in the schizophrenic brain, then two tasks that activate different portions of this network would be related to one another (in both controls and patients, although to different degrees). We thus hypothesized that a small number of joint components would capture differences between the patients and controls and reveal a joint network present in both groups. We expected this network to show decreased activation (i.e., diminished functional connectivity) in patients. We also expected to find other regions that were more

active in patients, perhaps contributing to (or due to) inefficiencies in cortical processing.

Theory

Our goal is to examine activation across multiple tasks, collected on the same participants, in a unified analytic framework by modeling potential coupling between the different fMRI task data (the features). In our case, the features are the SPM “contrast” images (i.e., the estimated regression parameters) generated from the two tasks. We assume that a given feature can be represented as a set of (subject-wise) linearly mixed, spatially independent maps. This is similar to the assumption commonly used for ICA of fMRI [McKeown and Sejnowski, 1998]; however, because in this case we use features generated from a fitted hemodynamic model, the results are biased heavily toward the modeled fMRI activity. We additionally model data from multiple tasks as having a shared dependence, mediated through the mixing coefficients.

Consider the general matter of estimating common parameters from multiple data sources. First, let $X^{(1)}$ and $X^{(2)}$ be two data sets (e.g., $X^{(1)}$ are data from fMRI Task 1 and $X^{(2)}$ are data from fMRI Task 2). An effective way to fuse this data is to form the joint statistical likelihood and estimate parameters and hidden variables in that setting, rather than combining the parameters estimated separately on each data set. This has the advantage of efficiently using the information presented by all the data and can thus provide more accurate estimation of shared parameters [Hallouli et al., 2002; Webster et al., 1998]. In our joint ICA (jICA) framework, these shared parameters are coefficients of images that have high intensity on distinct coordinates (distinct brain regions), and they provide a way to assess the coupling among those regions.

Shared Participant Weights

To illustrate the above point, consider the case that there are C component sources, which individuals possess with certain loading coefficients $A_{i,c}$ for individual i and component c (thus we have that A is a shared parameter). Additionally, it is modeled that there are distinct source component outcomes for two tasks, which are realized as $S_c^{(1)}$ and $S_c^{(2)}$. The model for the image response measured at voxel v thus is described with a pair of *mixing* equations:

$$x_{i,v}^{(1)} = \sum_{c=1}^C a_{i,c} S_{c,v}^{(1)} \text{ and } x_{i,v}^{(2)} = \sum_{c=1}^C a_{i,c} S_{c,v}^{(2)} \quad (1)$$

where $x_{i,v}^{(t)}$ is the feature data from task t , $a_{i,c}$ is the shared mixing coefficient for component c , and $S_{c,v}^{(t)}$ is the spatial component for task t . We thus utilize a shared mixing coefficient, $a_{i,c}$, to model the coupling between the tasks and refer to this approach as jICA.

Alternatively, if we for the moment assume we have only one source for each task and write the *unmixing* equations

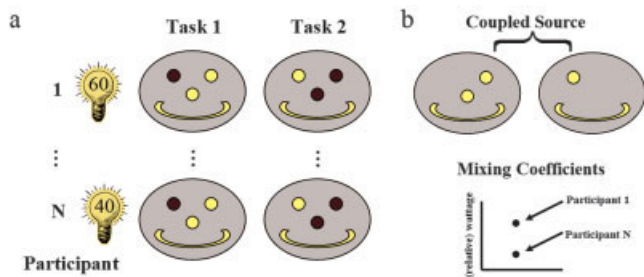


Figure 1.

A simple example of multitask or joint independent component analysis (ICA). Two “tasks” (left) have different portions of the smiley face that are illuminated, but share a common aspect (the light bulb wattage). Selected “watt-related” source on the right indicates the regions that are illuminated in each task and provides the shared mixing coefficients that reflect the wattage of the bulbs.

for the two tasks $s^{(1)} = wx^{(1)}$ and $s^{(2)} = wx^{(2)}$ (where $w = 1/a$), we can write the likelihood functions $p(x^{(1)}; w)$ and $p(x^{(2)}; w)$ [Kay, 1993]. If we were to run ICA on each task separately, we maximize these two likelihood functions in separate ICA analyses [Cardoso, 1997]. This would result in two sets of optimal unmixing coefficients, $w_1^* = \arg \max_{w_1} \log p(x^{(1)}; w_1)$ and $w_2^* = \arg \max_{w_2} \log p(x^{(2)}; w_2)$, one for each task. To interpret both results together, w_1^* and w_2^* would then have to be combined somehow as they are computed independently of one another. However, if we utilize a data fusion approach we determine a single optimal unmixing coefficient that maximizes the joint likelihood function, $w^* = \arg \max_w \log p(x(1), x(2); w)$. It makes intuitive sense not to compute the parameters independently, because the activation maps from the two tasks are coming from the same participant. We thus have a single w^* that fuses together the joint source (or alternatively, the basis vector common to the two tasks). The main advantage of this approach is that maximizing the joint likelihood function provides a different (and more reasonable) solution from one that does not utilize the joint statistics.

One may use existing ICA analysis algorithms to carry out a joint analysis by forming the overall data matrix $X = [X^{(1)}, X^{(2)}]$, stacking one beside the other (see Fig. 2), and likewise forming $S = [S^{(1)}, S^{(2)}]$ in which each of the original image component rows $S_c^{(1)}$ and $S_c^{(2)}$ are placed adjacent to form a total combined row of length $2V$ (the number of voxels in two images). Additional numbers of tasks may be handled similarly. The identification of components with shared loading parameters and the comparison of the associated maps is a key means to identify coupling between brain image components of different tasks.

For comparison purposes, we now briefly review fMRI group analysis carried out on the spatiotemporal fMRI data. In a spatial ICA analysis of group fMRI data, there are essentially two ways of staking the data, in the vertical (time) or horizontal (space) directions. Vertical stacking of

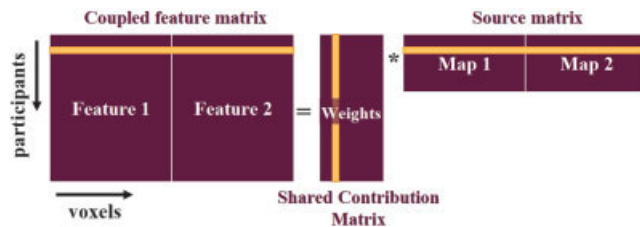


Figure 2.

Model in which loading parameters are shared for the hidden feature/source: The feature matrix is organized by placing the features (statistic parametric mapping (SPM) maps) from the two tasks side by side. This matrix is then modeled as separate task source images that share common mixing matrix parameters.

fMRI data (i.e., assuming common intragroup sources but unique time courses/weights) was proposed in Calhoun et al., [2001] (this approach also introduces a back-reconstruction step to generate source maps and time courses for individual subjects [Calhoun, 2004; Group ICA of fMRI Toolbox, online resource <http://icatb.sourceforge.net>]) whereas horizontal stacking (i.e. assuming shared time courses/weights, but unique components) was proposed in Svensen et al. [2002] and both were compared in Schmithorst and Holland [2004]. Alternative approaches assuming both shared and unique components [Lukic et al., 2002], clustering single-subject results [Esposito et al., 2005], or stacking into a cube via tensorial methods [Beckmann and Smith, 2005] have also been proposed. All of the approaches just mentioned utilize the entire multisubject spatiotemporal data set from a single task in various ways. In contrast, our second-level approach involves first selecting a single summary image (in our example a task-related activation map generated by *SPM2*) for each of two tasks and then stacking the task data side-by-side, thereby assuming a common mixing matrix. Additional approaches are also possible, for example, stacking the data vertically. Vertical stacking, however, imposes the rather strong assumption that there exists

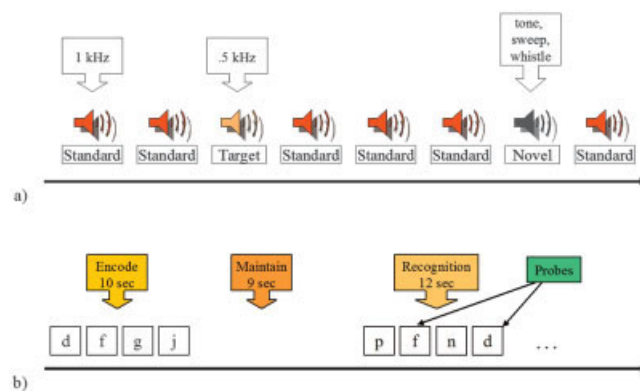


Figure 3.

Auditory oddball (AOD)/Sternberg (SB) paradigms: AOD (a) and SB (b) event-related fMRI tasks.

a common “source” map for both tasks. Because we already know that both tasks activate very different brain regions, this approach was not deemed appropriate for the current work.

SUBJECTS AND METHODS

Participants

Participants were recruited via advertisements, presentations at local universities, and by word-of-mouth. Fifteen healthy participants and fifteen outpatients with chronic schizophrenia, currently in complete or partial remission, provided written, informed, IRB-approved consent at Hartford Hospital and were compensated for their participation. Before inclusion in the study, healthy participants were screened to ensure they were free from Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) Axis I or Axis II psychopathology, assessed using the Structured Clinical Interview for DSM-IV [SCID; Spitzer et al., 1996], and were interviewed to determine that there was no history of psychosis in any first-degree relatives. Patients met criteria for schizophrenia in the DSM-IV based on a structured clinical interview and review of the case file [First et al., 1995]. All but one patient with schizophrenia were stabilized on atypical antipsychotic medications. There were equal numbers of males ($n = 12/12$) and females ($n = 3/3$) in the patient and control groups and all but two participants in each group were right-handed. There were no significant between-group differences in age (patients, 37 ± 11 years, range, 19–56 years; controls, 38 ± 11 years, range, 21–56 years). All participants had normal hearing (assessed by self-report) and were able to carry out both tasks successfully during practice before the scanning session.

Tasks

Auditory oddball

The auditory oddball (AOD) tasks consist of detecting an infrequent sound within a series of regular and different sounds. The tasks consisted of two runs of auditory stimuli presented to each participant by a computer stimulus presentation system (VAPP) via insert earphones embedded within 30-dB sound-attenuating MR-compatible headphones. The standard stimulus was a 500-Hz tone, the target stimulus was a 1,000-Hz tone, and the novel stimuli consisted of nonrepeating random digital noises (e.g., tone sweeps, whistles; Fig. 3). The target and novel stimuli each occurred with a probability of 0.10; the standard stimuli occurred with a probability of 0.80. The stimulus duration was 200 ms with a 1,000-, 1,500-, or 2,000-msec interstimulus interval (ISI). All stimuli were presented at approximately 80 dB above the standard threshold of hearing. All participants reported that they could hear the stimuli and discriminate them from the background scanner noise. The intervals between stimuli of interest (i.e., target and novel stimuli) were allocated in a pseudorandom manner to ensure that these stimuli had equal probability of occurring at 0, 1/3, and 2/3

after the beginning of a 1,500-ms image acquisition period. Because of this, the hemodynamic response to each type of stimulus of interest was sampled uniformly at 500-ms intervals. Before entry into the scanning room, each participant carried out a practice block of 10 trials to ensure understanding of the instructions. The participants were instructed to respond as quickly and accurately as possible with their right index finger every time they heard the target stimulus and not to respond to the nontarget stimuli or the novel stimuli. An MRI-compatible fiber optic response device (Lightwave Medical, Vancouver, BC, Canada) was used to acquire behavioral responses for both tasks. The stimulus paradigm data acquisition techniques and stimulus-related activation found previously are described more fully elsewhere [Kiehl and Liddle, 2001; Kiehl et al., 2005].

Sternberg working memory

The Sternberg working memory task (SB) consisted of a modified form of the Sternberg Item Recognition Paradigm [Manoach et al., 1997; Sternberg, 1966] that required subjects to memorize a list of alphabetic letters (consonants only), maintain the list in memory for several seconds, and then decide whether a probe letter was or was not a member of the memorized list. During each encoding phase, subjects were given a list of four, five, or six consonants, which were displayed sequentially for 1.5 s each with a 1-sec ISI. After a 9-s delay period, subjects saw a sequential series of probe letters (onscreen for 2.5 s with a 500-ms ISI) and were instructed to press a button with their dominant-hand index fingers for letters in the memorized list (targets) and to press a different button with their middle fingers for other letters (foils). The probes (targets) consisted of 4(2), 4(2), and 5(2 or 3) total probes (corresponding to the encoding of four, five, or six consonants, respectively). The total number of sets (consisting of an encode and probe condition) occurring over the 7-min run was four, five, and four, respectively. There was also a practice condition containing blocks of all possible memory loads. Before entering the MR scanner, all subjects were given complete task instructions and the practice condition. Practice and instructions were repeated as necessary until subjects achieved a high rate of correct responses on at least the blocks with lower memory loads.

Imaging Parameters

Scans were acquired at the Olin Neuropsychiatry Research Center at the Institute of Living on a 3T dedicated head scanner (Siemens Allegra; Siemens, Erlangen, Germany) equipped with 40mT/m gradients and a standard quadrature head coil. The functional scans were acquired using gradient-echo echo planar imaging (EPI) with the following parameters: repeat time (TR) = 1.50 sec (AOD)/1.86 s (SB), echo time (TE) = 27 ms, field of view = 24 cm, acquisition matrix = 64×64 , flip angle = 70 degrees, voxel size = $3.75 \times 3.75 \times 4$ mm, slice thickness = 4 (AOD)/3 (SB) mm, gap = 1 mm, 29 (AOD)/36 (SB) slices, ascending acquisition. Six

dummy scans were carried out at the beginning to allow for longitudinal equilibrium, after which the paradigm was automatically triggered to start by the scanner.

Data Analysis

Preprocessing

Data were preprocessed using the software package *SPM2* (<http://www.fil.ion.ucl.ac.uk/spm/>). Images were realigned using *INRIalign*, a motion correction algorithm unbiased by local signal changes [Freire and Mangin, 2001; Freire et al., 2002]. Data were spatially normalized into the standard Montreal Neurological Institute (MNI) space [Friston et al., 1995], spatially smoothed with a 12 mm^3 full width at half-maximum (FWHM) Gaussian kernel. The data, originally $3.4 \times 3.4 \times 4.5 \text{ mm}^3$, were slightly subsampled to 3 mm^3 , resulting in $53 \times 63 \times 46$ voxels. For display, even slices 8–38 are presented.

GLM analysis

Data for each subject were analyzed by multiple regression incorporating regressors for the novel, target, and standard (AOD) for the encode and recognition (SB) stimuli and their temporal derivatives plus an intercept term. Regressors were created by modeling the stimuli as delta functions convolved with the default *SPM2* hemodynamic response function (HRF). Only correct responses were modeled. The contrasts used in the jICA analysis included the AOD target-related contrast and the SB recognition-related contrast. The amplitude estimates from the first level analysis were corrected for amplitude bias due to spatially varying latencies [Calhoun et al., 2004c].

Joint ICA analysis

The algorithm for the jICA analysis proceeds with the following steps:

1. Feature selection: SPM contrast images for the AOD target stimuli and the SB recognition stimuli were generated using the *SPM2* software.
2. Feature normalization: To ensure similar consideration for both tasks, the features for the AOD and SB were normalized to have the same average sum-of-squares; the average sum-of-squares for each task was computed across all subjects and all voxels.
3. Feature matrix composition: In-brain voxels were analyzed and the two feature data sets were organized into a matrix as in Figure 2.
4. Dimensionality estimation: We used the minimum description length (MDL) [Calhoun et al., 2001; Rissanen, 1983] criteria to estimate the dimensionality of the feature matrix.
5. Dimensionality reduction: PCA was used to reduce the dimensionality of the data down to the estimated dimensionality (from step 4).
6. Spatial ICA decomposition: The extended-infomax algorithm [Bell and Sejnowski, 1995; Lee et al., 1999] was

used to decompose the reduced feature matrix to maximally independent component images and subject-specific mixing (loading) parameters.

7. Component selection: The eight resulting ICA mixing coefficients (columns of the shared contribution matrix in Fig. 2), each of which represents the relative degree to which an individual participant contributes to the joint component, were tested. We divided the participants into patients and controls, and tested within each component for a significant difference between patients and controls using a two-sample *t*-test. Only significant components were subsequently interpreted.
8. Component display: Joint components were re-converted into 3D images. The jICA analysis produces a set of different regions for each task (for clarity we call these jICA-AOD and jICA-SB regions).

Simulations

We examined the behavior of our algorithm by creating a hybrid data set in which a known source is mixed and added to actual SPM contrast images generated from the auditory oddball and Sternberg tasks. This data is then unmixed using the algorithm described in the previous section. Because the superimposed sources have a known pattern, it is straightforward to extract them from the unmixed data. Such an approach enables us to evaluate the performance of the jICA algorithm under a variety of contrast-to-noise (CNR) conditions, using data that has a complex structure (because it is partially composed of actual data). This is done by varying the amplitude of the known source. We can then evaluate the solution and different models by comparing the known unmixed sources to the ground truth using a measure such as correlation or Kullback-Leibler divergence between the known “truth” and the estimated joint sources [Calhoun et al., 2004a].

Figure 4a shows the generation of a hybrid-data experiment in which a known source (a 21×21 half-cycle sinusoid) was added to different parts of a single slice of auditory oddball and Sternberg data from 30 healthy individuals after multiplication by a random number drawn from a uniform distribution (the mixing parameter) with half of the individuals (the patients) having a mean shifted down by 0.5 from the other half of the individuals (the controls). The ranges used were [1.5–0.5] and [0–1], respectively. We generated hybrid data sets under a variety of CNRs (computed as the maximum value of the known source divided by the standard deviation of the fMRI data at the same voxel) by scaling the known source relative to the fMRI data. These data were then entered into a jICA analysis and the component that showed a significant difference between “groups” was examined.

Analysis of Patient and Control fMRI Data

For comparison with the jICA results, we carried out a standard random-effects analysis within *SPM2* by computing voxel-wise *t*-tests between the patient and control con-

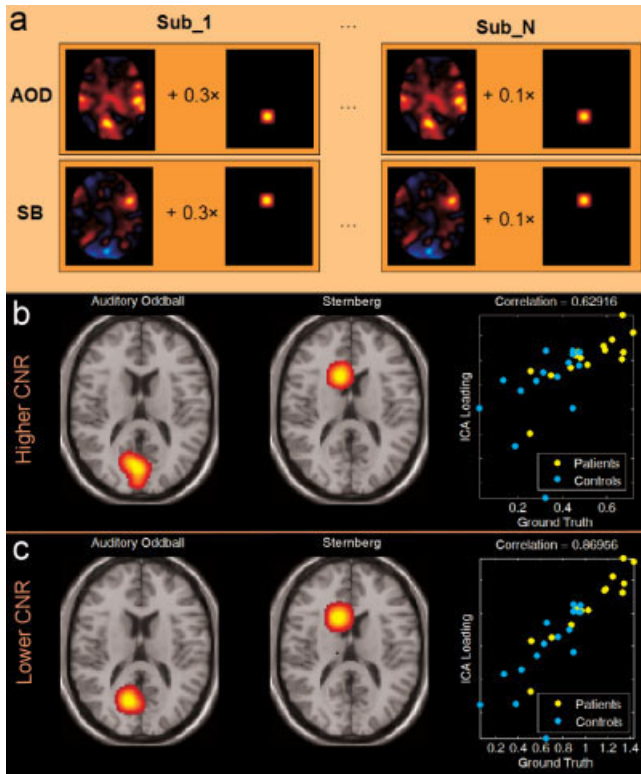


Figure 4.

Hybrid data generation and simulation results. Generation of hybrid data is depicted in Figure 4. **b, c:** Results from a lower and higher noise environment. The source that revealed the greatest difference between the two groups is shown for the auditory oddball (AOD) part of the joint source (left of Fig. 4b, c) and the Sternberg (SB) part of the joint source (middle of Fig. 4b, c). Mixing coefficient vs. the ground truth values are shown on the far right of Figure 4b, c.

trast images for each task [Holmes and Friston, 1998; Woods, 1996]. We then carried out a jICA analysis of the fMRI data from healthy controls and patients with schizophrenia. After the jICA analysis, we further examined the multitask data using a task-by-task histogram analysis. Voxels that were significant in the jICA analysis for either of the two tasks were used to generate a joint histogram of the AOD and SB fMRI data. These histograms were examined in individuals and as group averages.

RESULTS

Simulations

Results for a hybrid data set under a lower CNR situation (CNR = 0.5), comparable to the size of the signals found in our data, and a higher CNR situation (CNR = 1), are shown in Figure 4. The component that showed the largest group difference is displayed, with the AOD part of the joint source on the left, the SB part of the joint source in the middle, and

the correlation of the loading parameter with the ground truth on the right (patients are coded in cyan, controls in yellow). In both cases, the correct joint component was found and the controls showed a lower mean than the patients did (as expected). The jICA analysis thus selects out the coupled source into a separate component and enables us to visualize where in each data set the coupling occurs, as well as the loading parameters.

Behavioral Data

For the AOD task, performance and significances for whether controls and patients differed, was as follows: reaction time, controls 431.1 ± 102.2 ms and patients 522.3 ± 162.6 ms ($P < 0.01$); accuracy for target detection, controls $99.6 \pm 0.01\%$ and patients $97.7 \pm 0.05\%$ ($P > 0.2$). For the SB task performances were as follows: accuracy, controls 97.6% and patients 80.6% ($P < 0.0001$) and reaction time, controls 875.7 ± 140.3 ms and patients $1,067.5 \pm 281.2$ ms ($P < 0.0001$).

GLM Analysis

Translation and rotation corrections for each participant did not exceed half a voxel (i.e., 2 mm) or 2.0 degrees, respectively. We also qualitatively examined each statistical map to ensure there were no obvious motion artifacts (i.e., edge artifacts were not apparent, the pattern of activity looked in general like what was expected). There was no significant difference in movement between patients and controls. For comparison, both tasks were also analyzed using a standard random-effects analysis by entering the features into a voxel-wise one-sample *t*-test. Patients and controls analyzed separately ($P < 0.0005$, corrected for multiple comparisons using the false discovery rate [FDR; Genovese et al., 2002]) are shown in Figure 5 and differences between patients and controls (assessed with a two-sample voxel-wise *t*-test; $P < 0.05$, FDR corrected) are shown for the AOD and the SB tasks in Figure 6. Results are largely consistent with previous findings for both tasks [Kiehl and Liddle, 2001; Kiehl et al., 2005; Manoach et al., 1999, 2000].

Joint ICA Analysis

Results from the jICA analysis of both tasks are presented in Figure 7. Based upon the MDL criteria, eight components were estimated from the data. Only one component demonstrated significantly ($P < 0.00044$) different loadings in patients and controls (loading for controls was higher than that for patients), which we interpret as a difference in the degree/magnitude of functional connectivity in the two groups. Different spatial locations were identified for the two tasks. For display, AOD and SB sources were converted to empirical Z-values and thresholded at $|Z| > 3.5$. The AOD tasks showed only regions with controls versus patient increases (including temporal lobe structures and cerebellum) and the SB task showed some regions with control versus patient increases (basal ganglia and cerebellum) and others with control versus

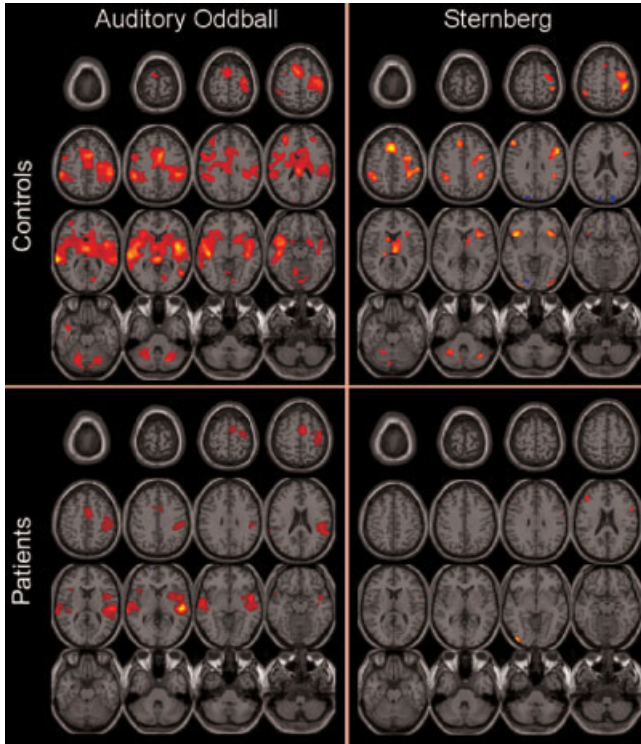


Figure 5.

Auditory oddball (AOD)/Sternberg (SB) group analyses. AOD target statistical parametric mapping (SPM) maps (left) and SB recognition SPM maps (right) for controls (top) and patients (bottom). Controls and patients both activated expected regions for both tasks, with patient activation being qualitatively less than controls.

patient decreases (visual cortex and occipitotemporal regions). Talairach coordinates for the AOD and SB jICA analyses are presented in Table I.

We can also examine more directly the relationship between the AOD and SB regions by going back to the SPM contrast images (i.e., the data that was input to the jICA analysis) and examining those regions that are contributing significantly to the jICA results. To examine the joint task activity in more detail, a joint histogram was computed as follows. Voxels surviving the threshold for the AOD part of the joint source were sorted in descending order by the component voxel values (the same was done for voxels in the SB part of the joint source). This procedure resulted in two sets of voxel coordinates. Histograms were then generated by pairing these two voxel sets. For example, the first point for Individual 1 is the voxel value for the AOD fMRI activation data (at the position that is maximum in the AOD part of the jICA source) versus the voxel value for the SB fMRI activation data (at the position that is maximum in the SB part of the jICA source). These pairings were used to generate single-subject 2D histograms of AOD fMRI signal (as estimated by the SPM contrast image) versus SB fMRI signal (as

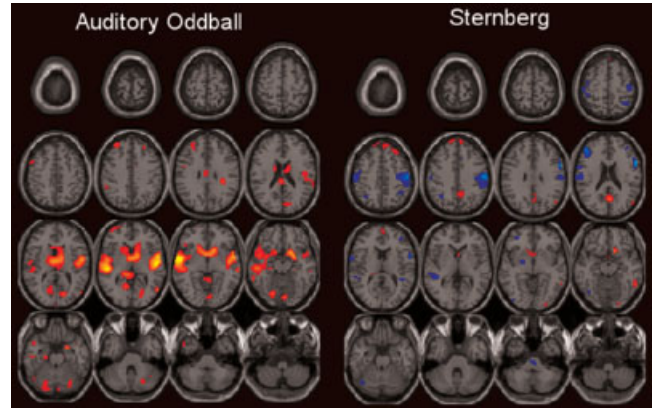


Figure 6.

Auditory oddball (AOD)/Sternberg (SB) group difference maps. Difference maps (controls minus patients) for the AOD (left) and SB (right) tasks. Patients demonstrated less activation in a variety of regions for the AOD tasks (consistent with previous findings) and demonstrated increased left DLPFC and decreased basal ganglia activation (also consistent with previous findings).

estimated by the SPM contrast image). The histogram image for each participant is shown in Figure 8a. In addition, we computed the within-group average of the histograms, and subtracted the control group average from the patient group average (shown in Fig. 8b with orange areas larger in controls and blue areas larger in patients). For the voxels included, the 2-D histogram can be considered an estimate of the joint distribution function for the two tasks (e.g., $p(f_{sb}, f_{aod})$, where $f_{sb,aod}$ indicates the fMRI signal amplitude for the AOD or SB tasks, respectively). We also computed the marginal estimated distributions $p(f_{sb}) = \sum_{aod} p(f_{sb}, f_{aod})$ and $p(f_{aod}) = \sum_{sb} p(f_{sb}, f_{aod})$ (Fig. 8c,d).

In general, more AOD task voxels were active in the controls and the SB task showed an increased kurtosis for

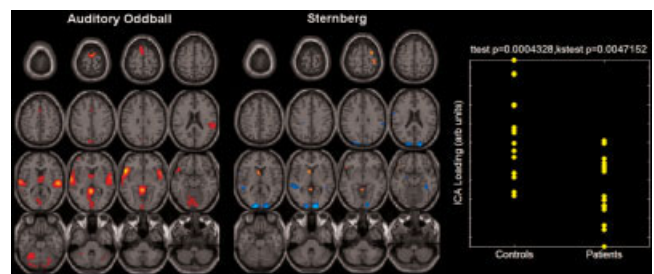


Figure 7.

Auditory oddball (AOD)/Sternberg (SB) joint independent component analysis (jICA). Only one component demonstrated a significant difference between patients and controls. The joint source map for the AOD (left) and SB (middle) tasks is presented along with the loading parameters for patients and controls (far right).

TABLE I. Talairach coordinates for auditory oddball and Sternberg joint independent component analysis (jICA)

Area	Brodman area	L/R volume (cc)	L/R random effects: Max T (x, y, z)
Auditory oddball			
Positive			
Cerebellum		2.3/0.6	9.5 (0, -44, 2)/8.3 (3, -44, 2)
Inferior frontal gyrus	47, *, 10, 45	3.3/0.5	8.8 (-53, 17, -6)/5.1 (53, 17, -6)
Superior temporal gyrus	38, 22, 41, *, 42, 21, 13, 29	11.0/10.8	8.8 (-53, 14, -6)/7.8 (56, -20, 9)
Parahippocampal gyrus	30, *, 27, 37	1.8/0.9	8.2 (-6, -44, 2)/6.8 (6, -41, 0)
Transverse temporal gyrus	41, 42	1.5/1.9	7.2 (-62, -14, 9)/7.9 (56, -17, 9)
Superior frontal gyrus	6	2.9/1.2	7.5 (0, 6, 66)/7.1 (3, 1, 69)
Postcentral gyrus	40, 43, *, 2	1.0/1.7	6.2 (-62, -17, 15)/6.4 (59, -23, 15)
Insula	40, 13, 29	0.3/2.4	4.7 (-45, 11, -3)/6.3 (53, -20, 15)
Precentral gyrus	13, 43, 6, 44	0.2/0.6	5.3 (-59, -8, 9)/6.0 (50, -14, 12)
Middle temporal gyrus	21, 22	1.2/0.1	6.0 (-62, 0, -5)/4.7 (59, -3, -5)
Medial frontal gyrus	6, 8, 32	0.8/0.1	5.9 (0, 3, 63)/5.1 (3, 3, 63)
Thalamus		2.5/2.1	5.8 (-6, -35, 2)/5.3 (6, -35, 2)
Lingual gyrus	18, 19, 17, *	5.7/2.2	5.6 (-3, -52, 0)/4.9 (6, -85, -11)
Fusiform gyrus	37, 18, 19, 20	3.9/1.0	5.5 (-33, -51, -20)/4.4 (30, -68, -19)
Inferior parietal lobule	40	0.0/0.5	ns/5.4 (59, -31, 21)
Inferior occipital gyrus	18, 17	0.4/0.1	5.1 (-24, -88, -13)/3.7 (9, -91, -8)
Caudate	Caudate head, Caudate body	0.5/0.6	4.5 (-3, 0, 3)/4.4 (9, 1, 11)
Cuneus	17, 19, 30, 18, 23, *, 7	3.6/1.0	4.3 (0, -81, 10)/4.2 (3, -84, 7)
Precuneus	7, 19	0.7/0.0	4.2 (-6, -77, 40)/ns
Middle frontal gyrus	10, *	0.4/0.0	4.2 (-42, 52, -3)/ns
Cingulate gyrus	32	0.6/0.0	4.0 (0, 16, 38)/ns
Sternberg			
Positive			
Cerebellum	—	0.4/0.1	5.9 (-3, -44, 2)/4.8 (3, -41, 2)
Caudate	—	1.5/0.1	5.3 (-6, 9, 5)/3.6 (3, 1, 11)
Inferior frontal gyrus	47, 46	0.6/0.0	5.0 (-48, 23, -11)/3.6 (48, 44, 12)
Precentral gyrus	4, 6	0.7/1.0	4.1 (-65, 4, 16)/4.4 (30, -23, 67)
Superior frontal gyrus	6	0.0/0.1	ns/4.2 (27, 0, 64)
Negative			
Cuneus	18, 17, 19, *	5.1/6.0	5.6 (-15, -98, 13)/5.4 (18, -98, 21)
Middle occipital gyrus	18, 19	1.7/2.3	5.5 (-18, -98, 13)/5.3 (21, -98, 19)
Parahippocampal gyrus	30, 37, 36, 19, amygdala	0.3/0.6	5.0 (-6, -44, 2)/4.0 (36, -38, -8)
Lingual gyrus	17, 18	0.1/1.1	3.7 (-15, -87, 4)/4.7 (12, -93, 2)
Middle frontal gyrus	6, 11	0.0/0.5	ns/4.5 (27, 3, 61)
Superior temporal gyrus	22, 38, 41, 29, 42, *	2.1/0.1	4.3 (-50, -38, 7)/3.6 (45, -41, 5)
Middle temporal gyrus	22, *, 21, 39	1.3/0.1	4.3 (-50, -40, 10)/3.6 (48, -41, 5)
Inferior parietal lobule	40	0.0/0.8	3.6 (-50, -59, 47)/4.1 (65, -33, 32)
Superior occipital gyrus	19	0.2/0.0	4.1 (-33, -86, 29)/ns
Fusiform gyrus	37, 20	0.0/0.6	ns/4.0 (39, -38, -8)

Voxels above the threshold for Figure 7 were converted from Montreal Neurological Institute (MNI) to Talairach coordinates and entered into a database to provide anatomic and functional labels for the left (L) and right (R) hemispheres. Both auditory oddball and Sternberg voxels are reported. The volume of activated voxels in each area is provided in cubic centimeters (cc). Within each area, the maximum T value and its coordinates are provided.

NS, not significant.

the patients (the number of voxels that were increased or decreased was greater). Upon further examination, these two “tails” were found to be driven by the blue regions in Figure 7 (which were negative in controls and positive in patients) and by the cerebellum and basal ganglia (which were positive in controls and negative in patients).

A final analysis was carried out after inspection of the histogram differences between the patients and controls. The correlation between the two tasks looked systematically different between the two. We thus fit a regression line to the voxels from the two tasks. Results are presented in Figure 9 and demonstrate, as seen from the histograms in Figure 8,

significantly more correlation between the two tasks in the patients ($P < 0.000085$).

DISCUSSION

We have demonstrated an approach for analyzing fMRI data from two tasks in a unified analytic framework using jICA. We also introduce the analysis of features generated from a subject-specific fMRI analysis. We utilize a well-established approach, *SPM2*, for generating fMRI activation maps, which are then decomposed into independent maps with shared loading parameters using jICA. This first-level

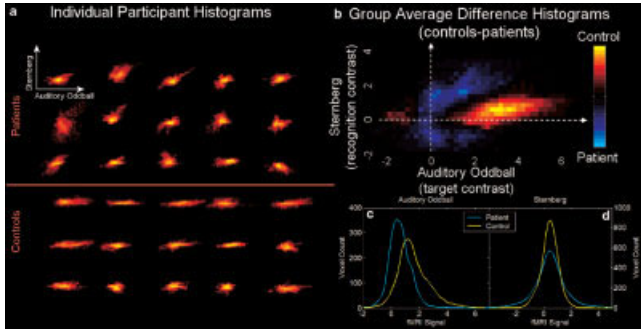


Figure 8.

Cross-task 2-D histograms. Joint 2-D histograms for voxels identified in the joint independent component analysis (jICA). Individual (a) and group average difference (b) histograms (with orange areas larger in controls and blue areas larger in patients) are provided along with the marginal histograms for the auditory oddball (statistical parametric mapping [SPM] contrast image for “targets”; c) and Sternberg (SPM contrast image for “recall”; d) data.

filtering effectively biases the results toward the modeled hemodynamic fMRI signal and is thus a reasonable way to begin to examine joint coupling between tasks. Using SPM to generate the features and then carrying out ICA is conceptually similar to assuming each subject has a fixed (modeled) hemodynamic time course (except for amplitude differences, which are modeled in the mixing matrix and vary for each participant).

Our novel data analytic approach revealed two interesting findings in the data that were missed with traditional analyses. First, consistent with our hypotheses, schizophrenia patients demonstrate decreased functional connectivity in the joint network identified using the jICA approach. This network includes regions in temporal lobe, cerebellum, thalamus, basal ganglia, and lateral frontal regions, and these findings are consistent with both the cognitive dysmetria [Andreasen et al., 1998] and frontotemporal disconnection [Liddle et al., 1992] models. The current analysis reveals a single network, including portions of all of the regions mentioned above, which is diminished/attenuated in activation amplitude. This supports the idea that the pathophysiology of schizophrenia includes impaired brain connectivity [Friston, 1999]. Contrary to our hypotheses, a component that was more active in patients than in controls was not revealed by the jICA analysis. Our findings thus argue against new, less-efficient areas being recruited by patients due to impaired connectivity between regions that are normally utilized. A second finding is that for the voxels identified by the jICA analysis, the correlation between the two tasks was significantly higher in patients than in controls. This finding suggests that schizophrenia patients activate “more similarly” for both tasks than controls. The degree to which a brain activation map is different from that of another task may reflect the degree to which performance on a task is specialized to a certain set of regions. A possible synthesis of

both findings is that patients are activating less, but also activating with a less-unique set of regions for these very different tasks. This suggests both a global attenuation of activity as well as a breakdown of specialized wiring between cognitive domains. Alternative explanations are also possible, and methods such as we propose may prove useful for further study. Nonetheless, both of the findings described support the claim that examination of joint activation across multiple tasks can provide insight into the misconnections in schizophrenia. Future studies will need to address the impact of performance differences between patients and controls (for example, by matching patients and controls on behavior as opposed to matching on task difficulty as done in this work) as well as other potential confounds common to patient studies, e.g., medication effects.

Identification of multinode networks can provide a useful tool for understanding the pathophysiology of schizophrenia as well as other psychiatric illnesses. This is a different goal than the useful process of classification using nodes that are known to be affected, such as the temporal lobe [Calhoun et al., 2004b]. In a previous work [Calhoun et al., 2004b], using solely temporal lobe function, we were able to correctly classify a new group of patients and controls with high accuracy using subject specific ICA maps of functional connectivity during the performance of a single fMRI task. The current work can be considered a precursor to developing potentially more sensitive methods for classification utilizing multiple tasks. Using subject-specific selected ICA features would also be a natural way to utilize the proposed

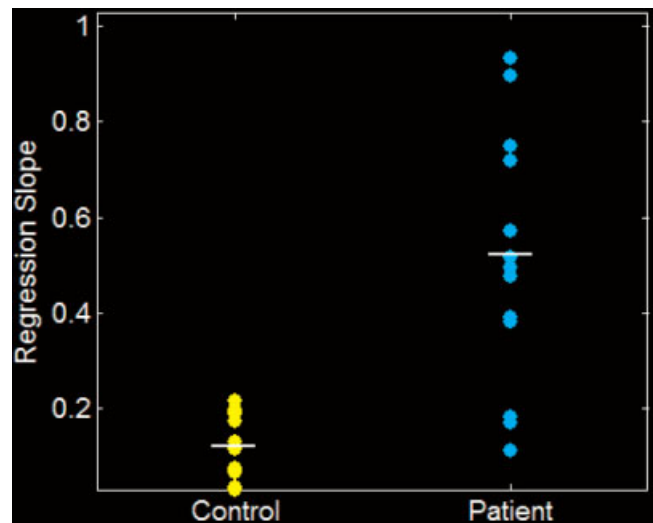


Figure 9.

Groupwise Sternberg (SB)/auditory oddball (AOD) regression slope. Fitted regression slope for patients (cyan) and controls (yellow) for the statistical parametric mapping (SPM) contrast data from regions that were revealed by the joint independent component analysis (jICA) approach. The SB task data was regressed onto the AOD task data and revealed a significant increase in slope for patients compared to that in controls.

framework, although for the present we have focused upon a model-based approach at the subject level.

Our proposed method has several advantages. First, our approach enables the joint analysis of multiple task data in a unified analytic framework. The framework presented, demonstrated on two fMRI tasks, is straightforward to use with two or more tasks or with other feature types such as gray matter segmentation images [Calhoun et al., 2005]. Secondly, we utilize a feature-based approach, providing a straightforward way to take advantage of data modeled at the subject level. These features are then queried for shared dependence, which is not detectable with a simple voxel-wise subtractive or conjunctive approach. Finally, the shared mixing coefficient provides a way to examine individual or group differences in coupling (which can be considered a measure of the relative degree of functional connectivity). Indeed, because we are looking directly at coupled differences between the two tasks, we need test only a single parameter (instead of thousands of voxels). Additionally, although our approach of carrying out ICA at the second (group) level is useful for a single task, carrying out a joint analysis improves the estimated sources, because more data is used in the estimation [Calhoun et al., 2005]. For the current data set, the significance also is increased when using the tasks in a joint analysis, providing evidence that joint task information is contributing to the results. These results provide additional evidence in support of the utility of jointly analyzing data from multiple tasks. Although we demonstrate our methodology on a data set in which vastly different tasks were used to examine widespread deficits related to schizophrenia expected to transcend cognitive/sensory domains, it may also be useful to apply the proposed methods to a set of tasks that are theoretically related to one another or involve common cognitive/sensory processes. The use of an approach that can examine joint information across multiple tasks may lead to improved understanding of a particular functional domain.

Although the modeling assumptions inherent in ICA have been explored to some degree for fMRI data [McKeown and Sejnowski, 1998], this has yet to be explored for our jICA approach. The main assumptions made are: (1) independence of the brain networks for both AOD and SB data, (2) a linear relationship between the subjects and these networks via the mixing parameters, and (3) the assumption of a common distribution for the AOD and SB joint sources. The first assumption has been used previously [McKeown et al., 1998] with good success with brain data and despite the fact that brain regions are not functioning independently due to the many brain regions and the sparsity of the sources, the independence assumption does seem to have some value. With spatial ICA, the stability of the results is dependent mainly upon the number of voxels/samples and thus even for a small number of individuals, the ICA solution is stable (although the significance testing between groups will of course suffer if the group sizes are too small). The linear relationship between subjects restated is that we assume

common networks/sources present in all subjects to (linearly) varying degrees. In this study, we test the hypothesis that patients and controls show differing network strengths, and find one component that is consistent with this hypothesis. The third assumption of a common joint distribution for AOD and SB sources is a sensible thing to do if one is interested in the examination of joint information. This assumption also is able to handle the case where a source may be contributed to by only the AOD data, in this case the SB part of the joint source would have insignificant voxel values. The aforementioned assumptions, however, although convenient, are possibly too limiting in several ways. First, the AOD and SB distributions may need to be modeled explicitly with different marginal distributions. To mitigate this concern in this initial jICA approach, we normalize the AOD and SB data and utilize the extended infomax algorithm that adaptively models the sources as having either supergaussian (e.g., a distribution with positive kurtosis) or subgaussian distribution. This algorithm has shown to be quite robust to violations of the underlying model for a wide variety of data types [Lee et al., 1999] and enables some flexibility in the source distributions. This is confirmed in our own data because upon examination of the distributions of the joint sources, we find that the distribution of the AOD and SB parts of the same source do show some variation (they have different means, variances, etc.), which is to be expected for these data. Finally, the linearity of the subject-wise contribution is an assumption of convenience, which, although it has worked very well thus far for ICA of many different data types, we would like to relax in future work because it is possible that the joint sources may also show nonlinear relationships.

There are also some additional limitations that should be mentioned. First, we are carrying out a joint second-level (group) analysis of features computed from separate first-level (subject-specific) analyses and thus some of the potentially useful information from the first level is not utilized in the source separation. Incorporation of a joint first-level analysis is thus an attractive goal for future work. The current framework, for practical reasons, assumes that both voxels and tasks are independent and identically distributed. Indeed, most ICA models used for fMRI data also make this assumption and perform quite well despite the known spatial correlation between voxels, provided appropriate data reduction is utilized [Hyvarinen et al., 2001]. Nonetheless, it can be potentially useful to incorporate some additional prior information on the voxels (such as spatial correlation) as well as incorporate different distributions for different features into the model. Our feature normalization step, which matches the variance of the two tasks, can potentially be improved by matching higher-order moments through, e.g., histogram equalization [Hansen et al., 2001]. In addition, the simulations using hybrid data have their limitations as well, although they demonstrate the utility and assumptions of the methods we propose and provide a measure of their performance in various conditions. We

hope in future work to utilize approaches related to those that have been used for fMRI to carry out cross-validation sampling [Strother et al., 2002]. Finally, although we have demonstrated some evidence supporting the usefulness of our model, our choice of modeling the shared dependence between the tasks with the mixing parameters should be examined in more studies before the true utility of this assumption will be known.

CONCLUSIONS

We have demonstrated a novel method for examining joint activation across multiple tasks. This approach has enabled us to ask novel questions about fMRI data and revealed several interesting findings in an application to data collected from healthy controls and patients with schizophrenia that were missed by a standard analysis approach. The development of models for jointly analyzing multitask fMRI data has been largely overlooked and may be a useful tool for assessing how brain function during different cognitive probes and in different regions can vary systematically between tasks.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (1 R01 EB 000840-01 and 1 R01 EB 005846-01 to V.C., 5 R01 MH43775-08, 5 R01 MH43775-08 to G.P.). We thank the research staff at the Olin Neuropsychiatry Research Center who helped collect and process the data, especially Nicole Giuliani, Matthew Johnson, and Nick Morris, as well as Matt Kurtz and Danny Mathalon for providing helpful input.

REFERENCES

Andreasen NC, Nopoulos P, O'Leary DS, Miller DD, Wassink T, Flaum M (1999): Defining the phenotype of schizophrenia: cognitive dysmetria and its neural mechanisms. *Biol Psychiatry* 46:908–920.

Andreasen NC, Paradiso S, O'Leary DS (1998): "Cognitive dysmetria" as an integrative theory of schizophrenia: a dysfunction in cortical-subcortical-cerebellar circuitry? *Schizophr Bull* 24:203–218.

Beckmann CF, Smith SM (2005): Tensorial extensions of independent component analysis for multisubject fMRI analysis. *Neuroimage* 25:294–311.

Bell AJ, Sejnowski TJ (1995): An information maximisation approach to blind separation and blind deconvolution. *Neural Comput* 7:1129–1159.

Bullmore E, Brammer M, Williams SC, Curtis V, McGuire P, Morris R, Murray R, Sharma T (1999): Functional MR imaging of confounded hypofrontality. *Hum Brain Mapp* 8:86–91.

Calhoun D, et al. (2004): Group ICA of fMRI Toolbox (GIFT). Online at <http://icatb.sourceforge.net>.

Calhoun VD, Adali T, Giuliani N, Pekar JJ, Pearlson GD, Kiehl KA (2005): Method for multimodal analysis of independent source differences in schizophrenia: combining gray matter structural and auditory oddball functional data. *Hum Brain Mapp*. DOI 10.1002/hbm.20166.

Calhoun VD, Adali T, Hansen JC, Larsen J, Pekar JJ (2003): ICA of fMRI: an overview. *Proc Int Conf on ICA and BSS*. Nara, Japan: IEEE.

Calhoun VD, Adali T, Pearlson GD (2004a): Independent components analysis applied to fMRI Data: a generative model for validating results. *J VLSI Signal Process Syst* 37:281–291.

Calhoun VD, Adali T, Pearlson GD, Pekar JJ (2001): A method for making group inferences from functional mri data using independent component analysis. *Hum Brain Mapp* 14:140–151.

Calhoun VD, Kiehl KA, Liddle PF, Pearlson GD (2004b): Aberrant localization of synchronous hemodynamic activity in auditory cortex reliably characterizes schizophrenia. *Biol Psychiatry* 55: 842–849.

Calhoun VD, Stevens M, Pearlson GD, Kiehl KA (2004c): FMRI analysis with the general linear model: removal of latency-induced amplitude bias by incorporation of hemodynamic derivative terms. *Neuroimage* 22:252–257.

Cardoso JF (1997): Infomax and maximum likelihood for source separation. *IEEE Lett Signal Process* 4:112–114.

Derbyshire SW, Vogt BA, Jones AK (1998): Pain and Stroop interference tasks activate separate processing modules in anterior cingulate cortex. *Exp Brain Res* 118:52–60.

Esposito F, Scarabino T, Hyvarinen A, Himberg J, Formisano E, Comani S, Tedeschi G, Goebel R, Seifritz E, Di SF (2005): Independent component analysis of fMRI group studies by self-organizing clustering. *Neuroimage* 25:193–205.

First MB, Spitzer RL, Gibbon M, Williams JBW. 1995. *Structured Clinical Interview for DSM-IV axis I disorders—patient edition (SCID-I/P, Version 2.0)*. New York: Biometrics Research Department, New York State Psychiatric Institute.

Freire L, Mangin JF (2001): Motion correction algorithms may create spurious brain activations in the absence of subject motion. *Neuroimage* 14:709–722.

Freire L, Roche A, Mangin JF (2002): What is the best similarity measure for motion correction in fMRI time series? *IEEE Trans Med Imaging* 21:470–484.

Friston K, Ashburner J, Frith CD, Poline JP, Heather JD, Frackowiak RS (1995): Spatial registration and normalization of images. *Hum Brain Mapp* 2:165–189.

Friston K, Poline JP, Strother S, Holmes A, Frith CD, Frackowiak RS (1996): A multivariate analysis of PET activation studies. *Hum Brain Mapp* 4:140–151.

Friston KJ (1999): Schizophrenia and the disconnection hypothesis. *Acta Psychiatr Scand Suppl* 395:68–79.

Friston KJ, Harrison L, Penny W (2003): Dynamic causal modelling. *Neuroimage* 19:1273–1302.

Genovese CR, Lazar NA, Nichols T (2002): Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15:870–878.

Hallouli K, Likforman-Sulern L, Sigelle M (2002): A comparative study between decision fusion and data fusion in Markovian printed character recognition. *Int Conf on Pat Recog*:30147.

Hansen LK, Nielsen FA, Strother SC, Lange N (2001): Consensus inference in neuroimaging. *Neuroimage* 13:1212–1218.

Hirsch J, Ruge MI, Kim KH, Correa DD, Victor JD, Relkin NR, Labar DR, Krol G, Bilsky MH, Souweidane MM, DeAngelis LM, Gutin PH (2000): An integrated functional magnetic resonance imaging procedure for preoperative mapping of cortical areas associated with tactile, motor, language, and visual functions. *Neurosurgery* 47:711–721.

Holmes AP, Friston KJ (1998): Generalizability, random effects, and population inference. *Neuroimage* 7(Suppl):754.

- Hyvarinen A, Karhunen J, Oja E. 2001. Independent component analysis. New York, NY: Johns Wiley & Sons. p 1–481.
- Kay SM. 1993. Maximum likelihood estimation. In: Gettman K, editor. Fundamentals of statistical signal processing, Vol 1. Upper Saddle River, NJ: Prentice Hall. p 157–204.
- Kiehl KA, Laurens KR, Duty TL, Forster BB, Liddle PF (2001): An event-related fMRI study of visual and auditory oddball tasks. *J Psychophysiol* 15:221–240.
- Kiehl KA, Liddle PF (2001): An event-related functional magnetic resonance imaging study of an auditory oddball task in schizophrenia. *Schizophr Res* 48:159–171.
- Kiehl KA, Stevens M, Laurens KR, Pearlson GD, Calhoun VD, Liddle PF (2005): An adaptive reflexive processing model of neurocognitive function: supporting evidence from a large scale (n=100) fMRI study of an auditory oddball task. *Neuroimage* 25:899–915.
- Laurens KR, Ngan ET, Bates AT, Kiehl KA, Liddle PF (2003): Rostral anterior cingulate cortex dysfunction during error processing in schizophrenia. *Brain* 126:610–622.
- Lee TW, Girolami M, Sejnowski TJ (1999): Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput* 11:417–441.
- Liddle PF, Friston KJ, Frith CD, Hirsch SR, Jones T, Frackowiak RS (1992): Patterns of cerebral blood flow in schizophrenia. *Br J Psychiatry* 160:179–186.
- Lim KO, Hedehus M, Moseley M, de Crespigny A, Sullivan EV, Pfefferbaum A (1999): Compromised white matter tract integrity in schizophrenia inferred from diffusion tensor imaging. *Arch Gen Psychiatry* 56:367–374.
- Liu X, Banich MT, Jacobson BL, Tanabe JL (2004): Common and distinct neural substrates of attentional control in an integrated Simon and spatial Stroop Task as assessed by event-related fMRI. *Neuroimage* 22:1097–1106.
- Lukic AS, Wernick MN, Hansen LK, and Strother SC (2002): An ICA algorithm for analyzing multiple data sets. *Int Conf on Image Processing (ICIP)*; Rochester, NY.
- Manoach DS, Gollub RL, Benson ES, Searl MM, Goff DC, Halpern E, Saper CB, Rauch SL (2000): Schizophrenic subjects show aberrant fMRI activation of dorsolateral prefrontal cortex and basal ganglia during working memory performance. *Biol Psychiatry* 48:99–109.
- Manoach DS, Press DZ, Thangaraj V, Searl MM, Goff DC, Halpern E, Saper CB, Warach S (1999): Schizophrenic subjects activate dorsolateral prefrontal cortex during a working memory task, as measured by fMRI. *Biol Psychiatry* 45:1128–1137.
- Manoach DS, Schlag G, Siewert B, Darby DG, Bly BM, Benfield A, Edelman RR, Warach S (1997): Prefrontal cortex fMRI signal changes are correlated with working memory load. *Neuroreport* 8:545–549.
- McIntosh AR, Bookstein FL, Haxby JV, Grady CL (1996): Spatial pattern analysis of functional brain images using partial least squares. *Neuroimage* 3:143–157.
- McIntosh AR, Gonzalez-Lima F (1994): Structural equation modeling and its application to network analysis in functional brain imaging. *Hum Brain Mapp* 2:2–22.
- McKeown MJ, Hansen LK, Sejnowski TJ (2003): Independent component analysis of functional MRI: what is signal and what is noise? *Curr Opin Neurobiol* 13:620–629.
- 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.
- McKeown MJ, Sejnowski TJ (1998): Independent component analysis of fMRI Data: examining the assumptions. *Hum Brain Mapp* 6:368–372.
- Mesulam MM (1998): From sensation to cognition. *Brain* 121:1013–1052.
- Meyer-Lindenberg A, Kohn P, Mervis CB, Kippenhan JS, Olsen RK, Morris CA, Berman KF (2004): Neural basis of genetically determined visuospatial construction deficit in Williams syndrome. *Neuron* 43:623–631.
- Pearlson GD (1997): superior temporal gyrus and planum temporale in schizophrenia: a selective review. *Prog Neuropsychopharmacol Biol Psychiatry* 21:1203–1229.
- Rajah MN, McIntosh AR (2005): Overlap in the functional neural systems involved in semantic and episodic memory retrieval. *J Cogn Neurosci* 17:470–482.
- Rissanen J (1983): A universal prior for integers and estimation by minimum description length. *Ann Stat* 11:416–431.
- Schmithorst VJ, Holland SK (2004): Comparison of three methods for generating group statistical inferences from independent component analysis of functional magnetic resonance imaging data. *J Magn Reson Imaging* 19:365–368.
- Schroder J, Essig M, Baudendistel K, Jahn T, Gerdens I, Stockert A, Schad LR, Knopp MV (1999): Motor dysfunction and sensorimotor cortex activation changes in schizophrenia: a study with functional magnetic resonance imaging. *Neuroimage* 9:81–87.
- Spitzer RL, Williams JB, Gibbon M. 1996. Structured Clinical interview for DSM-IV: Non-patient edition (SCID-NP). New York: Biometrics Research Department, New York State Psychiatric Institute.
- Sternberg S (1966): High-speed scanning in human memory. *Science* 153:652–654.
- Stevens AA, Goldman-Rakic PS, Gore JC, Fulbright RK, Wexler BE (1998): Cortical dysfunction in schizophrenia during auditory word and tone working memory demonstrated by functional magnetic resonance imaging. *Arch Gen Psychiatry* 55:1097–1103.
- Strother SC, Anderson J, Hansen LK, Kjems U, Kustra R, Sidtis J, Frutiger S, Muley S, LaConte S, Rottenberg D (2002): The quantitative evaluation of functional neuroimaging experiments: the NPAIRS data analysis framework. *Neuroimage* 15:747–771.
- Svensen M, Kruggel F, Benali H (2002): ICA of fMRI group study data. *Neuroimage* 16:551–563.
- Webster AM, Bridle SL, Hobson MP, Lansby AN, Lahav O, Rocha G (1998): Joint estimation of cosmological parameters from cosmic microwave background and IRAS data. *Astrophys J* 509:65–68.
- Woods RP (1996): Modeling for intergroup comparisons of imaging data. *Neuroimage* 4(Suppl):84–94.