NeuroImage: Clinical 12 (2016) 429–441

Contents lists available at ScienceDirect

NeuroImage: Clinical

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

Using concurrent EEG and fMRI to probe the state of the brain in schizophrenia

Judith M. Ford[a,b,*], Brian J. Roach[a], Vanessa A. Palzes[a], Daniel H. Mathalon[a,b]

[a]San Francisco VA Medical Center, 4150 Clement St, San Francisco, CA 94121, United States
[b]University of California, San Francisco, 505 Parnassus Ave, San Francisco, CA 94143, United States

ARTICLE INFO

Article history:
Received 31 March 2016
Received in revised form 20 July 2016
Accepted 9 August 2016
Available online 10 August 2016

Keywords:
Concurrent EEG + fMRI
Perception
N100
P200
Schizophrenia
Avolition/apathy

ABSTRACT

Perceptional abnormalities in schizophrenia are associated with hallucinations and delusions, but also with negative symptoms and poor functional outcome. Perception can be studied using EEG-derived event related potentials (ERPs). Because of their excellent temporal resolution, ERPs have been used to ask *when* perception is affected by schizophrenia. Because of its excellent spatial resolution, functional magnetic resonance imaging (fMRI) has been used to ask *where* in the brain these effects are seen. We acquired EEG and fMRI data simultaneously to explore *when* and *where* auditory perception is affected by schizophrenia.

Thirty schizophrenia (SZ) patients and 23 healthy comparison subjects (HC) listened to 1000 Hz tones occurring about every second. We used joint independent components analysis (jICA) to combine EEG-based event-related potential (ERP) and fMRI responses to tones.

Five ERP-fMRI joint independent components (JIC) were extracted. The "N100" JIC had temporal weights during N100 (peaking at 100 ms post-tone onset) and fMRI spatial weights in superior and middle temporal gyri (STG/MTG); however, it did not differ between groups. The "P200" JIC had temporal weights during P200 and positive fMRI spatial weights in STG/MTG and frontal areas, and negative spatial weights in the nodes of the default mode network (DMN) and visual cortex. Groups differed on the "P200" JIC: SZ had smaller "P200" JIC, especially those with more severe avolition/apathy. This is consistent with negative symptoms being related to perceptual deficits, and suggests patients with avolition/apathy may allocate too few resources to processing external auditory events and too many to processing internal events.

1. Introduction

1.1. Perception and schizophrenia

Perception involves the identification and interpretation of sensory information in the service of understanding and navigating the environment (Schacter et al., 2015). It is influenced by expectations and attention (Schroger et al., 2015) and results from a convergence of bottom-up and top-down processes (Joos et al., 2014), as the brain predicts the content and arrival of information (Friston, 2010). Its disruption in schizophrenia has been associated with a range of symptoms, the most obvious being hallucinations (Woodruff et al., 1995) and the inability to distinguish between what is real and what is not (e.g., Brebion et al., 1996). Recently, disruptions in perception have been associated with defeatist beliefs, negative symptoms, and poor functional outcome (Green et al., 2012).

Assessing perception has traditionally been done with a variety of behavioral methods, but is also studied using EEG-derived event related potentials (ERPs) (Joos et al., 2014; Woodman, 2010). Because of their excellent millisecond temporal resolution, ERPs have been used to inform us about *when* auditory processing is affected by a variety of psychological variables, such as attention and distraction (Hillyard et al., 1973; Hillyard et al., 1971; Näätänen and Picton, 1987). The different components of the ERP can also provide information about the transition from sensation to perception (Joos et al., 2014).

1.2. Brief history of ERPs to study perception

Over 50 years ago, ERPs were used in audiometry to assess hearing in people whose behavioral reports could not be obtained or trusted. A negative going potential, peaking 100 milliseconds (ms) after stimulus onset, was called N1 or N100. Because its amplitude increased with increasing loudness, N100 was considered a reasonable index of hearing. N100 is followed by P2 or P200, a positive going potential, peaking at about 200 ms. The N100-P200 complex was largest at the

* Corresponding author at: Mental Health Service, 116D, San Francisco VA Medical Center, 4150 Clement Street, San Francisco, CA 94121, United States.
E-mail addresses: Judith.Ford@ucsf.edu (J.M. Ford), Brian.Roach@ncire.org (B.J. Roach), Vanessa.Palzes@ncire.org (V.A. Palzes), Daniel.Mathalon@ucsf.edu (D.H. Mathalon).

http://dx.doi.org/10.1016/j.nicl.2016.08.009

vertex of the head and called the "vertex potential." For years, it was measured as the peak-to-peak difference between N100 and P200 (i.e. "N100-P200" or "N1-P2").

Although P200 invariably follows N100, they can be distinguished both experimentally (Ford et al., 1976; Ford et al., 1999; Oades et al., 1997; Wang et al., 2014) and topographically on the scalp (Vaughan et al., 1980; Verkindt et al., 1994; Wang et al., 2014). Yet, N100 and P200 often co-vary (Paiva et al., 2016). Thus, although structures that generate them may overlap to some extent, N100 and P200 waves are unlikely to reflect a single underlying neural process, and therefore, are best measured and studied independently of each other.

1.3. N100

In the 1970s, its sensitivity to attention (Hillyard et al., 1973) and arousal (Naatanen and Michie, 1979) shifted N100 out the realm of audiometry and into cognitive neuroscience. There is general consensus that N100 to a tone is augmented by selective attention, when that tone is in an attended channel (Hillyard et al., 1973). Data from a variety of sources suggest that N100 emanates generally from primary and secondary auditory cortical areas, namely superior temporal gyrus (STG) and middle temporal gyrus (MTG) (Chen et al., 2011; Flinker et al., 2010; Hari et al., 1987; Krumbholz et al., 2003; Ozaki et al., 2003; Pantev et al., 1996; Reite et al., 1994; Sams et al., 1985; Verkindt et al., 1995; Verkindt et al., 1994; Zouridakis et al., 1998), indirectly suggesting that attention to auditory events is associated with increased activity in these regions of the temporal lobe.

1.4. P200

The functional significance of P200 is poorly understood (Crowley and Colrain, 2004; Woodman, 2010). P200 may reflect an attention-modulated process required for the performance of an auditory discrimination task (Novak et al., 1992), or when elicited by a non-target stimulus in an oddball paradigm, it may reflect an attentional shift towards the stimulus and some aspects of the classification process (Garcia-Larrea et al., 1992). The brain areas responsible for P200 generation are also less well studied, but likely include both STG and MTG (Crowley and Colrain, 2004). Thus, although both are obligatory responses to tones, N100 and P200 might be considered reflections of different perceptual stages in the auditory processing stream.

1.5. N100, P200, and schizophrenia

N100 amplitude is typically, but not always, reduced in schizophrenia patients (see review by (Rosburg et al., 2008)). Indeed, its reduction has been proposed as a trait marker of functional brain changes related to genetic predisposition to schizophrenia (Ahveninen et al., 2006). N100 to probe tones may also be a state marker of the illness, as it is reduced during auditory hallucinations (Hubl et al., 2007), perhaps reflecting distraction by the voices. It may also reflect a readiness to attend to voices rather than probe tones in patients who tend to have auditory hallucinations (Ford et al., 2009). In spite of its prominence in the schizophrenia literature and its importance for understanding the pathophysiology of schizophrenia, the precise neural generators of N100 have not been adequately explored.

P200 amplitude reductions are sometimes (Ethridge et al., 2015; Roth et al., 1991; Roth et al., 1980; Salisbury et al., 2010), but not always (Potts et al., 1998) reported in schizophrenia. A meta-analysis indicated that P200s elicited by infrequent target tones are larger in schizophrenia, while P200s elicited by standard frequent tones are smaller (Ferreira-Santos et al., 2012), contributing to ongoing confusion about the relevance of this ERP component to understanding the pathophysiology of schizophrenia.

1.6. Goals of this experiment

To understand the different neural basis of N100 and P200, and how they are differentially affected by schizophrenia, we recorded EEG and fMRI data concurrently from patients and age-matched healthy controls. The EEG data provided the millisecond temporal precision needed to distinguish between rapidly resolving reflections of early (N100) and later (P200) stages of information processing, and the fMRI data provided the spatial/neuroanatomical precision needed to distinguish between areas of the brain involved in all stages of processing tones. Joint group Independent Components Analysis (jICA) allowed us to determine the patterns of spatial (with fMRI) and temporal (with ERP) covariance associated with processing tones. We focused on components in the temporal domain that load on traditional ERP waves (e.g. N100, P200) and their association with brain regions activated or even inhibited by tones in the spatial/neuroanatomical domain. Using jICA to integrate ERP and fMRI data allowed us to identify temporal-spatial relationships and their potential disruption in schizophrenia. We used a simple passive listening task to avoid confounding diagnostic effects with differences in cognition and motivation.

JICA is a blind source separation, unsupervised learning technique used to explain the underlying structure of multi-modal data. Thus, it is exploratory and data-driven in nature. Nevertheless based on the literature, we predicted that an "N100" joint component would reflect covariation of N100 amplitude and activity in STG, that a "P200" joint component would reflect covariation of P200 amplitude and activity in higher order cortical association areas. We also predicted that patients with schizophrenia would have smaller "N100" and "P200" joint independent components.

2. Materials and methods

2.1. Participants

Data are reported here from 30 patients with DSM-IV schizophrenia (N = 24) and schizoaffective disorder (N = 6) (hereinafter referred to as schizophrenia (SZ) patients), and 23 age- and gender-matched healthy comparison (HC) subjects (see below for description of why 8 SZ and 5 HC were dropped from the initial sample of 66 subjects.) Diagnoses were based on the Structured Clinical Interview for DSM-IV (First et al., 1995). Community outpatient clinicians referred SZ to us; both groups were recruited by advertisements and word of mouth.

Exclusion criteria for HC included any past or current major DSM-IV Axis I disorder based on a Structured Clinical Interview for DSM-IV Disorders, or having a first-degree relative with a psychotic disorder. For both groups, exclusion criteria were a history of a significant medical or neurological illness, head injury resulting in loss of consciousness, or substance abuse in the past 3 months. Additionally, HC did not have history of substance dependence (except nicotine), whereas SZ did not meet criteria for substance dependence within the past year. A psychiatrist or clinical psychologist conducted all interviews. Institutional Review Boards at the University of California at San Francisco and the San Francisco Veterans Affairs Medical Center approved the study, and all participants provided written informed consent. Clinical and demographic data are presented in Table 1 for those subjects included in the final analysis.

2.2. Clinical ratings

A trained research assistant, along with a psychiatrist or clinical psychologist, rated schizophrenia symptoms using the Scale for the Assessment of Positive Symptoms (SAPS) (Andreasen, 1984) and the Scale for the Assessment of Negative Symptoms (SANS) (Andreasen, 1983).

Table 1
Demographics and behavioral statistics of healthy controls and schizophrenia patients.

	Healthy controls N = 23 (6 women)					Schizophrenia patients N = 30 (7 women)				
	Mean	Median	SD	Min	Max	Mean	Median	SD	Min	Max
Age (years)[a]	37.4	38.0	13.6	21.7	60.7	39.2	38.9	14.5	19.1	63.2
Personal socioeconomic status[b]	30.8	29.0	10.7	11.0	54.0	46.5	47.0	14.2	22.0	69.0
Parental socioeconomic status[c]	32.0	29.0	15.8	11.0	63.0	29.7	30.0	14.4	11.0	65.0
Mean motion[d]	0.087	0.066	0.040	0.040	0.183	0.109	0.094	0.067	0.041	0.405
SAPS/SANS										
Global delusions						1.8	2.0	1.2	0	4
Global hallucinations						2.1	2.0	1.6	0	4
Global avolition/apathy						2.6	3.0	1.0	0	4
Global anhedonia/asociality						2.8	3.0	0.9	0	4
Global affective flattening						1.4	2.0	1.2	0	3
Global alogia						0.7	0	1.0	0	3
Global attention						1.4	2.0	0.9	0	3
Global thought disorder						1.4	2.0	1.3	0	4
Global bizarre behavior						0.8	1.0	0.9	0	3
Medication										
CPZ equivalents						442.4	250.0	416.7	50.0	1666.7
Antipsychotic medications						18	Atypical antipsychotic			
						5	Typical antipsychotic			
						0	Both antipsychotic			
						7	No antipsychotic			
Other psychiatric medications						10	Antidepressants			
						5	Anti-Parkinson's			
						8	Anxiolytics			
						2	Sedatives			
						2	Mood stabilizers			
						7	>1 other psychiatric medication			
						10	No other psychiatric medication			
Handedness	20 right, 2 left, 1 ambidextrous					28 right, 1 left, 1 ambidextrous				

[a] Groups did not differ in age, $t(51) = -0.455$, $p = 0.651$.
[b] HC had higher personal socioeconomic status than SZ, $t(51) = -4.402$, $p < 0.001$.
[c] Groups did not differ in parental socioeconomic status, $t(51) = 0.543$, $p = 0.589$.
[d] Groups did not differ in mean motion, $W = 538$, $p = 0.139$.

2.3. Experimental tasks

Subjects were instructed to, "Rest with your eyes open, look at a fixation cross, and listen to the sequence of tones you just generated." They listened to a series of 500 ms, 1000 Hz tones, presented in 16 seconds (s) blocks alternating with 16 s blocks of rest. The first and last rest blocks lasted 30 s each to facilitate EEG artifact correction procedures described below. There were two runs with 10 Listen and 10 Rest blocks. The individual sequences of tones presented during the Listen task had been generated by each subject earlier when subjects pressed a button every 1 to 2 s to deliver a tone (the Self task). EEG responses to tones collected in the MR scanner during the Self task cannot be analyzed because a button press artifact occurs in the EEG at the exact moment of tone onset. However, the fMRI data were not affected and will be described in a separate paper.

There were no significant differences in the number of tones between the groups, $t(51) = 0.80$, $p = 0.43$; HC (mean = 182 ± 35), SZ (mean = 171 ± 58). The interval between tones (in seconds) also did not differ between groups, $t(51) = -1.3$, $p = 0.22$, HC (mean = 0.95 ± 0.22), SZ (mean = 1.05 ± 0.35).

2.4. MRI data acquisition, preprocessing and analysis

We collected structural and functional MRI data using a 3 T Siemens Skyra scanner. The structural imaging protocol was a magnetization-prepared rapid gradient-echo (MPRAGE) T1-weighted high-resolution image (2300 ms TR, 2.98 ms TE, 1.20 mm slice thickness, 256 mm field of view, 1.0 × 1.0 × 1.2 voxel size, flip angle 9°, sagittal orientation, 9:14 min). The fMRI protocol was an AC-PC aligned echo planar imaging (EPI) sequence (2000 ms TR, 30 ms TE, flip angle 77°, 30 slices collected sequentially in ascending order, 3.4 × 3.4 × 4.0 mm voxel size, 182 frames, 6:08 min).

Image preprocessing was done using Statistical Parametric Mapping 8 (SPM8; http://www.fil.ion.ucl.ac.uk/spm/software/spm8/). First, motion correction was performed via affine registration of all runs, where the first image of each run was realigned to the first image of the first run, and then re-alignment proceeded within each run. Next, images were slice-time corrected with respect to the middle slice to adjust for timing differences of individual slices within each TR. To further denoise the data, we implemented aCompCor (anatomic component based noise correction) (Behzadi et al., 2007), a principal components-based approach to noise reduction of fMRI time-series data. ACompCor derives principal components from the time series of voxels within noise regions of interest (ROIs) defined on eroded white matter and cerebrospinal fluid (CSF) parcels from participants' segmented high-resolution T1-weighted anatomical images. To derive these white matter and CSF noise ROIs, the structural MRI scan was segmented using SPM 8, yielding a white matter map that was thresholded at $p > 0.99$ and eroded by 2 voxels, and a CSF map that was thresholded at $p > 0.9$ and eroded using a 3D nearest neighbor clustering criterion of at least two neighbors. A binary union mask of these white matter and CSF noise ROIs was generated and co-registered to the mean functional scan. Before proceeding with a principal component analysis (PCA) of the functional time series data of voxels contained in this co-registered white matter/CSF noise ROI mask, voxels in the mask that showed even weak relationships with the task regressors ($p < 0.2$) were excluded. Time series data for the remaining voxels in the noise ROI mask were then subjected to a PCA, and a number of noise (principal) components comprising weighted averages of white matter and CSF voxel time series were extracted. The number of components extracted was determined for each subject using a version of the broken stick method, a bootstrap procedure (Behzadi et al., 2007).

For individual participant (first-level) modeling of the fMRI data, an event-related analysis was implemented. SPM's canonical hemodynamic response function (a double gamma function) was convolved with

task event vectors to create a first-level task regressor representing the predicted response of a voxel activated by the presented tones. After applying a high pass temporal filter (128 s cut-off) to remove low-frequency noise, a general linear model was implemented voxel-wise, regressing each voxel's time series on the task regressor, as well as on nuisance regressors consisting of the aCompCor noise components. Parameters (i.e. beta coefficients) representing the fit of the task regressor to a voxel's time series were estimated, and resulting beta images, reflecting the contrast of tones versus implicit baseline (rest blocks), were then averaged over the two task runs. Next, mean beta images were normalized by applying the spatial transformation matrix derived from normalization of the mean functional image (generated during motion correction) to the Montreal Neurological Institute's EPI template (http://www.bic.mni.mcgill.ca). Normalized beta images were resliced to 3 mm^3 isotropic voxel dimensions using a fourth degree B-spline and were then spatially smoothed with a 6 mm full-width-half-maximum Gaussian kernel. Subsequently, each subject's first-level analysis brain mask image was normalized and applied to the smoothed beta image to mask out regions with insufficient signal intensity, as determined by SPM during first level modeling. In addition, a ventricular mask, generated using the Talairach Daemon database in Wake Forrest University's Pick Atlas, was used to mask out the lateral ventricles in the single subject normalized smoothed beta images.

2.5. EEG data acquisition, preprocessing, and analysis

Continuous EEG data were collected from 32 sites using BrainAmp MR plus, with high-input impedance specifically designed for recordings in high magnetic fields (BrainProducts, Munich, Germany). We used sintered Ag/AgCl ring electrodes with 5 kΩ resistors embedded in an electrode cap according to the 10–20 system (Falk Minow Services, Herrsching, Germany). An electrode was placed on the lower back to monitor electrocardiograms (ECG). Electrode impedances were kept below 10 kΩ. The nonmagnetic, battery powered, EEG amplifier was placed behind the MRI head coil and stabilized with sandbags. The subject's head was immobilized using cushions. EEG data were transmitted via a fiber optic cable to a BrainAmp USB Adapter that synchronized the EEG acquisition clock to the MRI master clock via a SyncBox (BrainProducts) before transferring data via USB to a laptop computer placed outside the scanner room.

All 32 channels were recorded with FCz as reference and AFz as ground to minimize the distance between reference and recording sites and to prevent amplifier saturation. The data were recorded with a bandpass filter of 0.01–250 Hz and digitized at a rate of 5 kHz with 0.5 μV resolution (16 bit dynamic range, 16.38 mV).

2.5.1. Removal of artifacts from EEG

EEG data were corrected for MR gradient artifacts by applying modified algorithms proposed by Allen et al. (2000) as implemented in Brain Vision Analyzer 2.0.4.368 software (BrainProducts). Next, the ECG channel was bandpass filtered 1–20 Hz to facilitate heartbeat detection and subsequent ballistocardiac artifact correction (see Appendix A for details.)

2.5.2. Canonical correlation-based denoising of EEG

Canonical correlation analysis (CCA) was used as a blind source separation technique to remove broadband or electromyography (EMG) noise from single trial EEG data using a method similar to that used by others (De Clercq et al., 2006; Ries et al., 2013) with some important differences. The CCA de-noising procedure involves correlating time series data from all channels with the one-sample time-lagged series from all channels, which is the multivariate equivalent of auto-regressive time series correlation. Each set of canonical correlation coefficients (one for each scalp electrode resulting in 31 for this study) has an associated time series (i.e. linear function of the coefficients and raw data called canonical variates). The fast Fourier transformed (FFT) power spectra of these canonical variate time series have been used to identify EMG components by taking the ratio of high (e.g. 15 to 30 Hz) to low (e.g. <15 Hz) power and removing components with ratios greater than a pre-determined limit (e.g. if high/low > 1/5 in (Ries et al., 2013)). This is a very rough heuristic for determining if a canonical variate's power spectrum has power-law scaling (e.g. $1/f^{\beta}$ or f^{α}, where $-\beta = \alpha$) where log-transformed power decreases linearly with increasing log-transformed frequency. Previous studies (Freeman et al., 2003; Pereda et al., 1998) have suggested that the exponent, α, is less than −1 in human EEG, while white noise or EMG would have an exponent of approximately zero.

Using simple linear regression, we estimated α by predicting log-power with log-frequency. For each trial and canonical variate, a bootstrap confidence interval was constructed for the estimated α by randomly sampling, without replacement, half of the frequency bins between 1 and 125 Hz from the FFT one thousand times to avoid potential contamination by a few frequencies (i.e. 60 Hz or alpha-band). If the interval contained values less than −1, the component was retained while all others were algebraically removed during back-projection to the original EEG epoch space (see Appendix A for details.)

2.5.3. Final EEG denoising, artifact removal, and ERP derivation

Single trial EEG data were re-referenced to an average reference, and data were obtained for the prior reference channel, FCz. Shorter, −100 to 250 ms epochs were created for the final processing steps. Outlier trials were rejected based on previously established criteria (Ford et al., 2014; Nolan et al., 2010). Prior to conducting an independent components analysis (ICA) of the EEG data, we initially conducted a principal components analysis (PCA) on the stacked single trial 350 time-point × 32 channel EEG data matrices to reduce the data. The number of principal components to extract was estimated separately for each subject using a parallel test approach (Parmet et al., 2010). This involved comparing the eigenvalues from the PCA of the stacked, single trial EEG data correlation matrix to eigenvalues from a PCA of a thousand completely simulated random normal data matrices of equal rank. A 95% cutoff, derived from the simulated data eigenvalue distribution, was used to determine how many components to extract from the EEG data. All subjects had 3, 4, or 5 components with eigenvalues exceeding this cut-off, and neither the number of components (HC: 3.61 ± 0.737, SZ: 3.45 ± 0.602; $t(64) = 0.9686$, $p = 0.3364$), nor the percentage of variance explained by those selected components (HC: 91.1 ± 3.7%, SZ: 92.6 ± 2.8%; $t(64) = -1.74$, $p = 0.0824$) significantly differed between groups. ICA was then performed on each subject's single trial EEG data in EEGLAB (Delorme and Makeig, 2004) with initial PCA dimension reduction (i.e. number of components to extract from EEG) set according to the previous estimate. Two researchers independently inspected individual component properties (i.e. fronto-central N100-P200 spatial loading, auditory evoked potential time course of ERP average from the component's single trial epochs, and 1/f power spectrum) to identify ERP-related independent components (ICs) and to exclude artifact-related ICs, and then they conferred to achieve consensus on which ICs to retain. Data from 13 (5 HC, 8 SZ) subjects were dropped from further analysis because zero components were selected (i.e. none of the ICs satisfied the criteria used to identify EEG/ERP ICs). The mean framewise displacement in this group of 13 participants suggested significantly greater mean motion values (median = 0.1656) than the 53 participants included in the analysis (median = 0.0856) using a Wilcoxon rank sum test ($Z = 3.466$, $W = 651$, $p = 0.0005$). Retained IC single trial data were then back-projected and averaged to generate a separate cleaned ERP for each subject and channel. Resulting ERP waveforms were low-pass filtered at 30 Hz.

2.6. Integration of EEG and fMRI data

We conducted a joint independent components analysis (jICA) using the Matlab-based Fusion ICA Toolbox or "FIT" (Calhoun et al., 2006). Joint ICA is a multi-modal data reduction technique that has been used to combine ERP and fMRI data collected simultaneously (Liebenthal et al., 2013) or in separate sessions (Wynn et al., 2015). FIT uses each subject's ERP and fMRI data (i.e. ERP time courses and fMRI beta maps) concatenated side-by-side. The ERP time course is resampled to a higher rate such that the number of ERP time samples and the number of fMRI spatial samples are equal. Sources associated with these modalities are assumed to co-vary the same way across groups and subjects (i.e. equal linear covariation).

Prior to calculating the jICA, the number of components in each imaging modality was estimated. FIT estimated the number of fMRI components to be 5 based on the minimum description length (Rissanen, 1978). The parallel test approach was used on the ERP data, indicating that there were 3 components in that modality. We chose the larger of these two numbers ($N = 5$) in the jICA, and each of the 5 components estimated had fMRI and ERP loadings that represented different patterns of spatio-temporal covariance between the modalities. Each subject had coefficients or scores associated with these 5 components' loadings that could be used in statistical tests. However, we first checked the validity of each component with a method similar to that used by others (Edwards et al., 2012). Specifically, the grand average ERP time course was regressed on the IC time course, and a peak (i.e. maximum absolute value) was picked from the fitted values of that regression model. There were only two components with monophasic waveforms and peaks that exceeded one standard deviation (0.6724 μV) of the grand average ERP. These two components captured the N100 and P200 peaks (Figs. 1 and 2), and are referred to as the "N100" and "P200" linked components to distinguish them from the N100 and P200 ERP components. Individual subject joint component scores for these linked components were retained for subsequent analyses. These scores were compared between groups using independent samples *t*-tests.

2.7. Relating clinical data to jICA data

To assess the relationship between symptoms and neurophysiological data, we correlated the "N100" and "P200" joint independent component scores (JICs) to the 9 global ratings from SANS and SAPS, listed in Table 1. The significance level for each component was Bonferroni corrected to $0.05/9 = 0.006$.

3. Results

Results from the jICA of fMRI and ERP responses to tones are presented in Figs. 1 and 2, for the "N100" and "P200" linked components, respectively.

3.1. "N100" linked component

As can be seen in Fig. 1 (left), a JIC reflected the joint activity of auditory cortex and the N100 component of the ERP. Activity in auditory areas in both the left and right STG and MTG (Fig. 1, right) were linked to N100 amplitude. A one-sample t-test against zero indicated that this ERP-fMRI association was robust ($t(52) = 15.1$, $p < 0.0001$). The details of the spatial maps are listed in Table 2. The groups did not differ in this JIC ($p = 0.76$). Importantly, no brain region was inversely linked to N100. That is, no region was significantly more active when N100 was smaller in magnitude.

3.2. "P200" linked component

Only the "P200" linked component showed a significant between-group difference ($t(51) = 2.05$, $p < 0.05$, Cohen's $d = 0.572$) on the group loading parameters, suggesting a difference in the combined activation of the linked fMRI/EEG brain features. Fig. 2 shows that the ERP portion of this linked component maps onto the P200 ERP component, which is larger in controls. Fig. 2 shows the anatomical activation associated with this linked component for all subjects. The regions depicted in hot colors reflect areas where the P200 ERP amplitude is positively correlated with activity in these areas. These areas include right and left STG and MTG; inferior, middle, and superior frontal gyri; inferior parietal lobule; and the declive and posterior lobe of the cerebellum. Details about these regions are listed in Table 3.

Unlike the N100 ERP component, the P200 ERP component was inversely linked to activity in several regions, depicted in cool colors in Fig. 2. This is prominent in two areas considered part of the default mode network (DMN): medial prefrontal cortex (MPFC) and posterior cingulate/pre-cuneus (PCC). That is, subjects with smaller P200 amplitudes have more activation in DMN regions. Also, smaller P200s are linked to greater occipital activation, suggesting that less auditory activity is related to more visual activity, in spite of the visual demands being

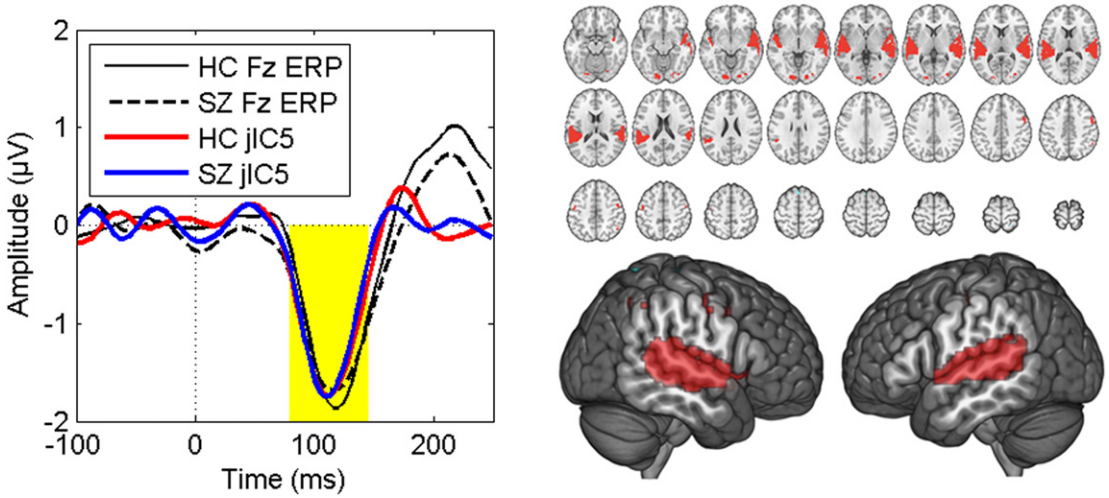

Fig. 1. Results from the fMRI/ERP jICA analysis at electrode Fz, showing the "N100" JIC. On the left are shown average ERP waveforms for HC (black solid line) and SZ (black dotted line) overlaid onto the temporal aspect of the "N100" JIC for HC (red) and SZ (blue). On the right, regions depicted in red reflect areas where the "N100" JIC is positively correlated with BOLD activation.

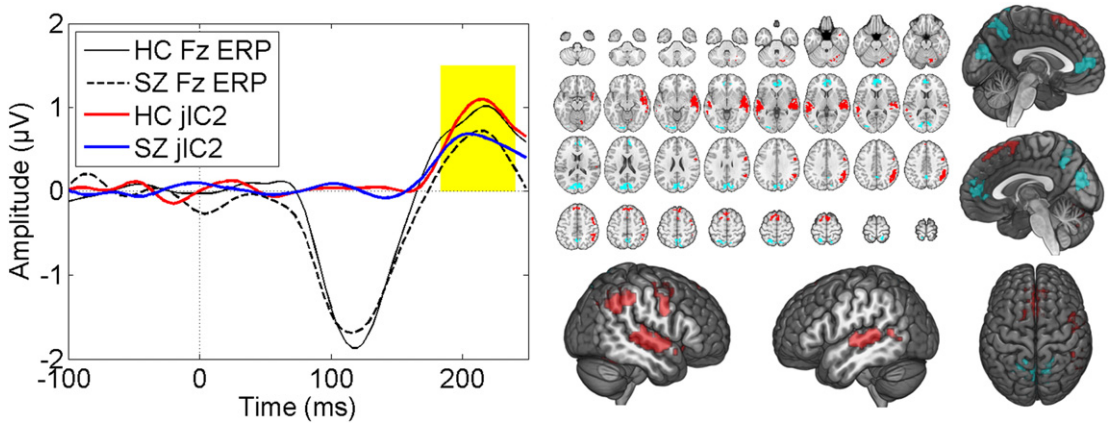

Fig. 2. Results from the fMRI/ERP jICA analysis at electrode Fz, showing the "P200" JIC. On the left are shown average ERP waveforms for HC (black solid line) and SZ (black dotted line) overlaid onto the temporal aspect of the "P200" joint independent component (JIC) for HC (red) and SZ (blue). On the right, regions depicted in red reflect areas where the "P200" JIC is positively correlated with BOLD activation, and those in cyan reflect areas where there is negative covariation between P200 amplitude and BOLD activation.

minimal. Table 4 lists anatomical details of these clusters. Together, the pattern of positive and negative associations between P200 amplitudes and cortical activations suggests that people with the biggest P200 amplitudes had both the *greatest* activation of STG/MTG and the *least* activation of the DMN and occipital cortex. While this relationship is also seen in SZ, as a group their mean "P200" JIC is smaller, due to more SZ than HC having negative "P200" JIC scores. That is, as a group, SZ have reduced P200 amplitudes and reduced STG/MTG activity along with increased DMN and occipital cortical activity.

3.3. "N100" vs. "P200" linked components

A comparison of the positive clusters from the "N100" and "P200" JICs can be seen in Fig. 3 where they are overlaid, in yellow and red, respectively. Although these components were not statistically compared to each other, this picture illustrates both their overlap and dissociability: "N100" and "P200" JICs are both linked to both STG and MTG. Larger "P200" JICs link to greater frontal and inferior parietal lobe activity, suggesting the contribution of higher-order cortical regions.

Table 2
Neuroanatomical areas positively linked to N100 ERP amplitude.[a]

Temporal lobe areas			
Left auditory cortex		Right auditory cortex	
# voxels	Region	# voxels	Region
630	Temporal lobe	720	Temporal lobe
526	STG	504	STG
106	Parietal lobe	147	MTG
95	BA 22	118	BA 22
81	Insula	71	Insula
56	BA 41	60	Parietal lobe
54	BA 13	47	BA 41
46	Postcentral gyrus	47	Postcentral gyrus
46	IPL	46	BA 21
42	TTG	37	TTG
29	BA 40	34	BA 13
27	MTG	33	Frontal lobe
21	Precentral gyrus	27	Precentral gyrus
18	Frontal lobe	17	BA 42
17	BA 42	17	BA 40
17	Supramarginal gyrus	10	BA 43
14	BA 43		

BA = Brodmann area; STG = Superior temporal gyrus; MTG = Middle temporal gyrus; IPL = Inferior parietal lobule; TTG = Transverse temporal gyrus; SMG = Supramarginal gyrus.
[a] Only clusters with ≥100 voxels and regions with ≥10 voxels are listed.

3.4. Clinical relationships with the "P200"linked component

The magnitude of the "P200" linked component correlated negatively with avolition/apathy ($r = -0.66$, $p < 0.0001$), such that SZ with the highest symptom scores had the lowest scores for this component. This relationship is plotted in Fig. 4.

To address which spatial features of the JIC contributed to the relationship, we extracted the mean parameter estimates from each subject's modeled data by first z-scoring the ICA loading spatial map, then by identifying all clusters with $|z| > 3.29$ ($p < 0.001$), and with at least 100 voxels. The zero-order correlation coefficients are presented on the left side of Table 5. Activity in MPFC showed the strongest relationship ($p = 0.01$), but it did not meet our Bonferroni significance threshold ($p < 0.006$).

To address whether each cluster made contributions to avolition/apathy while controlling for the others, we calculated the semi-partial correlations (right side Table 5). These revealed a trend for MPFC activation to correlate with avolition/apathy ($p = 0.05$, uncorrected for multiple comparisons) with more severely apathetic SZ having greater MPFC activity.

We also assessed the relationship between the temporal aspect of the JIC and avolition/apathy. To this end, we estimated P200 amplitude by first z-scoring the ICA loading waveform, then by identifying all points with $|z| > 1.96$ ($p < 0.05$) threshold in this waveform, which resulted in 9 samples centered on 212 ms (196 ms to 228 ms), and finally by averaging the data over this window. This estimate of P200 amplitude was also related to avolition/apathy, $r = -0.61$, $p < 0.0001$, suggesting that SZ with more severe avolition/apathy had smaller P200 amplitudes. None of the relationships between the other global symptom domains and the "P200" JIC survived the Bonferroni-corrected threshold, nor did the relationships between the "N100" JIC and the 9 symptom ratings.

4. Discussion

Our data suggest that large N100 amplitudes in response to a tone during passive listening are associated with strong, bilateral auditory cortex activation. This link between the auditory N100 and STG/MTG is consistent with magnetoencephalography (MEG) recordings of the N100m (Hari et al., 1987; Krumbholz et al., 2003; Ozaki et al., 2003; Pantev et al., 1996; Reite et al., 1994; Sams et al., 1985; Verkindt et al., 1994; Zouridakis et al., 1998), EEG-based N100 recorded from scalp

Table 3
Neuroanatomical areas positively linked to P200 ERP amplitude.[a]

Temporal lobe areas				Cerebellum	
Right auditory cortex		Left auditory cortex		Right cerebellum	
# voxels	Region	# voxels	Region	# voxels	Region
539	Temporal lobe	258	Temporal lobe	94	Posterior cerebellum
323	STG	174	STG	83	Declive
160	MTG	77	MTG		
76	BA 22	52	BA 22		
47	BA 21	11	BA 21		
23	Insula				
22	BA 38				
14	BA 13				
10	BA 41				

Frontal-parietal areas					
Right frontal lobe		Right parietal lobe		Superior frontal lobe	
# voxels	Region	# voxels	Region	# voxels	Region
101	Frontal lobe	273	Parietal lobe	137	Frontal lobe
101	Right cerebrum	189	IPL	123	SFG
51	Precentral gyrus	108	BA 40	39	BA 6
35	MFG	60	SMG	34	BA 8
28	BA 6	17	Angular gyrus	10	MFG
15	IFG	11	BA 39		

BA = Brodmann area; STG = Superior temporal gyrus; MTG = Middle temporal gyrus; IPL = Inferior parietal lobule; SFG = Superior frontal gyrus; MFG = Middle frontal gyrus; IFG = Inferior frontal gyrus; SMG = Supramarginal gyrus.
[a] Only clusters with ≥100 voxels and regions with ≥10 voxels are listed.

electrodes (Pantev et al., 1995; Verkindt et al., 1995), and intracranial recordings from the auditory cortical surface of patients being evaluated for treatment resistant epilepsy (Chen et al., 2011; Flinker et al., 2010). The lack of a group difference for the "N100" JIC indicates that both controls and patients process tones similarly during passive listening, 100 ms after stimulus onset, in the noisy scanner environment. There was no evidence that subjects with *smaller* N100s had *greater* activity in other areas of the brain, as was the case for P200. Although we predicted N100 amplitude would co-vary with activity in STG, our data suggest this extends to MTG, as well.

By 200 ms, there was a consistent pattern of co-variation between *increased* P200 amplitude, *increased* activation of bilateral STG/MTG, *increased activation* of frontal and parietal regions, and *reduced* activation of MPFC, PCC, and visual cortex. The involvement of frontal cortex in the "P200" linked component was unexpected, but was consistent with auditory evoked responses being recorded from the frontal eye fields in patients undergoing evaluation for surgical resection of epileptogenic tissue (Kirchner et al., 2009). To the extent that P200 reflects the allocation of perceptual resources, our data show that people who expend more perceptual resources to the tone during passive listening have less activation of the DMN and visual cortex. This is consistent with P200 amplitude reduction during meditation (Cahn and Polich, 2009).

Although the "P200" JIC was derived from all subjects, it was reduced on average in the SZ especially in those with more severe avolition-apathy. This is consistent with negative symptoms being related to perceptual deficits (Green et al., 2012), suggesting patients with avolition/apathy may allocate too few resources to processing the auditory events and too many resources to processing internal events and visual information of limited importance, like the fixation cross. This extends the earlier studies based on visual psychophysics to ERP-assessed perception during passive listening.

Although N100 recorded in ERP studies is often reduced in SZ (Rosburg et al., 2008), the "N100" JIC was not reduced in this sample. Given the nature of the "N100" JIC, reflecting both N100 and auditory cortical activation, it is difficult to know which is contributing to the

Table 4
Neuroanatomical areas negatively linked to P200 ERP amplitude.[a]

Midline areas				Parietal lobe	
Medial prefrontal cortex		Precuneus/occipital		Parietal/precuneus	
# voxels	Region	# voxels	Region	# voxels	Region
133	Frontal lobe	373	Occipital lobe	149	Parietal lobe
126	MPFC	258	Cuneus	94	Precuneus
103	Right cerebrum	73	BA 18	76	BA 7
71	Anterior cingulate	45	BA 19	39	SPL
71	Limbic lobe	44	MOG	14	Postcentral gyrus
57	BA 10	37	Precuneus		
30	BA 32	33	Lingual gyrus		
		22	BA 31		
		22	Parietal Lobe		
		11	IOG		
		11	BA 17		

MPFC = Medial pre-frontal cortex; BA = Brodmann area; MOG = Middle occipital gyrus; IOG = Inferior occipital gyrus; SPL = Superior parietal lobule.
[a] Only clusters with ≥100 voxels and regions with ≥10 voxels are listed.

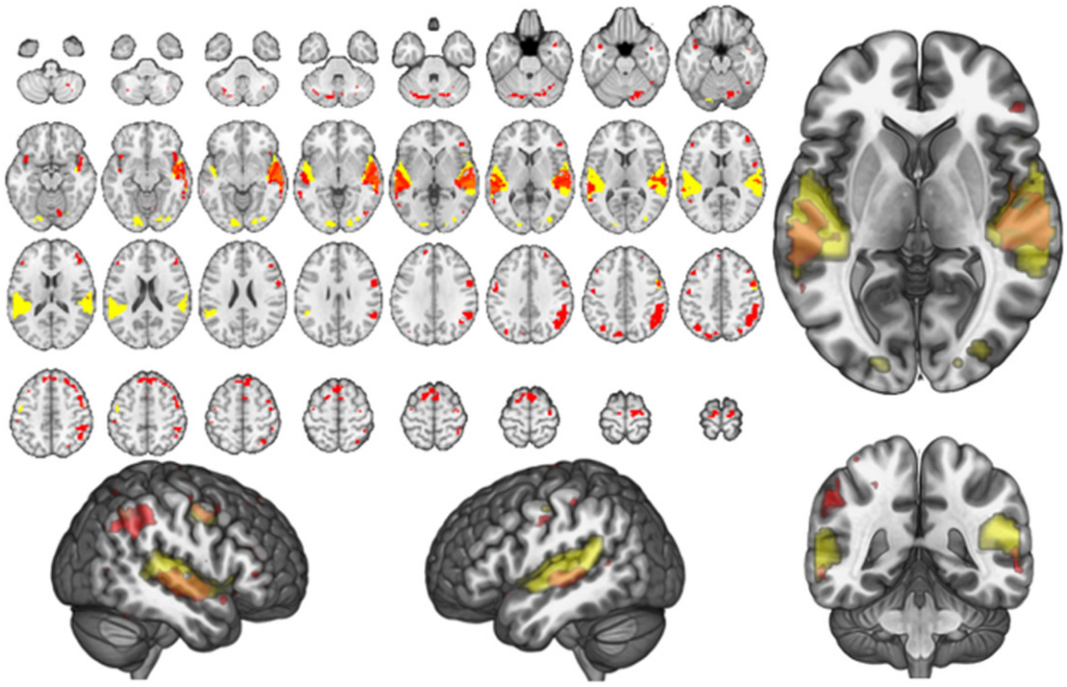

Fig. 3. "N100" and "P200" JICs are overlaid, in yellow and red, respectively, to illustrate their overlap and dissociability.

lack of a group effect. It is noteworthy that an fMRI study of schizophrenia patients passively listening to tones reported *more* auditory cortex activity in patients than controls (Mayer et al., 2013). Whether that increase in auditory cortical activity would have been related to increases in N100 or other ERP components occurring before or after N100 is difficult to know because of the poor temporal resolution of the BOLD response. Unlike N100, P200 amplitude reduction is not characteristic of schizophrenia, and group effects are seldom reported. Our finding of P200 amplitude reduction in schizophrenia joins a small literature reviewed in a recent meta-analysis (Ferreira-Santos et al., 2012).

The fact that both the "N100" and the "P200" linked components map onto auditory cortex may explain why they are sometimes considered a single component, the "N1-P2." However, our data suggest some important differences: while the "P200" linked component is positively linked with activity in frontal areas, it is also negatively associated with activity in the medial prefrontal and precuneus and occipital cortices, suggesting that perceptual processes in the auditory system are enhanced when activity in visual cortex and some nodes of the DMN are less active. Thus, this analysis strengthens the argument that P200 reflects more complex processes than N100, and it further distinguishes N100 from P200.

The term "DMN" has been used to describe a network of brain areas that are active during self-reflective activities such as "Random Episodic Silent Thinking" (REST) (Andreasen et al., 1995) or "stimulus-independent thought" (Fransson, 2005; Greicius et al., 2003; Raichle et al., 2001). However, Binder (2012) pointed out, "The term nicely captures the fact that these regions seem to return to an active state spontaneously whenever attention is not directed to an extrinsic input. On the other hand, the label says nothing about the nature of the information processing that characterizes this state, nor does it capture anything about the adaptive value of engaging these processes during resting

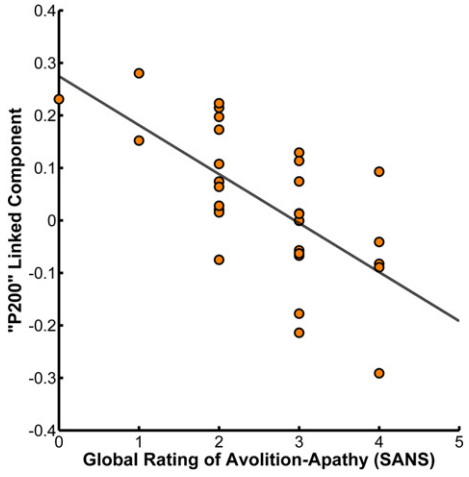

Fig. 4. The scatterplot showing a significant relationship between SANS global score on avolition/apathy and the "P200" linked component mixing matrix coefficients (arbitrary units), shown in Fig. 2. SZ with more severe avolition/apathy had smaller "P200" JIC magnitude. This also illustrates the point that while SZ have an overall smaller "P200" linked component than HC, some have positive values.

Table 5
Zero order and semi-partial correlations between ROI activations and avolition-apathy.

Cluster	Zero-order correlations		Standardized beta coefficients	Semi-partial correlations[a]	
	Coefficient	p-Value		Coefficient	p-Value
MPFC	0.46	0.01	0.68	0.36	0.05
PPC/occipital	0.23	0.22	0.17	0.08	0.66
Parietal/PPC	0.10	0.58	−0.29	−0.15	0.39
Cerebellum	−0.11	0.56	0.23	0.08	0.63
Right auditory cortex	0.18	0.34	−0.13	−0.06	0.72
Left auditory cortex	0.15	0.44	0.23	0.11	0.54
Right frontal	0.00	1.00	0.04	0.02	0.90
Right parietal	0.06	0.75	0.35	0.24	0.18
SPL	−0.18	0.36	−0.59	−0.29	0.10

MPFC = Medial pre-frontal cortex; PCC = Precuneus cortex; SPL = Superior frontal lobe.
[a] Each ROI controls for all remaining ROIs.

and predictable states. The term is also a misnomer because the cognitive processes that characterize this state are not unique to 'resting' and 'passive' conditions, but are also clearly engaged during many active tasks." Although there are reasons to question the assumption that the DMN is a "task-negative" network, nevertheless, it is normally suppressed during performance of difficult tasks. Consistent with that, patients with schizophrenia have been shown to have deficits in processing external information, proportional to activity in the default mode network (Whitfield-Gabrieli et al., 2009). Our results are consistent with Whitfield-Gabrieli et al. (2009), but extend those findings to tasks that put few cognitive demands on the participants. Specifically, our results suggest that some patients with schizophrenia allocate too few resources to processing external auditory events while allocating too many to processing internal events and visual information. This is especially true in patients with perceptual deficits, as reflected in P200 amplitude.

The DMN traditionally includes the midline structures, MPFC and PCC, as well as the inferior parietal lobule (IPL). However, in our data, P200 was negatively associated with MPFC and PCC but positively associated with IPL. In mindfulness trained subjects, IPL activates at the initiation of a thought, with MPFC providing subsequent elaboration (Ellamil et al., 2016), suggesting that these regions do not necessary co-activate temporally. Because we did not attempt to manipulate or probe internal states, it is difficult to know why IPL was positively associated with P200 in this analysis.

One limitation of this analysis is subject-loss due to our ICA component selection procedure. If this resulted in the loss of more clinically severe patients, group differences would be underestimated. Importantly, we lost both controls and patients in this procedure. Another limitation is our sparse recording montage (32 scalp sites); a denser montage might have provided greater opportunity to identify ERP components. Given our focus on auditory responses, clustered MRI acquisition might have improved our ability to image auditory responses. Although there was no behavioral read-out with our passive listening task (e.g. a button press response), it is naturalistic and similar to daily experiences. Finally, more than three-fourths of the patients were medicated, which might have affected the results; however, reductions of N100 (Rosburg et al., 2008) and P200 (Ferreira-Santos et al., 2012) amplitudes with antipsychotic medication are seldom found.

Acknowledgements and disclosures

This work was supported by grants from National Institute of Mental Health (MH58262 to JMF) and the VA (I01 CX000497 to JMF). DHM consults for Boehringer Ingelheim. The authors have declared that there are no conflicts of interest in relation to the subject of this study.

Appendix A

A.2.5.1. Removal of MR gradient artifacts from EEG

The correction of MR gradient artifacts in the EEG data involved subtracting an artifact template from each two-second EEG epoch time-locked to the onset of its concurrently acquired fMRI volume. A separate template was generated for each EEG epoch by averaging the 21 consecutive de-meaned EEG epochs centered on the to-be-corrected epoch, thereby including artifacts from the current epoch, the 10 preceding epochs, and the10 subsequent epochs, in the resulting template. By sliding this camped average of 21 EEG epochs along the time series of successive fMRI volume acquisitions and subtracting the resulting epoch-specific artifact templates from each EEG epoch, MR gradient artifacts were effectively removed. Subsequently, the EEG data were down-sampled to 250 Hz.

A.2.5.2. Removal of ballistocardiac artifacts from EEG

Data collected from an electrocardiogram (ECG) electrode was used to identify when heartbeats occurred in the continuous EEG signal. Semi-automatic heartbeat detection was performed in Brain Vision Analyzer 2, where heartbeats were identified in the ECG channel that exhibited a high temporal correlation ($r > 0.5$) with a heartbeat template (a heartbeat chosen by the algorithm occurring about 30 s into the run) and within-range amplitude (0.6–1.7 μV); trained research assistants adjusted templates and search windows, manually identifying pulses as needed. Ballistocardiac artifact removal was achieved in the same manner as gradient artifact removal, but now the event of interest was the heartbeat as originally proposed (Allen et al., 1998) and done by others (Schulz et al., 2015; Shams et al., 2015). Individual tone trials were then segmented into epochs from −1000 to 500 ms around the tone onset, baseline-corrected using −100 to 0 ms, and exported for additional processing in Matlab 2013a (Mathworks, Natick, MA).

A.2.5.3. Canonical correlation-based denoising of EEG

Canonical correlation analysis (CCA) was used as a blind source separation technique to remove broadband or electromyographic noise from single trial electroencephalographic (EEG) data, generating denoised EEG epochs. Our approach is similar to the CCA method described by others (De Clercq et al., 2006; Ries et al., 2013), with some important differences. The method is based on the concepts that true EEG data tend to show high auto-correlation and exhibit power-law scaling (i.e., power is proportional to 1/frequency), but that high frequency random noise in EEG (e.g., muscle artifact, electromyographic (EMG)) tends to show low auto-correlation and violates power-law scaling (i.e., inappropriately high power at higher frequencies relative to low frequencies). The CCA de-noising procedure is performed separately for each subject on the single trial EEG epoch data. For each 1500 ms epoch sampled at 250 Hz, the ($S \times X$) matrix containing the time series of $S = 375$ EEG amplitude values ($s_1, s_2, s_3 \ldots s_{375}$) at each of $X = 31$ scalp electrodes ($x_1, x_2, x_3 \ldots x_{31}$) is subjected to a CCA with the ($S \times Y$) matrix containing the $s + 1$ time-lagged series of 375 EEG amplitude values ($s_2, s_3, s_4 \ldots s_{375}, s_{376}$, where $s_{376} = 0$) at each of the same 31 electrodes ($y_1, y_2, y_3 \ldots y_{31}$). This is the multivariate equivalent of auto-regressive time series correlation. Since both the X and Y vectors each contain 31 electrodes, a total of 31 independent canonical correlations can be extracted. Each canonical correlation coefficient expresses the correlation of a time series of values representing the weighted sums of the X electrodes with a $s + 1$ time series of values representing the weighted sums of the Y electrodes, with weights chosen to yield the largest canonical correlation that accounts for variance independent of the variance accounted for by all previously extracted canonical correlations. Thus, each canonical correlation coefficient has an associated time series of values that constitutes the canonical variate, X (i.e. each time point has a value that is a linear function of the canonical weights and raw data associated with the 31 electrodes), as well as a similar canonical variate, Y. The current CCA de-noising method only makes use of the set of 31 canonical X variates, one for each of the 31 extracted canonical correlations. When the time series represented by a canonical variate is subjected to a fast Fourier transformation (FFT), the resulting power spectrum can be evaluated to determine whether the canonical variate conforms to the power-law expected from EEG data, in which case it should be retained, or whether it violates the power law as would be expected for high-frequency noise (e.g. EMG contamination), in which case it should be excluded. With this approach, the retained canonical variates are those showing the strongest canonical correlations, whereas the rejected canonical variates are those showing the weakest canonical correlations. The specific criterion used to make these retain/reject decisions is where our CCA denoising approach differs from previously published approaches (e.g. (De Clercq et al., 2006; Ries et al., 2013)).

Previously described CCA denoising approaches have made decisions about which canonical variates to retain or reject by taking the ratio of high frequency (e.g. 15 to 30 Hz) to low frequency (e.g. <15 Hz) power, rejecting canonical variates with ratios greater than a pre-determined limit (e.g. if high/low > 1/5 in (Ries et al., 2013)). Such ratios provide a very rough heuristic for determining whether a canonical variate's power spectrum has power-law scaling (i.e. $1/f^{\beta}$ or f^{α}, where $-\beta = \alpha$) where log-transformed power decreases linearly with increasing log-transformed frequency. Previous studies (Freeman et al., 2003; Pereda et al., 1998) have suggested that the power-law exponent, α, which corresponds to the slope of the linear regression of log-power on log-frequency, is -1 or less for human EEG, whereas white noise or EMG has an exponent of approximately zero. In order to develop a precise and conservative criterion for deciding which canonical variates to retain, we used simple linear regression to estimate α by regressing log-power on log-frequency. However, rather than running this regression once using all of the available log power and log frequency values obtained from the FFT of the canonical variate, thereby yielding a single estimate of α that could be unduly influenced by contamination from a few frequencies (e.g. 60 Hz or alpha band), a bootstrap procedure was used to generate a distribution of α estimates. Specifically, for each canonical variate and each trial epoch, the linear regression of log-power on log-frequency used to estimate α was repeated 1000 times by randomly sampling, without replacement, half of the frequency bins between 1 and 125 Hz from the FFT. If the distribution of 1000 α estimates, which constitutes a 99.9% bootstrap confidence interval for α, contains a value of -1 or smaller, then the canonical variate was retained, whereas if the smallest value in the bootstrap distribution is greater than -1, the canonical variate is rejected. All rejected canonical variates are then algebraically removed during back-projection to the original EEG epoch space, generating a denoised EEG epoch.

A.2.6.1. Joint independent analysis components that were not subjected to group or correlation analyses

Three of the five joint ICs were not included in group comparisons and correlation analyses based on the criteria described in the main text. It is worth describing these components for reference. Two of the three components quite clearly capture noise. For example, component 1 (Fig. A1) has a ring of "deactivation" around the edges of the brain and "activation" in the ventricles, consistent with a motion-related noise component, and the corresponding ERP time courses are relatively flat.

An additional component (component 3, Fig. A2) seems to have captured some low-level ERP activity in the first 100 ms and broadly distributed fMRI "activation" that may once again capture motion noise as well as bilateral auditory and visual cortex "deactivation".

Finally, component 4 (Fig. A3) appeared to capture N1-P2 transitions with many, poorly-defined clusters that to some extent overlapped with both the component 5 (N100 component) and component 2 (P200 component) spatial maps. Specifically, deactivation of visual cortex, precuneus, and pre-frontal cortex (similar to component 2) along with activation of bilateral STG/MTG (components 2 and 5) are present at the Z height threshold of $p < 0.001$. The waveforms have a positive peak at about 92 ms and a negative trough at approximately 156 ms. These waveforms do not appear to differ between subject groups, and since the loading waveforms appear to track the first temporal derivative of the grand average ERPs, they may capture some latency variability in the peak of either the N1 or P2 components.

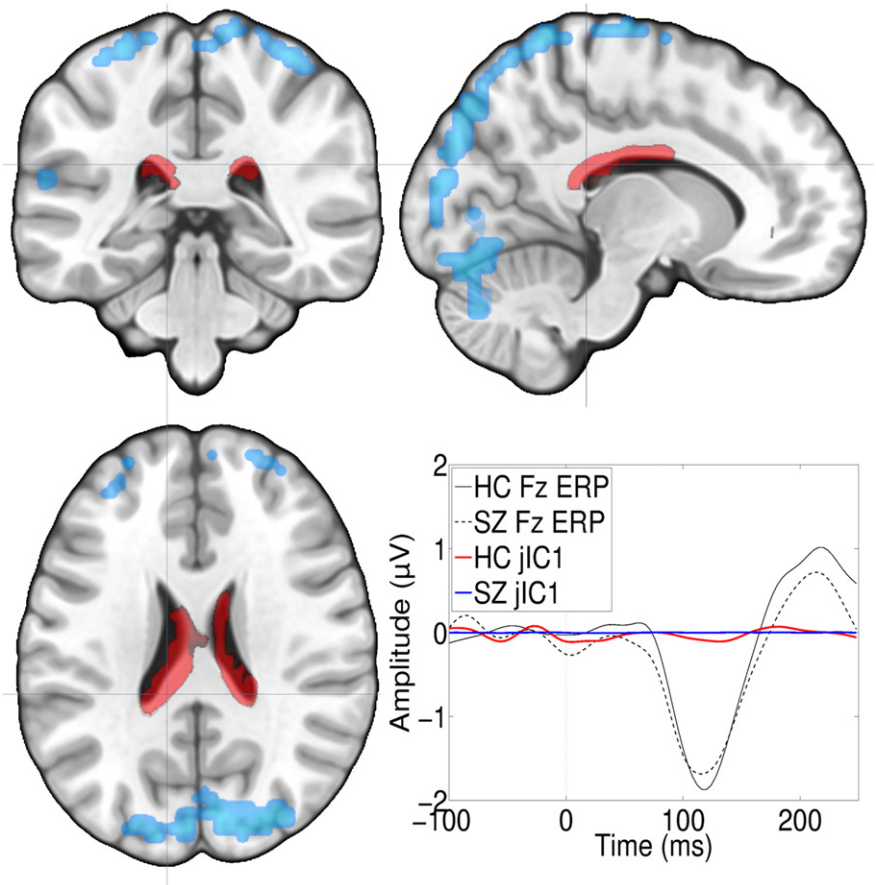

Fig. A1. The spatial loadings and time courses for joint independent component 1 (jIC1). The flat line blue waveform shows the ERP loading for the patients, and the red waveform shows the ERP loading for the controls.

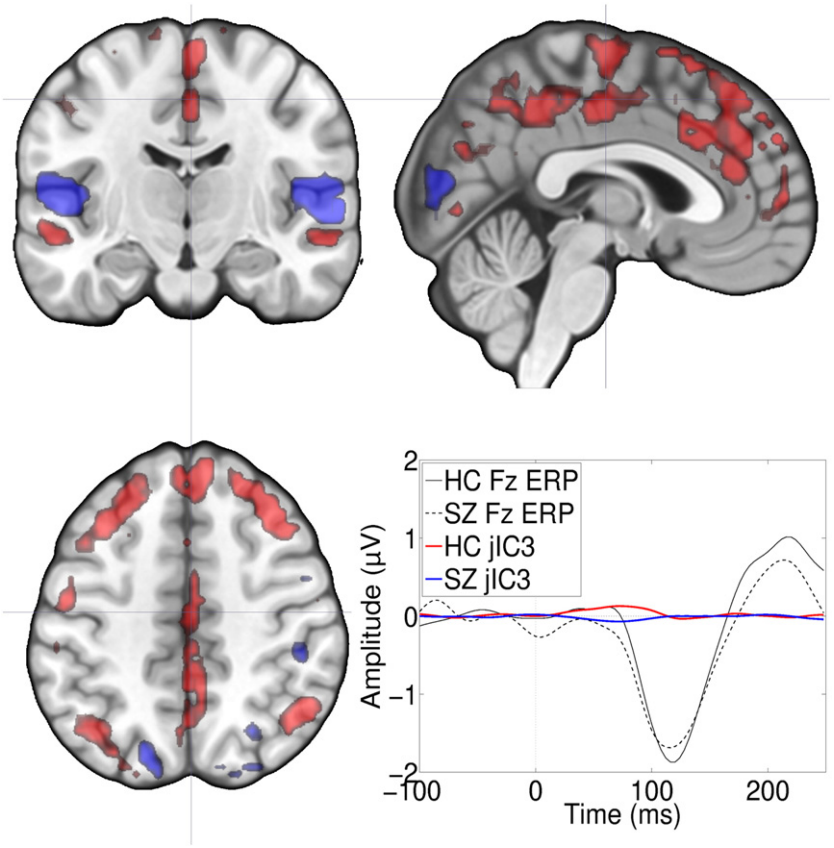

Fig. A2. The spatial loadings and time courses for joint independent component 3 (jIC3). The relatively flat line blue waveform shows the ERP loading for the schizophrenia patients (SZ), and the red waveform shows the ERP loading for the controls (HC), which is slightly positive in the 0–100 ms interval.

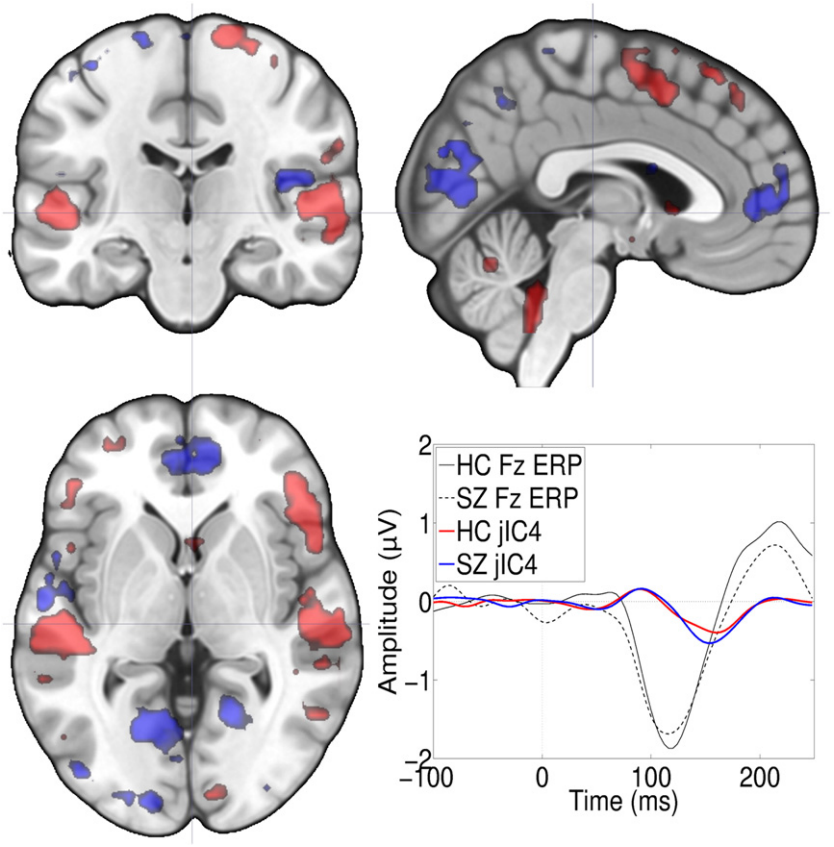

Fig. A3. The spatial loadings and time courses for joint independent component 4 (jIC4). The ERP loading waveforms for the schizophrenia patients (SZ, blue) and the controls (HC, red) track one another closely, with a positive peak around 92 milliseconds (ms) and negative peak around 156 ms.

Appendix B. Appendix references

Allen, P.J., Polizzi, G., Krakow, K., Fish, D.R., Lemieux, L., 1998. Identification of EEG events in the MR scanner: the problem of pulse artifact and a method for its subtraction. Neuroimage 8, 229–239.

De Clercq, W., Vergult, A., Vanrumste, B., Van Paesschen, W., Van Huffel, S., 2006. Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. IEEE Trans. Biomed. Eng. 53, 2583–2587.

Freeman, W.J., Holmes, M.D., Burke, B.C., Vanhatalo, S., 2003. Spatial spectra of scalp EEG and EMG from awake humans. Clin. Neurophysiol. 114, 1053–1068.

Pereda, E., Gamundi, A., Rial, R., Gonzalez, J., 1998. Non-linear behaviour of human EEG: fractal exponent versus correlation dimension in awake and sleep stages. Neurosci. Lett. 250, 91–94.

Ries, S., Janssen, N., Burle, B., Alario, F.X., 2013. Response-locked brain dynamics of word production. PLoS One. 8, e58197.

Schulz, M.A., Regenbogen, C., Moessnang, C., Neuner, I., Finkelmeyer, A., Habel, U., Kellermann, T., 2015. On utilizing uncertainty information in template-based EEG-fMRI ballistocardiogram artifact removal. Psychophysiology. 52, 857–863.

Shams, N., Alain, C., Strother, S., 2015. Comparison of BCG artifact removal methods for evoked responses in simultaneous EEG-fMRI. Journal of Neuroscience Methods. 245, 137–146.

References

Ahveninen, J., Jaaskelainen, I.P., Osipova, D., Huttunen, M.O., Ilmoniemi, R.J., Kaprio, J., Lonnqvist, J., Manninen, M., Pakarinen, S., Therman, S., Naatanen, R., Cannon, T.D., 2006. Inherited auditory-cortical dysfunction in twin pairs discordant for schizophrenia. Biol. Psychiatry 60, 612–620.

Allen, P.J., Josephs, O., Turner, R., 2000. A method for removing imaging artifact from continuous EEG recorded during functional MRI. NeuroImage 12, 230–239.

Andreasen, N.C., 1983. The Scale for the Assessment of Negative Symptoms (SANS). University of Iowa, Iowa City, IA.

Andreasen, N.C., 1984. Scale for the Assessment of Positive Symptoms. University of Iowa, Iowa City, IA.

Andreasen, N.C., O'Leary, D.S., Cizadlo, T., Arndt, S., Rezai, K., Watkins, L., Ponto, L.L.B., Hichwa, R.D., 1995. Remembering the past: two facets of episodic memory explored with positron emission tomography. Am. J. Psychiatr. 152, 1576–1585.

Behzadi, Y., Restom, K., Liau, J., Liu, T.T., 2007. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. NeuroImage 37, 90–101.

Binder, J.R., 2012. Task-induced deactivation and the "resting" state. NeuroImage 62, 1086–1091.

Brebion, G., Smith, M.J., Gorman, J.M., Amador, X., 1996. Reality monitoring failure in schizophrenia: the role of selective attention. Schizophr. Res. 22, 173–180.

Cahn, B.R., Polich, J., 2009. Meditation (Vipassana) and the P3a event-related brain potential. Int. J. Psychophysiol. 72, 51–60.

Calhoun, V.D., Adali, T., Kiehl, K.A., Astur, R., Pekar, J.J., Pearlson, G.D., 2006. A method for multitask fMRI data fusion applied to schizophrenia. Hum. Brain Mapp. 27, 598–610.

Chen, C.M., Mathalon, D.H., Roach, B.J., Cavus, I., Spencer, D.D., Ford, J.M., 2011. The corollary discharge in humans is related to synchronous neural oscillations. J. Cogn. Neurosci. 23, 2892–2904.

Crowley, K.E., Colrain, I.M., 2004. A review of the evidence for P2 being an independent component process: age, sleep and modality. Clin. Neurophysiol. 115, 732–744.

De Clercq, W., Vergult, A., Vanrumste, B., Van Paesschen, W., Van Huffel, S., 2006. Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram. IEEE Trans. Biomed. Eng. 53, 2583–2587.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.

Edwards, B.G., Calhoun, V.D., Kiehl, K.A., 2012. Joint ICA of ERP and fMRI during error-monitoring. NeuroImage 59, 1896–1903.

Ellamil, M., Fox, K.C., Dixon, M.L., Pritchard, S., Todd, R.M., Thompson, E., Christoff, K., 2016. Dynamics of neural recruitment surrounding the spontaneous arising of thoughts in experienced mindfulness practitioners. NeuroImage http://dx.doi.org/10.1016/j.neuroimage.2016.04.034.

Ethridge, L.E., Hamm, J.P., Pearlson, G.D., Tamminga, C.A., Sweeney, J.A., Keshavan, M.S., Clementz, B.A., 2015. Event-related potential and time-frequency endophenotypes for schizophrenia and psychotic bipolar disorder. Biol. Psychiatry 77, 127–136.

Ferreira-Santos, F., Silveira, C., Almeida, P.R., Palha, A., Barbosa, F., Marques-Teixeira, J., 2012. The auditory P200 is both increased and reduced in schizophrenia? A meta-analytic dissociation of the effect for standard and target stimuli in the oddball task. Clin. Neurophysiol. 123, 1300–1308.

First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 1995. Structured Clinical Interview for DSM-IV Axis I Disorders. Biometrics Research Department, New York State Psychiatric Institute, New York, NY.

Flinker, A., Chang, E.F., Kirsch, H.E., Barbaro, N.M., Crone, N.E., Knight, R.T., 2010. Single-trial speech suppression of auditory cortex activity in humans. J. Neurosci. 30, 16643–16650.

Ford, J.M., Roth, W.T., Kopell, B.S., 1976. Attention effects on auditory evoked potentials to infrequent events. Biol. Psychol. 4, 65–77.

Ford, J.M., Roth, W.T., Menon, V., Pfefferbaum, A., 1999. Failures of automatic and strategic processing in schizophrenia: comparisons of event-related brain potential and startle blink modification. Schizophr. Res. 37, 149–163.

Ford, J.M., Roach, B.J., Jorgensen, K.W., Turner, J.A., Brown, G.G., Notestine, R., Bischoff-Grethe, A., Greve, D., Wible, C., Lauriello, J., Belger, A., Mueller, B.A., Calhoun, V., Preda, A., Keator, D., O'Leary, D.S., Lim, K.O., Glover, G., Potkin, S.G., Mathalon, D.H., 2009. Tuning in to the voices: a multisite FMRI study of auditory hallucinations. Schizophr. Bull. 35, 58–66.

Ford, J.M., Palzes, V.A., Roach, B.J., Mathalon, D.H., 2014. Did I do that? Abnormal predictive processes in schizophrenia when button pressing to deliver a tone. Schizophr. Bull. 40, 804–812.

Fransson, P., 2005. Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. Hum. Brain Mapp. 26, 15–29.

Freeman, W.J., Holmes, M.D., Burke, B.C., Vanhatalo, S., 2003. Spatial spectra of scalp EEG and EMG from awake humans. Clin. Neurophysiol. 114, 1053–1068.

Friston, K., 2010. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11, 127–138.

Garcia-Larrea, L., Lukaszewicz, A.C., Mauguiere, F., 1992. Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. Neuropsychologia 30, 723–741.

Green, M.F., Hellemann, G., Horan, W.P., Lee, J., Wynn, J.K., 2012. From perception to functional outcome in schizophrenia: modeling the role of ability and motivation. Arch. Gen. Psychiatry 69, 1216–1224.

Greicius, M.D., Krasnow, B., Reiss, A.L., Menon, V., 2003. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc. Natl. Acad. Sci. U. S. A. 100, 253–258.

Hari, R., Pelizzone, M., Makela, J.P., Hallstrom, J., Leinonen, L., Lounasmaa, O.V., 1987. Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. Audiology 26, 31–43.

Hillyard, S.A., Squires, K.C., Bauer, J.W., Lindsay, P.H., 1971. Evoked potential correlates of auditory signal detection. Science 172, 1357–1360.

Hillyard, S.A., Hink, R.F., Schwent, V.L., Picton, T.W., 1973. Electrical signs of selective attention in the human brain. Science 182, 177–180.

Hubl, D., Koenig, T., Strik, W.K., Garcia, L.M., Dierks, T., 2007. Competition for neuronal resources: how hallucinations make themselves heard. Br. J. Psychiatry 190, 57–62.

Joos, K., Gilles, A., Van de Heyning, P., De Ridder, D., Vanneste, S., 2014. From sensation to percept: the neural signature of auditory event-related potentials. Neurosci. Biobehav. Rev. 42, 148–156.

Kirchner, H., Barbeau, E.J., Thorpe, S.J., Regis, J., Liegeois-Chauvel, C., 2009. Ultra-rapid sensory responses in the human frontal eye field region. J. Neurosci. 29, 7599–7606.

Krumbholz, K., Patterson, R.D., Seither-Preisler, A., Lammertmann, C., Lutkenhoner, B., 2003. Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. Cereb. Cortex 13, 765–772.

Liebenthal, E., Sabri, M., Beardsley, S.A., Mangalathu-Arumana, J., Desai, A., 2013. Neural dynamics of phonological processing in the dorsal auditory stream. J. Neurosci. 33, 15414–15424.

Mayer, A.R., Ruhl, D., Merideth, F., Ling, J., Hanlon, F.M., Bustillo, J., Canive, J., 2013. Functional imaging of the hemodynamic sensory gating response in schizophrenia. Hum. Brain Mapp. 34, 2302–2312.

Naatanen, R., Michie, P.T., 1979. Early selective-attention effects on the evoked potential: a critical review and reinterpretation. Biol. Psychol. 8, 81–136.

Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24, 375–425.

Nolan, H., Whelan, R., Reilly, R.B., 2010. FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection. J. Neurosci. Methods 192, 152–162.

Novak, G., Ritter, W., Vaughan Jr., H.G., 1992. Mismatch detection and the latency of temporal judgements. Psychophysiology 29, 398–411.

Oades, R., Dittmann-Balcar, A., Zerbin, D., 1997. Development and topography of auditory event-related potentials (ERPs): mismatch and processing negativity in individuals 8–22 years of age. Psychophysiology 34, 677–693.

Ozaki, I., Suzuki, Y., Jin, C.Y., Baba, M., Matsunaga, M., Hashimoto, I., 2003. Dynamic movement of N100m dipoles in evoked magnetic field reflects sequential activation of isofrequency bands in human auditory cortex. Clin. Neurophysiol. 114, 1681–1688.

Paiva, T.O., Almeida, P.R., Ferreira-Santos, F., Vieira, J.B., Silveira, C., Chaves, P.L., Barbosa, F., Marques-Teixeira, J., 2016. Similar sound intensity dependence of the N1 and P2 components of the auditory ERP: averaged and single trial evidence. Clin. Neurophysiol. 127, 499–508.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., Elbert, T., 1995. Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. Electroencephalogr. Clin. Neurophysiol. 94, 26–40.

Pantev, C., Eulitz, C., Hampson, S., Ross, B., Roberts, L.E., 1996. The auditory evoked "off" response: sources and comparison with the "on" and the "sustained" responses. Ear Hear. 17, 255–265.

Parmet, Y., Schechtman, E., Sherman, M., 2010. Factor analysis revisited—how many factors are there? Communications in Statistics - Simulation and Computation 39, 1893–1908.
Pereda, E., Gamundi, A., Rial, R., Gonzalez, J., 1998. Non-linear behaviour of human EEG: fractal exponent versus correlation dimension in awake and sleep stages. Neurosci. Lett. 250, 91–94.
Potts, G.F., Hirayasu, Y., O'Donnell, B.F., Shenton, M.E., McCarley, R.W., 1998. High-density recording and topographic analysis of the auditory oddball event-related potential in patients with schizophrenia. Biol. Psychiatry 44, 982–989.
Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. Proc. Natl. Acad. Sci. U. S. A. 98, 676–682.
Reite, M., Adams, M., Simon, J., Teale, P., Sheeder, J., Richardson, D., Grabbe, R., 1994. Auditory M100 component 1: relationship to Heschl's gyri. Brain Res. Cogn. Brain Res. 2, 13–20.
Ries, S., Janssen, N., Burle, B., Alario, F.X., 2013. Response-locked brain dynamics of word production. PLoS One 8, e58197.
Rissanen, J., 1978. Paper: modeling by shortest data description. Automatica 14, 465–471.
Rosburg, T., Boutros, N.N., Ford, J.M., 2008. Reduced auditory evoked potential component N100 in schizophrenia - a critical review. Psychiatry Res. 161, 259–274.
Roth, W.T., Horvath, T.B., Pfefferbaum, A., Kopell, B.S., 1980. Event related potentials in schizophrenics. Electroencephalogr. Clin. Neurophysiol. 48, l27–l39.
Roth, W.T., Goodale, J., Pfefferbaum, A., 1991. Auditory event-related potentials and electrodermal activity in medicated and unmedicated schizophrenics. Biol. Psychiatry 29, 585–599.
Salisbury, D.F., Collins, K.C., McCarley, R.W., 2010. Reductions in the N1 and P2 auditory event-related potentials in first-hospitalized and chronic schizophrenia. Schizophr. Bull. 36, 991–1000.
Sams, M., Hamalainen, M., Antervo, A., Kaukoranta, E., Reinikainen, K., Hari, R., 1985. Cerebral neuromagnetic responses evoked by short auditory stimuli. Electroencephalogr. Clin. Neurophysiol. 61, 254–266.
Schacter, D.L., Gilbert, D.T., Wegner, D.M., 2015. Introducing Psychology, Second Edition, Special Update for DSM-5. Worth Publishers, New York, NY.
Schroger, E., Marzecova, A., SanMiguel, I., 2015. Attention and prediction in human audition: a lesson from cognitive psychophysiology. Eur. J. Neurosci. 41, 641–664.
Vaughan Jr., H.G., Ritter, W., Simson, R., 1980. Topographic analysis of auditory event-related potentials. Prog. Brain Res. 54, 279–285.
Verkindt, C., Bertrand, O., Thevenet, M., Pernier, J., 1994. Two auditory components in the 130–230 ms range disclosed by their stimulus frequency dependence. Neuroreport 5, 1189–1192.
Verkindt, C., Bertrand, O., Perrin, F., Echallier, J.-F., Pernier, J., 1995. Tonotopic organization of the human auditory cortex: N100 topography and multiple dipole model analysis. Electroencephalogr. Clin. Neurophysiol. 96, 143–156.
Wang, J., Mathalon, D.H., Roach, B.J., Reilly, J., Keedy, S.K., Sweeney, J.A., Ford, J.M., 2014. Action planning and predictive coding when speaking. NeuroImage 91, 91–98.
Whitfield-Gabrieli, S., Thermenos, H.W., Milanovic, S., Tsuang, M.T., Faraone, S.V., McCarley, R.W., Shenton, M.E., Green, A.I., Nieto-Castanon, A., LaViolette, P., Wojcik, J., Gabrieli, J.D., Seidman, L.J., 2009. Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. Proc. Natl. Acad. Sci. U. S. A. 106, 1279–1284.
Woodman, G.F., 2010. A brief introduction to the use of event-related potentials in studies of perception and attention. Atten. Percept. Psychophys. 72, 2031–2046.
Woodruff, P., Brammer, M., Mellers, J., Wright, I., Bullmore, E., Williams, S., 1995. Auditory hallucinations and perception of external speech. Lancet 346, 1035.
Wynn, J.K., Jimenez, A.M., Roach, B.J., Korb, A., Lee, J., Horan, W.P., Ford, J.M., Green, M.F., 2015. Impaired target detection in schizophrenia and the ventral attentional network: findings from a joint event-related potential-functional MRI analysis. Neuroimage Clin. 9, 95–102.
Zouridakis, G., Simos, P.G., Papanicolaou, A.C., 1998. Multiple bilaterally asymmetric cortical sources account for the auditory N1m component. Brain Topogr. 10, 183–189.