

# The Neural Substrates of Impaired Prosodic Detection in Schizophrenia and Its Sensorial Antecedents

David I. Leitman, Ph.D.

Matthew J. Hoptman, Ph.D.

John J. Foxe, Ph.D.

Erica Saccente, B.S.

Glenn R. Wylie, Ph.D.

Jay Nierenberg, M.D.

Maria Jalbrzikowski, B.A.

Kelvin O. Lim, M.D.

Daniel C. Javitt, M.D., Ph.D.

**Objective:** Individuals with schizophrenia show severe deficits in their ability to decode emotions based upon vocal inflection (affective prosody). This study examined neural substrates of prosodic dysfunction in schizophrenia with voxelwise analysis of diffusion tensor magnetic resonance imaging (MRI).

**Method:** Affective prosodic performance was assessed in 19 patients with schizophrenia and 19 comparison subjects with the Voice Emotion Identification Task (VOICEID), along with measures of basic pitch perception and executive processing (Wisconsin Card Sorting Test). Diffusion tensor MRI fractional anisotropy values were used for voxelwise correlation analyses. In a follow-up experiment, performance on a nonaffective prosodic perception task was assessed in an additional cohort of 24 patients and 17 comparison subjects.

**Results:** Patients showed significant deficits in VOICEID and Distorted Tunes Task performance. Impaired VOICEID performance correlated significantly with lower

fractional anisotropy values within primary and secondary auditory pathways, orbitofrontal cortex, corpus callosum, and peri-amygdala white matter. Impaired Distorted Tunes Task performance also correlated with lower fractional anisotropy in auditory and amygdalar pathways but not prefrontal cortex. Wisconsin Card Sorting Test performance in schizophrenia correlated primarily with prefrontal fractional anisotropy. In the follow-up study, significant deficits were observed as well in nonaffective prosodic performance, along with significant intercorrelations among sensory, affective prosodic, and nonaffective measures.

**Conclusions:** Schizophrenia is associated with both structural and functional disturbances at the level of primary auditory cortex. Such deficits contribute significantly to patients' inability to decode both emotional and semantic aspects of speech, highlighting the importance of sensorial abnormalities in social communicatory dysfunction in schizophrenia.

(*Am J Psychiatry* 2007; 164:474–482)

Schizophrenia is associated with deficits in the ability to decode emotion based upon modulation of intonation (affective prosody) (1, 2). Such deficits, along with disturbances in facial affect recognition, contribute substantially to impaired social and global outcome in schizophrenia (3, 4). Traditionally, such deficits have been attributed to generalized neurocognitive dysfunction, particularly involving processes such as executive function and working memory (5), as well as limbic dysfunction (6). More recently, however, specific contributions of auditory processing deficits have been noted as well, with deficits in simple tone matching correlating with deficits in prosodic identification (4). These findings predict both sensory-level and cognitive-level contributions to impaired prosodic processing. This study investigates the neural substrates of impaired auditory emotion detection with a combined behavioral and structural imaging approach.

Neuroanatomical abnormalities in schizophrenia have been extensively documented and shown to involve both white and gray matter structures (e.g., references 7, 8). In order to further evaluate the neural substrates of im-

paired prosodic processing in schizophrenia, voxelwise correlation analyses were performed relative to magnetic resonance diffusion tensor imaging studies. Diffusion tensor imaging is sensitive to white matter disturbances in schizophrenia and has previously been used to evaluate structure-function correlations within frontal cortical regions (7–9). To our knowledge, this is the first study to evaluate structure-function relationships within auditory sensory brain regions in schizophrenia with diffusion tensor imaging.

Diffusion tensor imaging studies are typically analyzed with fractional anisotropy, a parameter that reflects the relative diffusion of water parallel to the long axis of structural boundaries, such as axonal membranes or myelin, relative to diffusion perpendicular to those boundaries (9). Reduced fractional anisotropy in schizophrenia is thought to reflect either axonal or myelin-related pathology (9), both of which have been documented in schizophrenia (10). Reduction in fractional anisotropy may also reflect disorganization of fiber bundles (“disconnectivity”), although in areas of crossing fibers, such disconnectivity

This article is featured in this month's AJP **Audio** and is discussed in an editorial by Dr. Freedman on p. 385.

tivity may lead to paradoxical increases in fractional anisotropy (11).

Regardless of underlying etiology, fractional anisotropy has proved effective for evaluating structure-function relationships. For example, increased impulsivity has been found to correlate selectively with reduced fractional anisotropy in inferior frontal white matter (12, 13), whereas impairments in executive processing have been found to correlate selectively with reduced fractional anisotropy in anterior prefrontal regions (14, 15). Impairments in visual processing have been found to correlate selectively with reduced fractional anisotropy in optic radiations (16).

In this study, we analyzed relationships between regional fractional anisotropy in schizophrenia patients and healthy comparison subjects and performance on two separate tasks: the Distorted Tunes Task (17), which measures the ability to detect incorrect notes within common melodies, and the Voice Emotion Identification Task (VOICEID) (2), which measures the ability to decode emotions based upon tone of voice. The Distorted Tunes Task was originally developed to assess genetic contributions to musical pitch perception abilities and shows high heritability within families (17). The VOICEID task has been used by ourselves (4) and others (2, 3), and it is highly sensitive to affective prosodic perception deficits in schizophrenia.

In the brain, auditory projection paths begin at the level of the medial geniculate nucleus and project to the primary auditory cortex (Heschl's gyrus, A1) through superolaterally projecting thalamocortical (acoustic) radiations. From the auditory cortex, fibers project to higher brain regions along both ventral and dorsal divisions of the arcuate fasciculus (18). The ventral stream is primarily involved in acoustic feature analysis (19), whereas the dorsal stream is thought to process spatial and spectral motion (20), including speech (19).

Affective prosodic comprehension can be conceptualized as involving a "three-stage processing chain" that begins with sensation (stage 1) in the primary auditory cortex and continues with integration (stage 2) within ventral aspects of the temporal cortex and the superior temporal sulcus. There, aspects of the acoustical information are tagged as affective. Processing proceeds finally to the cognitive stage (stage 3) in inferior frontal regions, where this information is evaluated both semantically and contextually (reviewed in reference 21).

The present study used two specific tasks—the Distorted Tunes Task and the VOICEID—to evaluate functioning of sensation and integration phases of affective prosodic comprehension. We have previously demonstrated that individuals with schizophrenia show deficits in auditory sensory-level performance, as reflected in impaired tone-matching ability (22), as well as reduced auditory event-related potential generation (23). For this study, we hypothesized that joint impairments in Distorted Tunes Task and VOICEID performance in patients would

correlate significantly with reduced fractional anisotropy primarily in basic auditory brain regions (i.e., acoustic radiations) that subserve sensation and that VOICEID impairments would show additional correlations in ventral and dorsal stream projection regions that subserve integration and cognitive evaluation.

As a control condition, we evaluated correlations between fractional anisotropy and performance levels on the Wisconsin Card Sorting Test, a widely used, visually based test of executive/prefrontal performance that would not be expected to show correlations with auditory sensory regions. In a prior study, an increased perseverative error rate on the Wisconsin Card Sorting Test was found to correlate with reduced fractional anisotropy in specific regions of the cingulum (14). Here, patterns of voxelwise correlations to the Wisconsin Card Sorting Test were compared to patterns observed for the Distorted Tunes Task/VOICEID.

To characterize the further extent of prosodic processing dysfunction in schizophrenia, a follow-up study examined the integrity of nonaffective prosodic processing, such as the ability to distinguish statements from questions (declarative versus interrogative prosody) or to distinguish between alternative meanings based upon which syllables within a sentence are emphasized (stress prosody). We hypothesized that impairments in pitch perception, which were statistically associated with impaired affective prosodic performance (i.e., emotion recognition deficits) in a prior study (4), might also be correlated with impairments in the ability to make even nonaffective discriminations, such as the ability to differentiate statements and questions (i.e., declarative versus interrogative intent), based upon tone of voice. Furthermore, such findings would provide convergent evidence for the contribution of sensorial abnormalities to social communicatory disturbances in schizophrenia.

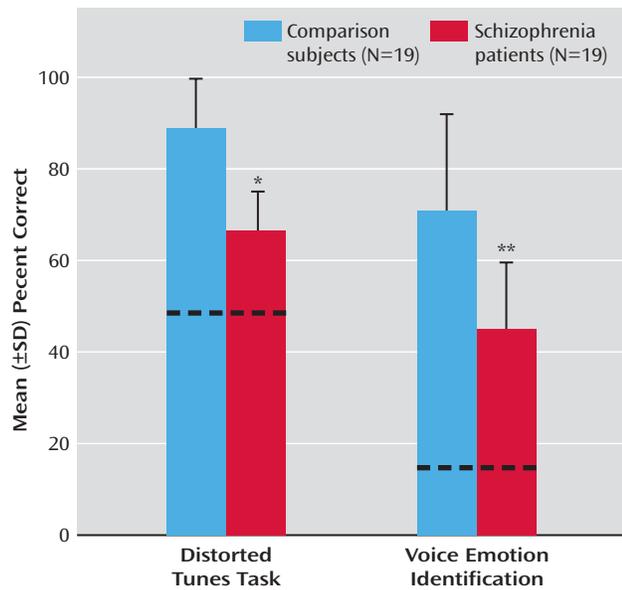
## Methods

### Experiment 1

**Participants.** Nineteen patients (one woman) meeting DSM-IV criteria for either schizophrenia (N=17) or schizoaffective disorder (N=2) participated in this study. Diagnoses were based upon the Structured Clinical Interview for DSM-IV (SCID), with all available clinical information used. Twelve patients were receiving only second-generation antipsychotics (primarily risperidone or olanzapine), two patients were receiving clozapine, two patients were receiving only traditional antipsychotics (haloperidol), and three patients were receiving combination treatment. The mean chlorpromazine equivalency dose was 1298.3 mg/day (SD=780.4). The mean illness duration was 15.7 years (SD=8.7). The clinical ratings of patients followed the methods described previously (4). The patients had a mean rating of 35.5 (SD=7.1) on the Brief Psychiatric Rating Scale and a mean rating of 32.5 (SD=12.8) on the Schedule for the Assessment of Negative Symptoms.

The healthy comparison group consisted of 19 (six women) staff volunteers or individuals who responded to local advertisements. The comparison subjects had a mean age of 36 years (SD=9). The healthy subjects and the patients differed significantly in mean IQ (112, SD=9, versus 95, SD=16, respectively;  $t=3.9$ ,  $df=32$ ,  $p<0.001$ )

**FIGURE 1. Behavioral Task Performance of Comparison Subjects and Schizophrenia Patients<sup>a</sup>**



<sup>a</sup> Effect sizes (d) for Distorted Tunes Task and VOICEID were SD=1.2 and SD=1.6, respectively. Dashed lines represent chance performance levels. \*p<0.01. \*\*p<0.001.

and grades achieved (16 years, SD=2, versus 11, SD=3; t=6.05, df=36, p<0.001).

The local institutional review boards approved all experimental procedures, and all subjects provided written informed consent after study procedures were explained fully. The participants received \$10/hour for participation.

All subjects except one patient were right-handed, as assessed by methods described previously (4). Within this group, 12 of 19 patients and seven of 19 healthy comparison subjects had been in a prior study (4).

**Behavioral measures.** Behavioral measures included the VOICEID (2), which consists of 21 spoken sentences each conveying one of six different emotions (happiness, anger, fear, sadness, surprise, or shame), the Distorted Tunes Task (17), which consists of 26 popular tunes of which 17 are rendered melodically incorrect by changing the pitch of specific notes within the tune, and the Wisconsin Card Sorting Test (24). For the VOICEID and the Distorted Tunes Task, the primary dependent measure was percentage of stimuli identified correctly. For the Wisconsin Card Sorting Test, which was administered to patients only, the perseverative error rate was used as a primary dependent measure. All behavioral tasks were presented through a stereo player at a comfortable hearing level. For all subjects, behavioral task data were collected within 6 months of magnetic resonance imaging (MRI) acquisition.

**MRI.** Scanning was performed on a 1.5-T Siemens Vision system (Erlangen, Germany) at the Nathan Kline Institute Center for Advanced Brain Imaging. Three main sequences were acquired: a magnetization prepared rapidly acquired gradient echo scan (TR/TE=11.4/4.9 msec, matrix=256×256, field of view=300 mm, number of excitations=1, 1.17-mm slice thickness, 172 slices, no gap), a turbo spin echo scan (TR=5000 msec, TE=22/90 msec, matrix=256×256, field of view=224 mm, number of excitations=1, 5-mm slice thickness, 26 slices, no gap), and a diffusion tensor imaging sequence. The diffusion tensor imaging sequence has been

described elsewhere (25) (TR/TE=6000/100 msec, matrix=128×128 [interpolated to 256×256], field of view=240 mm, 5-mm slice thickness, 20 slices, no gap) and employs a double echo pulse to minimize eddy current effects (26). The sequence entailed four acquisitions of six diffusion-weighted images (b=1000 s/mm<sup>2</sup>) for 20 slices. In addition, two acquisitions without diffusion weighting (b=0 s/mm<sup>2</sup>) were acquired.

Fractional anisotropy was calculated with custom software. The b=0 images were corrected for susceptibility-induced distortion and were transformed into Montreal Neurological Institute space with methods described elsewhere (13, 27). Images were matched to a template in Montreal Neurological Institute space, and the final voxel size was 2×2×2 mm<sup>3</sup>. A white matter mask was computed from the mean spatially normalized patient fractional anisotropy image with a nonparametric image segmentation algorithm (28) and was applied to all of the standardized images. This approach limited the voxels to white matter and resulted in fewer statistical comparisons, thereby lowering the probability of false positive tests.

After transformation into Talairach space, the images were masked such that only voxels with data present for all subjects were included in the analyses. This ensured that missing data, which would have zero values, would not drive correlations.

**Statistical analysis.** Between-group comparisons of prosodic (VOICEID) and pitch (Distorted Tunes Task) performance were performed with repeated-measures analysis of variance. Spearman correlation coefficients were used to measure the relationships between task performances within groups.

For neuroimaging data, a voxelwise correlation approach was used similar to that of Baudewig et al. (29), with thresholds as described previously (13, 15). This approach protects against false positive correlations with voxels that are significant at p≤0.05 that are grown from a seed voxel with a significance value of p<0.005. To supplement these criteria, only clusters with more than 11 contiguous voxels were considered significant. To assess areas of shared correlation across tasks, maps of each task-fractional anisotropy correlation cluster were overlaid, and a new map representing overlap regions (identically thresholded for each task) was generated.

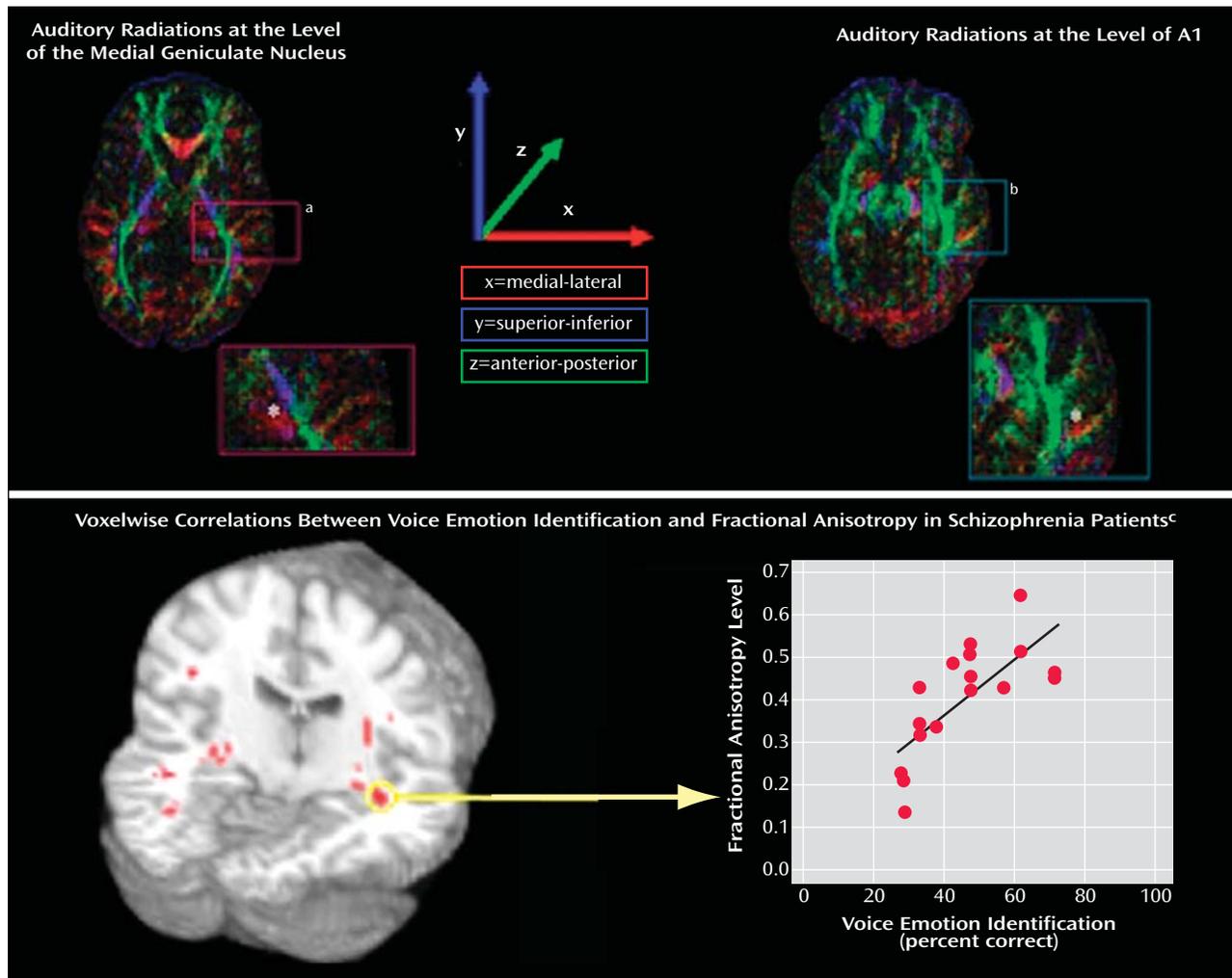
Our voxelwise correlation analysis was two-tailed; nevertheless, we focused a priori only on correlations in which worse performance correlated with fractional anisotropy reductions.

**Experiment 2**

**Participants.** The participants consisted of 24 patients with schizophrenia (three women) meeting DSM-IV criteria for either schizophrenia (N=21) or schizoaffective disorder (N=3) and 17 healthy volunteers (three women). Healthy subjects and patients were of similar age (mean=37.8 years, SD=10.2, versus mean=32.5, SD=10.6, respectively) but differed in verbal IQ (mean=109.7, SD=10.4, versus mean=94.1, SD=7.5; t=4.4, df=33, p>0.001) and education (mean=16, SD=1, versus mean=11, SD=2; t=-7.5, df=35, p>0.001). The patients were receiving typical and/or atypical antipsychotic medication (chlorpromazine dose: mean=1373 mg/day, SD=829).

**Behavioral measures.** In addition to the VOICEID (2), the Distorted Tunes Task (17), and the Wisconsin Card Sorting Test (24), nonaffective prosody was assessed with Weintraub's Sentence Discrimination and Semantic Comprehension Tasks (30): Twenty-five pairs of semantically neutral sentences, such as "Jack climbed the mountain," were repeated after a brief delay. Seventeen of the pairs differed because of either stress (where stressed emphasis shifted between the subject and object of the sentence) or declarative/interrogative differences. Eight pairs were identical. The subjects were asked whether the sentences were said in

**FIGURE 2.** Fractional Anisotropy Map at the Level of the Medial Geniculate Nucleus and the Primary Auditory Cortex (Heschl's Gyrus) (top left), Fiber Pathways at the Level of the Ventral Stream (top right), and Three-Dimensional Voxelwise Correlation Map for VOICEID Performance in Schizophrenia Patients (bottom)



<sup>a</sup> Magnified view showing acoustic radiations that project superolaterally from the medial geniculate nucleus to Heschl's gyrus (right).

<sup>b</sup> Magnified view showing medial lateral (red) radiations to Broca's area branching off the superior arcuate fibers (green).

<sup>c</sup> Within-patient scatterplot and correlation values between fractional anisotropy levels from auditory radiations and VOICEID performance.

the same or a different manner. The score reflected percent correct. Additionally, scores were broken down into percent correct for stress and declarative/interrogative distinctions.

The Semantic Comprehension Task consisted solely of 16 utterances, expressing either declarative (eight utterances) or interrogative (eight utterances) intent. The subjects were asked whether the speaker posed a question or a statement. The score reflected percent correct.

**Data analysis.** Between-groups effects across all auditory measures were assessed with multivariate analysis of variance, with post hoc contrasts for specific measures (t tests). In the event of ceiling or floor effects, Mann-Whitney nonparametric measures were employed. Nonparametric signal detection analyses used  $A'$  (31) and  $B''$  (32) as measures of sensitivity and bias, respectively.

Correlation matrices between nonaffective and affective prosody, as well as pitch perception, were calculated within the patient group only and submitted to principal components analysis. Principal components analysis, factor selection, and rotation were conducted on eigenvalues  $\geq 1$  (see reference 4). All statistical

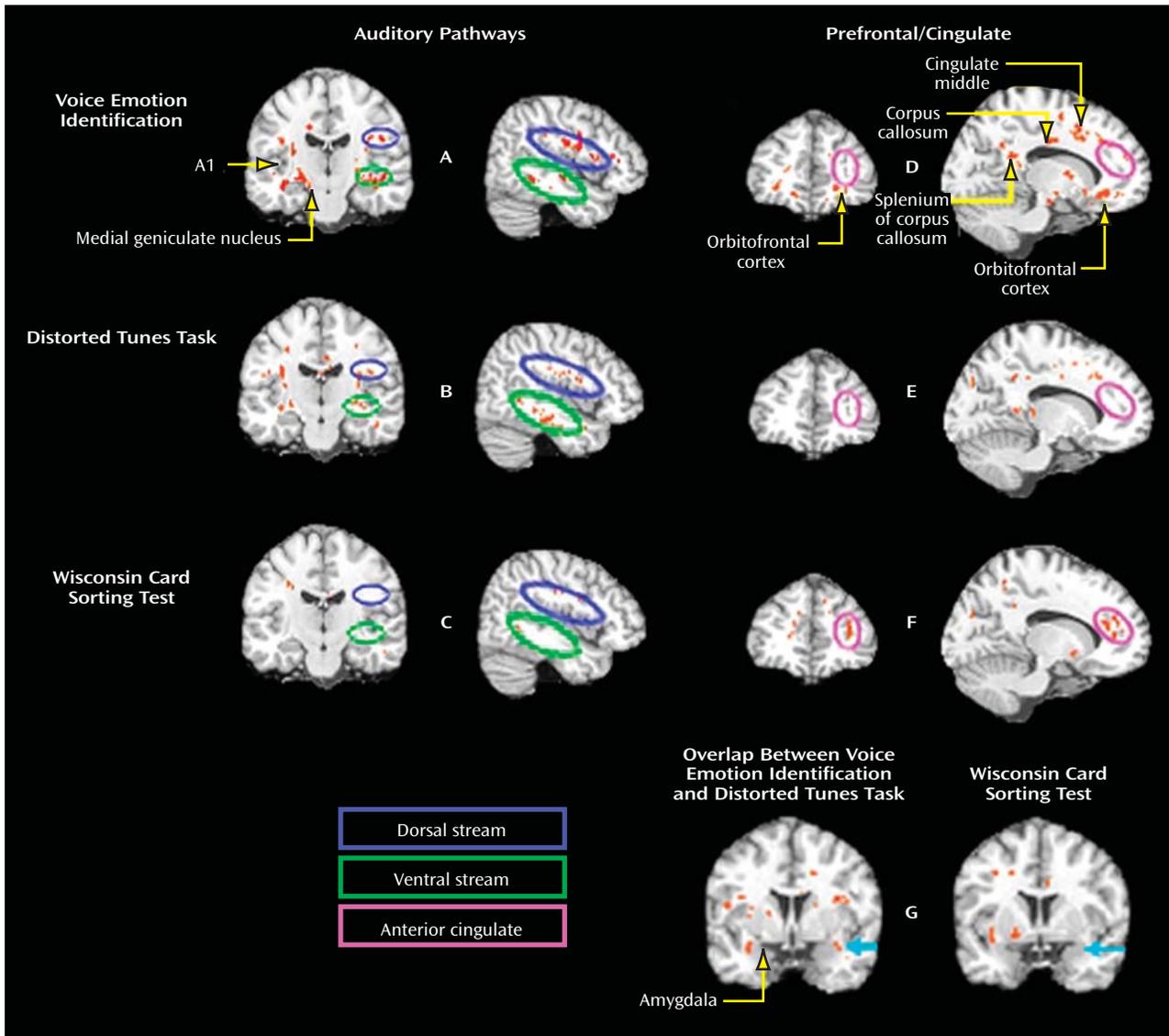
tests were two-tailed, with  $\alpha \leq 0.05$ , and computed with JMP software (SAS Institute, Cary, N.C.).

## Results

### Experiment 1

**Behavioral results.** The patients showed significantly impaired performance across both tests ( $F=56.6$ ,  $df=1$ ,  $36$ ,  $p<0.001$ ) (Figure 1). The group-by-task interaction was nonsignificant ( $F=2.5$ ,  $df=1$ ,  $36$ ,  $p<0.13$ ). Distorted Tunes Task and VOICEID scores were significantly correlated both across groups ( $r_s=0.55$ ,  $N=38$ ,  $p<0.001$ ) and within the patient group alone ( $r_s=0.54$ ,  $N=19$ ,  $p<0.02$ ) but were not significantly correlated for comparison subjects ( $r_s=0.25$ ,  $N=19$ ,  $p<0.30$ ). For patients, poorer Wisconsin Card Sorting Test performance (increased perseverative errors) correlated with poorer VOICEID ( $r_s=-0.55$ ,

FIGURE 3. Voxelwise Correlation Maps in Schizophrenia Patients for Dorsal and Ventral Auditory Pathways for Voice Emotion Identification (A), Distorted Tunes Task (B), and Wisconsin Card Sorting Test (C) (D–F) and Voxelwise Correlation Maps for Cingulate Fasciculus (D–F); and Voxelwise Correlation Maps in Schizophrenia Patients for Overlap Between Voice Emotion Identification and Distorted Tunes Task and Wisconsin Card Sorting Test at the Level of the Amygdala (G)<sup>a</sup>



<sup>a</sup> Arrows indicate periamygdala correlations.

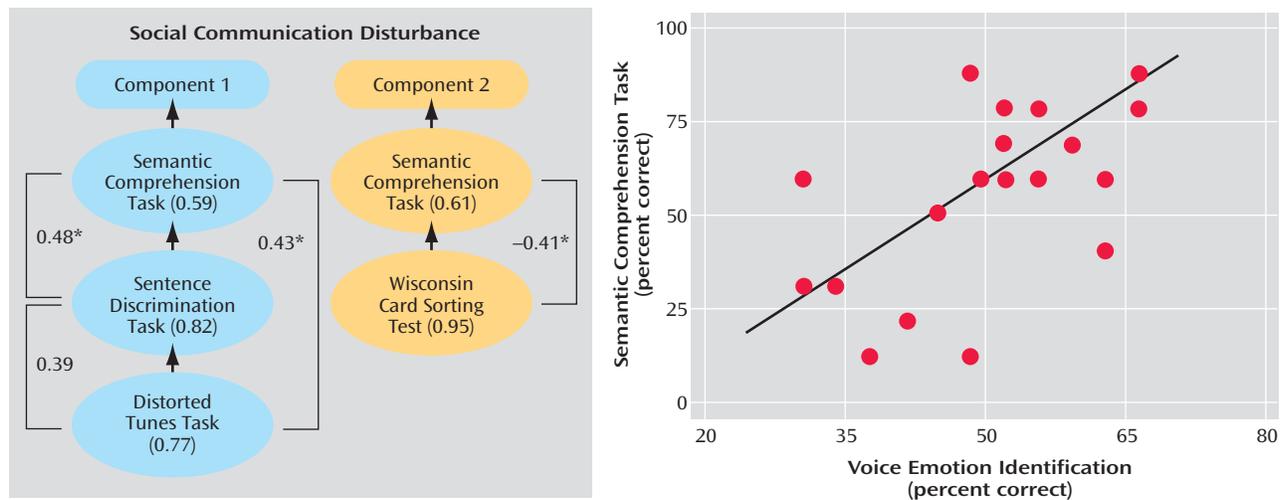
N=17,  $p < 0.02$ ) but not Distorted Tunes Task ( $r_s = -0.23$ , N=17,  $p < 0.34$ ) performance. Additionally, among patients, medication dosage did not correlate with neuropsychological measures (all  $p > 0.22$ ).

**Structural correlations within patients.** *VOICEID*. As predicted, impaired *VOICEID* performance was significantly correlated with fractional anisotropy in regions lying between the auditory thalamus (medial geniculate nucleus) and the primary auditory cortex (acoustic radiations, Figure 2, top). These regions are known to contain auditory radiations from Montreal Neurological Institute space to A1 (Figure 2, bottom). In addition to these regions, significant correlation clusters were observed bilat-

erally along the ventral and dorsal auditory pathway in temporal and frontal white matter (Figure 3A). Other areas in which correlation clusters were observed include the corpus callosum splenium and body, as well as the posterior commissure and the right cingulum (Figure 3D). Clusters were also observed adjacent to both left and right amygdala medial laterally (Figure 3, left). Additional areas of significant correlation included white matter in Brodmann's regions 44, 45, and 46 and the orbitofrontal cortex (Figure 3).  $R^2$  values for each region are shown in Table 1.

*Distorted Tunes Task*. Also as predicted, the pattern of correlations of fractional anisotropy with the Distorted Tunes Task closely resembled the pattern of correlations

**FIGURE 4. Principal Components Analysis on Correlations Between Pitch Perception and Nonaffective Prosody Performance in Schizophrenia Patients<sup>a</sup>**



<sup>a</sup> Schematic diagram of interrelationships between nonaffective prosody measures, pitch perception, and executive processing (left). Values inside circles represent rotated factor scores, while values outside circles represent Pearson correlation coefficients between indicated measures. The Sentence Discrimination and Distorted Tunes Tasks rely exclusively on component 1 (sensory processing), the Wisconsin Card Sorting Test relies exclusively on component 2 (executive processing), and the Semantic Comprehension Task relies on both components. On the right, the Semantic Comprehension Task by affective prosody (VOICEID).

\* $p < 0.05$ .

with VOICEID (Figure 3B) (see Table 1). Regions of overlap included primary auditory radiations, dorsal and ventral stream auditory projections (Figure 3), and the amygdala (Figure 4, left). However, no correlations were observed in regions 44, 45, or 46 or in the orbitofrontal cortex.

**Wisconsin Card Sorting Test.** As opposed to the Distorted Tunes Task and the VOICEID, no significant correlations of fractional anisotropy with the Wisconsin Card Sorting Test were observed in regions of the acoustic radiations or along either dorsal or ventral auditory radiations (Figure 3C) (see Table 1). Moreover, there were no significant areas of overlap between Wisconsin Card Sorting Test and VOICEID correlations (Figure 3A, Figure 3C). Significant correlation clusters were observed between the Wisconsin Card Sorting Test perseverative error scores and fractional anisotropy in white matter in the regions of the right anterior cingulate gyrus (Figure 3F). Even in frontal regions, however, little overlap was observed between correlation clusters for the VOICEID and the Wisconsin Card Sorting Test. Finally, in contrast to the VOICEID and the Distorted Tunes Task, no correlations in the vicinity of the amygdala were observed (Figure 3G).

## Experiment 2

The patients performed significantly worse than the comparison subjects across all prosodic measures (Table 2) with no significant group-by-task interaction ( $p > 0.50$ ). On the Sentence Discrimination Task, the patients showed significant decrements in performance on interrogative/declarative items as well as stress items (all  $p < 0.01$ ).

**TABLE 1. R<sup>2</sup> Values for Pitch (Distorted Tunes Task) and Affective Prosody (Voice Emotion Identification Task) Performance and Fractional Anisotropy Levels Within Dorsal and Ventral Auditory Streams and Anterior Cingulate Regions of Interest**

| Region of Interest/R <sup>2</sup>                           | Voice Emotion Identification Task | Distorted Tunes Task | Wisconsin Card Sorting Test |
|---|-----------------------------------|----------------------|-----------------------------|
| Right ventral stream (x=36, y=-21, z=-7) <sup>a</sup>       | 0.55**                            | 0.34**               | 0.07                        |
| Right dorsal stream (x=44, y=-4, z=15) <sup>a</sup>         | 0.45**                            | 0.31*                | 0.04                        |
| Left ventral stream (x=36, y=-21, z=-7) <sup>a</sup>        | 0.38**                            | 0.36**               | 0.03                        |
| Left dorsal stream (x=-47, y=-11, z=19) <sup>a</sup>        | 0.37**                            | 0.42**               | 0.16                        |
| Right anterior commissure (x=-47, y=-11, z=19) <sup>a</sup> | 0.01                              | 0.01                 | 0.43**                      |

<sup>a</sup> Talairach coordinates of region of interest locations for between-groups analysis.

\* $p \leq 0.01$ . \*\* $p \leq 0.005$ .

Within nonaffective prosody measures, the patients were significantly less sensitive (A') than comparison subjects on both the Sentence Discrimination Task (mean=0.86, SD=0.19, versus mean=0.98, SD=0.32, respectively) and the Semantic Comprehension Task (mean=0.83, SD=0.17, versus mean=0.98, SD=0.03) in detecting differing prosody or interrogative intent, respectively ( $p < 0.001$ ). However, there were no significant differences in terms of bias (B'') in the Sentence Discrimination Task (mean=0.54, SD=0.14, versus mean=0.68, SD=0.11) ( $p > 0.50$ ) or in the Semantic Comprehension Task (mean=0.37, SD=0.60, versus mean=0.72, SD=0.67) ( $p > 0.08$ ).

TABLE 2. Performance (% correct) on Indicated Tasks for 24 Schizophrenia Patients and 17 Healthy Comparison Subjects

| Measure 1 <sup>a</sup>                                  | Schizophrenia Patients |      | Comparison Subjects |      | t (df=34) | Effect size (d) <sup>b</sup> |
|---|------------------------|------|---------------------|------|-----------|------------------------------|
|   | Mean                   | SD   | Mean                | SD   |           |                              |
| Nonaffective prosody                                    |                        |      |                     |      |           |                              |
| Sentence Discrimination Task                            | 79.8                   | 20.6 | 98.1                | 2.5  | -3.8†     | 1.4                          |
| Sentence Discrimination Task: stress subfactor          | 79.1                   | 26.5 | 99.5                | 2.0  | -3.5†     | 1.1                          |
| Semantic Comprehension Task (interrogative/declarative) | 79.2                   | 16.3 | 96.3                | 5.9  | -3.9†     | 1.4                          |
| Affective prosody                                       |                        |      |                     |      |           |                              |
| Voice emotion identification (VOICE-ID)                 | 41.3                   | 14.7 | 64.1                | 13.5 | -5.0**    | 1.6                          |
| Voice emotion discrimination (VOICEDIS)                 | 70.0                   | 13.8 | 84.9                | 5.3  | -4.0**    | 1.4                          |
| Perceptual processing                                   |                        |      |                     |      |           |                              |
| Distorted Tunes Task                                    | 88.7                   | 17.1 | 69.7                | 19.0 | 3.1*      | 1.1                          |

<sup>a</sup> Between-group multivariate analysis of variance across measures ( $F=46.6$ ,  $df=4$ ,  $31$ ,  $p<0.0001$ ).

<sup>b</sup> Calculated after Cohen's (1988) conventions.

\* $p<0.05$ , between-group t test. \*\* $p<0.001$ , between-group t test. † $p<0.001$ , Mann-Whitney U test.

An examination of the interrelationship among prosody measures, pitch perception, and executive processing using principal components analysis (Figure 4, left) yielded only two criteria-meeting components, which when rotated revealed that the Distorted Tunes Task and the Sentence Discrimination Task loaded exclusively onto the first component (0.77 and 0.82, respectively) and the Wisconsin Card Sorting Test onto the second (0.95). The Semantic Comprehension Task, however, loaded significantly on both components (component 1=0.59, component 2=0.61). Correlations between affective prosody scores and their nonaffective counterparts were highly significant (Semantic Comprehension Task by VOICEID [ $r=0.64$ ,  $N=24$ ,  $p<0.001$ ], Sentence Discrimination Task by VOICEDIS [ $r_s=0.61$ ,  $N=21$ ,  $p<0.004$ ]) (Figure 4).

Finally, patient performance on nonaffective prosody measures did not significantly correlate with illness duration or medication dosage (all  $p>0.20$ ).

## Discussion

Emotion recognition deficits are associated with poor social and global functional outcome in schizophrenia (3, 4, 33), yet neural correlates have been investigated to only a limited degree. This study used a combined behavioral and voxelwise investigation to localize areas of potential relevance to impaired auditory prosodic processing. Significant correlations were observed between prosodic processing deficits and regions (e.g., prefrontal, peri-amygdalar) that are classically associated with neurocognitive dysfunction in schizophrenia (6). However, prominent deficits were also observed with reduced fractional anisotropy in regions such as the primary auditory radiations and the dorsal and ventral auditory streams, suggesting that impairments in voice emotion recognition arise from sensory-level disturbance in schizophrenia as well. Thus, these findings support our prior observations of processing deficits within, rather than across, sensory modalities (4) and suggest that functional and structural deficits within early sensory regions contribute to the overall pattern of cognitive dysfunction in schizophrenia.

In this study, structure-function relationships were assessed with voxelwise diffusion tensor imaging analysis. In contrast with volumetric approaches targeting gray matter regions, diffusion tensor imaging provides a measure of integrity of white matter tracts in the brain, which in turn may serve as a measure of dysconnectivity or dysmyelination within specific brain pathways (7, 8). Fractional anisotropy reductions in schizophrenia have been observed across brain regions, consistent with underlying reductions in oligodendrocytic markers (10). Furthermore, regionally specific correlations have already been observed for several well-validated measures (12, 13). Within this study, for example, reduced Wisconsin Card Sorting Test performance in schizophrenia correlated with reduced fractional anisotropy within the cingulate fasciculus, consistent with prior investigations in the field (14), as well as functional brain imaging studies (34). In contrast, no significant correlations were observed in auditory regions. Using the present approach, we have also previously demonstrated significant associations between verbal declarative memory, attention, and fractional anisotropy in task-relevant regions, attesting to the regional specificity of the current analysis approach (15).

The primary finding of this study was that reduced performance on the Distorted Tunes Task and the VOICEID correlated independently with reduced fractional anisotropy in brain regions containing primary auditory radiations from the medial geniculate nucleus of the thalamus to Heschl's gyrus and subsequent dorsal and ventral stream auditory projections. Additional areas of commonality included the genu and splenium of corpus callosum and the middle cingulate gyrus, consistent with lesion and neuroimaging studies implicating these regions in musical pitch and affective prosodic processing (19, 35, 36). Correlation clusters lateral to the amygdala were also observed in both tasks. As such, these findings indicate significant contributions of low-level auditory processing deficits to higher-order failures of neurocognition in schizophrenia.

In addition to areas of commonality, we also observed differences in correlation patterns between the pitch and prosodic tasks, particularly in the frontal cortex. Here, prosodic correlations extended more anteriorly to Brod-

mann's areas 44, 45, and 46, which are implicated particularly in speech perception (19) (Figure 2, bottom). Other areas involved in the affective evaluation of speech, such as the prefrontal and orbitofrontal cortex (21), also showed significant prosody-fractional anisotropy but not pitch-fractional anisotropy correlations (Figure 2, bottom; Figure 3B). The somewhat greater severity of prosodic versus pitch identification deficits observed in our patient group may reflect the greater extent of brain involvement engaged by the prosodic identification task.

The finding of sensory-level correlations in patients with schizophrenia is consistent with well-replicated deficits in auditory processing that have been demonstrated with both electrophysiological (e.g., reference 37) and behavioral (e.g., reference 22) approaches. Structural imaging studies of the auditory primary cortex are conflicting with some (38)—but not all (39)—studies, finding reduced volume of superior temporal auditory regions. However, postmortem changes analogous to those observed in the prefrontal cortex have been demonstrated in the auditory cortex as well (40), supporting auditory cortical involvement in the pathophysiology of schizophrenia.

In our second experiment, large effect size deficits ( $d > 1.1$ ) were observed in nonaffective prosodic perception and were correlated with deficits in affective prosodic perception. Thus, for instance, the subjects had difficulty in differentiating statements and questions based upon tone of voice, just as they had difficulty in differentiating sad from happy utterances. Both types of deficits are related to impaired pitch perception abilities, suggesting significant audiosensory antecedents. These findings indicate that dysprosodia, rather than being associated purely with emotional perceptual disturbances in schizophrenia, affects broader aspects of cognitive and social communication functioning.

---

Presented in part at the 12th annual meeting of the Society for Cognitive Neuroscience, New York, May 5–8, 2005. Received April 3, 2006; revisions received July 12 and Sept. 5, 2006; accepted Sept. 28, 2006. From the Program in Cognitive Neuroscience and Schizophrenia and the Cognitive Neurophysiology Laboratory, Nathan S. Kline Institute for Psychiatric Research; the Program in Cognitive Neuroscience, the City College of the City University of New York, New York; the Department of Psychiatry, New York University School of Medicine, New York; the Department of Psychiatry, University of Minnesota, Minneapolis, Minn.; and the Clinical Research Division, Nathan Kline Institute, Orangeburg, N.Y. Address correspondence and reprint requests to Dr. Javitt, Program in Cognitive Neuroscience and Schizophrenia, Nathan S. Kline Institute for Psychiatric Research, 140 Old Orangeburg Rd., Orangeburg, NY 10962; javitt@nki.rfmh.org (e-mail).

All authors report no competing interests.

Supported in part by NIMH grants NRSA F1-MH-067339 (to Dr. Leitman), K02-MH-01439 and R01-MH-49334 (to Dr. Javitt), R01-MH-64783 (to Dr. Hoptman), R01-MH-060662 (to Dr. Lim), and a Translational Research Scientist Award from the Burroughs Wellcome Fund (to Dr. Javitt).

The authors thank Dr. Denis Drayna for the use of his Distorted Tunes Test, Drs. Kerr and Neale for the use of their Voice Emotion Identification Task, and Raj Sangoi, R.T.(R.)M.R., for his help in MR scanning of participants.

---

## References

1. Bozidak VP, Kosmidis MH, Anezoulaki D, Giannakou M, Andreou C, Karavatos A: Impaired perception of affective prosody in schizophrenia. *J Neuropsychiatry Clin Neurosci* 2006; 18:81–85
2. Kerr SL, Neale JM: Emotion perception in schizophrenia: specific deficit or further evidence of generalized poor performance? *J Abnorm Psychol* 1993; 102:312–318
3. Brekke J, Kay DD, Lee KS, Green MF: Biosocial pathways to functional outcome in schizophrenia. *Schizophr Res* 2005; 80: 213–225
4. Leitman DI, Foxe JJ, Butler PD, Saperstein A, Revheim N, Javitt DC: Sensory contributions to impaired prosodic processing in schizophrenia. *Biol Psychiatry* 2005; 58:56–61
5. Green M: What are the functional consequences of neurocognitive deficits in schizophrenia? *Am J Psychiatry* 1996; 153: 321–330
6. Phillips ML, Drevets WC, Rauch SL, Lane R: Neurobiology of emotion perception, II: implications for major psychiatric disorders. *Biol Psychiatry* 2003; 54:515–528
7. Lim KO, Hedehus M, Moseley M, de Crespigny A, Sullivan EV, Pfefferbaum A: Compromised white matter tract integrity in schizophrenia inferred from diffusion tensor imaging. *Arch Gen Psychiatry* 1999; 56:367–374
8. Shenton ME, Dickey CC, Frumin M, McCarley RW: A review of MRI findings in schizophrenia. *Schizophr Res* 2001; 49:1–52
9. Kubicki M, McCarley R, Westin CF, Park HJ, Maier S, Kikinis R, Jolesz FA, Shenton ME: A review of diffusion tensor imaging studies in schizophrenia. *J Psychiatr Res* 2007; 41:15–30
10. Davis KL, Stewart DG, Friedman JI, Buchsbaum M, Harvey PD, Hof PR, Buxbaum J, Haroutunian V: White matter changes in schizophrenia: evidence for myelin-related dysfunction. *Arch Gen Psychiatry* 2003; 60:443–456
11. Pomara N, Crandall DT, Choi SJ, Johnson G, Lim KO: White matter abnormalities in HIV-1 infection: a diffusion tensor imaging study. *Psychiatry Res* 2001; 106:15–24
12. Hoptman MJ, Volavka J, Johnson G, Weiss E, Bilder RM, Lim KO: Frontal white matter microstructure, aggression, and impulsivity in men with schizophrenia: a preliminary study. *Biol Psychiatry* 2002; 52:9–14
13. Hoptman MJ, Ardekani BA, Butler PD, Nierenberg J, Javitt DC, Lim KO: DTI and impulsivity in schizophrenia: a first voxelwise correlational analysis. *Neuroreport* 2004; 15:2467–2470
14. Kubicki M, Westin CF, Nestor PG, Wible CG, Frumin M, Maier SE, Kikinis R, Jolesz FA, McCarley RW, Shenton ME: Cingulate fasciculus integrity disruption in schizophrenia: a magnetic resonance diffusion tensor imaging study. *Biol Psychiatry* 2003; 54: 1171–1180
15. Lim KO, Ardekani BA, Nierenberg J, Butler PD, Javitt DC, Hoptman MJ: Voxelwise correlational analyses of white matter integrity in multiple cognitive domains in schizophrenia. *Am J Psychiatry* 2006; 163:2008–2010
16. Butler PD, Zemon V, Schechter I, Saperstein AM, Hoptman MJ, Lim KO, Revheim N, Silipo G, Javitt DC: Early-stage visual processing and cortical amplification deficits in schizophrenia. *Arch Gen Psychiatry* 2005; 62:495–504
17. Drayna D, Manichaikul A, de Lange M, Snieder H, Spector T: Genetic correlates of musical pitch recognition in humans. *Science* 2001; 291:1969–1972
18. Parker GJ, Luzzi S, Alexander DC, Wheeler-Kingshott CA, Ciccarelli O, Lambon Ralph MA: Lateralization of ventral and dorsal auditory-language pathways in the human brain. *Neuroimage* 2005; 24:656–666
19. Arnott SR, Binns MA, Grady CL, Alain C: Assessing the auditory dual-pathway model in humans. *Neuroimage* 2004; 22:401–408

20. Belin P, Zatorre RJ: "What," "where" and "how" in auditory cortex. *Nat Neurosci* 2000; 3:965–966
21. Schirmer A, Kotz SA: Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn Sci* 2006; 10:24–30
22. Rabinowicz EF, Silipo G, Goldman R, Javitt DC: Auditory sensory dysfunction in schizophrenia: imprecision or distractibility? *Arch Gen Psychiatry* 2000; 57:1149–1155
23. Javitt DC, Doneshka P, Grochowski S, Ritter W: Impaired mismatch negativity generation reflects widespread dysfunction of working memory in schizophrenia. *Arch Gen Psychiatry* 1995; 52:550–558
24. Heaton RK, Chelune GJ, Talley JL, Kay GG, Curtis G: The Wisconsin Card Sorting Test. Odessa, Fla, Psychological Assessment Resources, 1993
25. Lim KO, Helpner JA: Neuropsychiatric applications of DTI: a review. *NMR Biomed* 2002; 15:587–593
26. Reese TG, Weisskopf RM, Wedeen VJ: Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. *Magn Reson Med* 2003; 49:177–182
27. Ardekani BA, Nierenberg J, Hoptman MJ, Javitt DC, Lim KO: MRI study of white matter diffusion anisotropy in schizophrenia. *Neuroreport* 2003; 14:2025–2029
28. Otsu N: A threshold selection model from gray-level histograms. *IEEE Trans (SMC)* 1979; 9:63–66
29. Baudewig J, Dechent P, Merboldt KD, Frahm J: Thresholding in correlation analyses of magnetic resonance functional neuroimaging. *Magn Reson Imaging* 2003; 21:1121–1130
30. Weintraub S, Mesulam MM, Kramer L: Disturbances in prosody: a right-hemisphere contribution to language. *Arch Neurol* 1981; 38:742–744
31. Snodgrass JG, Corwin J: Pragmatics of measuring recognition memory: applications to dementia and amnesia. *J Exp Psychol Gen* 1988; 117:34–50
32. Grier JB: Nonparametric indexes for sensitivity and bias: computing formulas. *Psychol Bull* 1971; 75:424–429
33. Murphy D, Cutting J: Prosodic comprehension and expression in schizophrenia. *J Neurol Neurosurg Psychiatry* 1990; 53:727–730
34. Buchsbaum BR, Greer S, Chang WL, Berman KF: Meta-analysis of neuroimaging studies of the Wisconsin Card-Sorting Task and component processes. *Hum Brain Mapp* 2005; 25:35–45
35. Ross ED, Thompson RD, Yenkosky J: Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain Lang* 1997; 56:27–54
36. Blood AJ, Zatorre RJ: Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc Natl Acad Sci USA* 2001; 98:11818–11823
37. Javitt DC, Shelley AM, Grochowski S, Ritter W: Mismatch negativity (MMN) as an index of impaired auditory sensory memory in schizophrenia. *Schizophr Res* 1995; 15:179
38. Hirayasu Y, McCarley RW, Salisbury DF, Tanaka S, Kwon JS, Frumin M, Snyderman D, Yurgelun-Todd D, Kikinis R, Jolesz FA, Shenton ME: Planum temporale and Heschl gyrus volume reduction in schizophrenia: a magnetic resonance imaging study of first-episode patients. *Arch Gen Psychiatry* 2000; 57:692–699
39. Kulynych JJ, Vladar K, Jones DW, Weinberger DR: Superior temporal gyrus volume in schizophrenia: a study using MRI morphometry assisted by surface rendering. *Am J Psychiatry* 1996; 153:50–56
40. Sweet RA, Bergen SE, Sun Z, Sampson AR, Pierri JN, Lewis DA: Pyramidal cell size reduction in schizophrenia: evidence for involvement of auditory feedforward circuits. *Biol Psychiatry* 2004; 55:1128–1137