Published in final edited form as:

Psychol Med. 2014 January ; 44(1): . doi:10.1017/S0033291713000834.

The 5% Difference: Early sensory processing predicts sarcasm perception in schizophrenia and schizoaffective disorder

Joshua T. Kantrowitz^{1,2,*}, Matthew J Hoptman^{1,3}, David I Leitman⁴, Gail Silipo¹, and Daniel C. Javitt^{1,2}

¹Schizophrenia Research Center, Nathan Kline Institute, Orangeburg, NY, United States

²Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY, United States

³Psychiatry, New York University, New York, NY, United States

⁴Neuropsychiatry, University of Pennsylvania, Philadelphia, PA, United States

Abstract

Background—Intact sarcasm perception is a critical component of social cognition and mentalizing (ability to understand the mental state of oneself and others). In sarcasm, tone of voice is used to negate the literal meaning of an utterance. In particular, changes in pitch are used to distinguish between sincere and sarcastic utterances. Schizophrenia patients show well-replicated deficits in auditory function and functional connectivity within and between auditory cortical regions. The present study investigates contributions of auditory deficits to sarcasm perception in schizophrenia.

Methods—Auditory measures including pitch-processing, auditory emotion recognition (AER) and sarcasm detection were obtained from 76 patients with schizophrenia/schizoaffective disorder and 72 controls. Resting state functional connectivity (rsFC) was obtained from a subsample, and was analyzed using seeds placed in both auditory cortex and meta-analysis-defined core-mentalizing regions relative to auditory performance.

Results—Patients showed large effect-size deficits across auditory measures. Sarcasm deficits correlated significantly with general functioning and impaired pitch-processing both across-groups and within the patient group alone. Patients also showed reduced sensitivity to alterations in mean pitch and variability. For patients, sarcasm discrimination correlated exclusively with level of rsFC within primary auditory regions, whereas for controls, correlations were observed exclusively within core-mentalizing regions (right posterior superior temporal gyrus, anterior superior temporal sulcus and insula, and left posterior medial temporal gyrus).

Conclusions—These findings confirm the contribution of auditory deficits to theory of mind impairments in schizophrenia, and demonstrate that functional connectivity within auditory-, but not core-mentalizing-, regions is rate limiting with respect to sarcasm detection in schizophrenia.

Keywords

Theory of mind; schizophrenia; auditory; resting state; sarcasm; MRI; theory of mind

*Contact info: Joshua Kantrowitz MD, Nathan Kline Institute for Psychiatric Research, 140 Old Orangeburg Rd Orangeburg, NY 10962, Phone: (212)-543-6711, jk3380@columbia.edu, Fax: (845) 398-6545

Drs. Hoptman and Leitman and Ms. Silipo have no conflicts.

Introduction

One critical method by which individuals communicate social information is via modulation of tone of voice (prosody). For example, in social situations, individuals sometimes choose to communicate information by saying the opposite of what they mean, while simultaneously modulating their tone of voice to communicate their counterfactual intent. This process - termed attitudinal prosody or “sarcasm” - is often done with hostility. Alternatively, however, it may be used adaptively to communicate displeasure or anger without using words or vocal characteristics (e.g. direct insults, yelling) that might otherwise be considered even more threatening. In either case, intact sarcasm perception is a critical component of social cognition and “mentalizing” [e.g. ability to understand the mental state of oneself and others—Theory of Mind (ToM)] (Winner *et al.*, 1998). In schizophrenia, social cognition is considered crucial for psychosocial impairment and rate limiting to social recovery (Fett *et al.*, 2011, Green *et al.*, 2008).

Although social cognition depends upon numerous forms of information that give rise to ToM, sarcasm may play a particularly crucial role in modern society across cultures [see (Cheang *et al.*, 2009) for review]. For example, sarcasm is important for reciprocal social interaction and the development of age-appropriate peer relations, is associated with decreased defensiveness and with effective problem solving skills, and is a common way to foster and conform to group membership in both the workplace and more causal settings (Gibbs Jr, 2000, Smith *et al.*, 1965). Because of its ambiguity vis-a-vis other forms of expressing anger, sarcasm may also be a less threatening way to convey displeasure or anger, and thus may serve an adaptive purpose (Fong, 2006, Jorgenson, 1996, Miron-Spektor *et al.*, 2011). Sarcasm differs from other forms of deception, such as “lying”, in that the sender is intending to have the receiver detect the true meaning, so that appropriate social interaction depends upon such detection. Sarcasm perception also requires more cognitive effort to discern, it is therefore more memorable than non-sarcastic speech and may enhance creative thinking (Gibbs Jr, 1986, Miron-Spektor *et al.*, 2011).

In schizophrenia, social cognitive deficits, including auditory (voice) emotion recognition (AER, “affective prosody”), have increasingly been linked to impairments in basic auditory function (Gold *et al.*, 2012, Kantrowitz *et al.*, In Press, Leitman *et al.*, 2010), over and above the contributions of general cognitive impairment. Furthermore, in schizophrenia, emotion recognition deficits correlate with neurophysiological dysfunction within sensory brain regions (Leitman *et al.*, 2007, Leitman *et al.*, 2011). The present study investigates impairments in sarcasm detection in schizophrenia from both a prosody/sensory and connectivity based perspective.

Our group first demonstrated sarcasm detection deficits in schizophrenia in 2006 (Leitman *et al.*, 2006), a finding that has since been replicated by several additional groups (Kern *et al.*, 2009, Mancuso *et al.*, 2011, Sparks *et al.*, 2010). Although sarcasm is typically studied in the context of ToM, similar to AER, sarcasm is often impossible to detect without making use of the psychophysical, non-verbal features that contradict the semantic features. In particular, proper detection of pitch modulations, such as mean voice pitch (F0M) and pitch variability (F0SD) is important for both AER and sarcasm (Banse *et al.*, 1996, Juslin *et al.*, 2001). Sarcasm is communicated by a reduction in F0M and F0SD in the range of 5-20% (Cheang *et al.*, 2008). Additional changes in duration, voice quality, intensity and tempo are observed more variably across studies (Cheang *et al.*, 2008, Rockwell, 2000, 2007). Thus, to the extent that ToM deficits in schizophrenia are driven by sensory-level impairments, as was suggested by our preliminary study (Leitman *et al.*, 2006), high correlation would be expected between sarcasm and AER performance.

At present, identifying neural substrates of ToM in general, and sarcasm detection in specific, remains an area of active research. ToM ability is linked to function within a widespread mentalizing network, consisting primarily of fronto-limbic brain regions. These regions are associated not only with ToM in general (Abu-Akel *et al.*, 2011, Loughhead *et al.*, 2010, Materna *et al.*, 2008, Pedersen *et al.*, 2012, Vollm *et al.*, 2006), but also with sarcasm in particular (Kipps *et al.*, 2009, Uchiyama, 2006, Uchiyama *et al.*, 2012). A recent meta-analysis of ToM neuroimaging studies identified a putative “core-mentalizing” network (Table 1) that includes primarily frontal and paralimbic/limbic areas, but also auditory cortex/temporal lobe (Mar, 2011), suggesting that these areas contribute significantly to the process of mentalization..

As previously, we assess sarcasm using the attitudinal subtest (APT) of the Aprosodia Battery (Orbelo *et al.*, 2005), adding an acoustic analysis of the individual items within the test, permitting evaluation of the degree to which variation in specific psychophysical parameters (F0M, F0SD, intensity) affected between-group performance. All subjects were also tested on simple pitch processing ability using the tone-matching test, on AER using the Juslin & Laukka battery (Juslin *et al.*, 2001). Based recent findings suggesting the relative importance of processing speed for community function (Bowie *et al.*, 2008, Kern *et al.*, 2011), all subjects were tested on the WAIS-3 Processing Speed Index (PSI), a component of the Performance IQ construct (Wechsler, 1997) as a proxy for general neurocognitive function.

In the present study, we also utilized resting state functional magnetic resonance imaging (rsfMRI) to evaluate the relationship of the brain connectivity of the auditory cortex/core-mentalizing region and sarcasm impairments in schizophrenia. rsfMRI is a recently developed, reliable (Turner *et al.*, 2012) technique that permits assessment of functional connectivity (rsFC) between brain regions by evaluating coherence of low frequency oscillations (0.01 – 0.1Hz) in blood-oxygen-level-dependent (BOLD) signal during the resting state (Biswal *et al.*, 1995, Biswal, 2010, Friston, 1994).

Previous studies (Das *et al.*, 2012a, Das *et al.*, 2012b)) have found impaired functional connectivity within putative mentalizing networks during a visual ToM task, but the specific role of sensory regions (primary auditory cortex) vs. “core-mentalizing regions” in ToM impairments was not assessed, and moreover, no auditory ToM tasks were assessed. In the present study, correlational seeds were placed within both auditory regions and regions identified from meta-analysis of recent ToM studies (Table 1) (Mar, 2011). Based upon our prior findings of significant auditory dysfunction in schizophrenia patients, we predicted that the schizophrenia group would show significant correlation between ToM deficits and rsFC within auditory regions, suggesting that auditory dysfunction might be rate limiting, whereas in controls the correlations would be to core-mentalizing regions as in prior studies.

Materials and Methods

Subjects

Subjects consisted of 76 medicated patients recruited from chronic inpatient (51%) and supervised residential outpatient sites (49%) associated with the Nathan Kline Institute (NKI) and 72 controls recruited from the healthy volunteer pool at NKI who had completed the sarcasm perception and all ancillary tasks (PSI, AER and tone-matching). All subjects signed informed consent, and patients met DSM-IV-TR criteria (First *et al.*, 1994) for either schizophrenia (n=61) or schizoaffective disorder (n=15), with no significant between diagnosis differences or hospital status seen on the auditory tasks (all $p>0.22$). We excluded controls with a history of an Axis I psychiatric disorder, as defined by the Structured Clinical Interview for DSM-IV. Patients and controls were excluded if they had any

neurological or auditory disorders noted on medical history or in prior records, or for alcohol or substance dependence within the last 6 months and/or abuse within the last month (First *et al.*, 1994). To assess the relationship with clinical symptoms and overall functioning, a subsample of subjects were interviewed using semi-structured clinical interviews [the Positive and Negative Symptom Scale (PANSS) (Kay *et al.*, 1987), the Global Assessment of Functioning (GAF) (Hall, 1995) and the Independent Living Scale (ILS) (Revheim *et al.*, 2004)]. Clinical ratings were consistent with moderate levels of illness.

Acoustic analysis of the psychophysical features of the individual stimuli of the sarcasm task was conducted on 52 patients and 61 controls for whom full item-level data were recorded. We also report on an imaging subset of 17 patients and 22 controls who completed the sarcasm task and participated in the MRI. The imaging subset included 2 patients and 8 controls who did not complete all the ancillary tasks and therefore were not included in the larger sample. See Supplemental Table 1 for details on demographics, clinical ratings and sub-sample sizes.

Auditory Tasks

Auditory tasks were presented on a CD player at a sound level that was comfortable for each listener in a sound-attenuated room.

Attitudinal Prosody (Sarcasm perception)—As previously (Leitman *et al.*, 2006), sarcasm perception was assessed using the attitudinal subtest (APT) of the Aprosodia Battery (Orbelo *et al.*, 2005). This battery consists of 10 semantically neutral sentences, such as ‘That was a smart thing to say’, that were recorded by a female speaker in both a sincere or sarcastic manner for a total of 20 unique utterances (10 pairs). These utterances were repeated twice for a total of 40 stimuli. Subjects were instructed to answer after each stimulus whether the speaker was being sincere or sarcastic. If subjects were confused by the instructions, further elaboration, using more commonplace synonyms, was provided. Subjects’ scores reflected overall percent correct (sarcasm) as the primary outcome, with “Hits”: correct detection of sarcastic utterances; and correct rejections (CR), i.e. correct detection of sincere utterances analyzed secondarily. As in the previous study (Leitman *et al.*, 2006), non-parametric signal detection measures of sensitivity (A') and Bias (B'') were calculated. Acoustic analysis of the individual stimuli was conducted with PRAAT software (Boersma, 2001). Mean (F0M) and variability (F0SD) of pitch were measured, as were mean and variability of intensity (volume).

Auditory emotion recognition (AER)—AER was assessed using 32 stimuli from Juslin and Laukka’s (Juslin *et al.*, 2001) emotional prosody task, as described previously (Gold *et al.*, 2012). The sentences were scored based on the speaker’s intended emotion (happy, sad, angry, fear or neutral). The sentences were semantically neutral and consisted of both statements and questions (i.e., “It is eleven o’clock”, “Is it eleven o’clock?”). Correct percent responses were analyzed across groups. These data represent a subsample that has been presented previously (Gold *et al.*, 2012).

Tone-matching task—Pitch processing was obtained using a simple tone-matching task (Leitman *et al.*, 2010). This task consists of pairs of 100-ms tones in series, with 500-ms intertone interval. Within each pair, tones are either identical or differed in frequency by specified amounts in each block (2.5%, 5%, 10%, 20%, or 50%). In each block, 12 of the tones are identical and 14 are dissimilar. Tones are derived from 3 reference frequencies (500, 1000, and 2000 Hz) to avoid learning effects. In all, the test consisted of 5 blocks of 26 pairs of tones.

Statistical analyses

Between-group comparisons were performed using multivariate ANOVA, with follow-up independent sample t- tests as required. Relationship among measures was determined by Pearson correlations and multivariate linear regression, as indicated. Two-tailed statistics are used throughout (α level of $p < 0.05$). Between group effect sizes (Cohen's d) were calculated (Cohen, 1988).

MRI Acquisition

Scanning took place on the 1.5T Siemens Vision Scanner (Erlangen, Germany) at the NKI Center for Advanced Brain Imaging. Participants received a magnetization prepared rapidly acquired gradient echo (MPRAGE) T1-weighted scan (TR=11.6 ms, TE=4.9 ms, TI=1122 ms, matrix=256x256, FOV=256 mm, slice thickness=1 mm, 190 slices, no gap, 1 acquisition), and a six minute resting state fMRI scan (TR=2000 ms, TE=50 ms, matrix=64x64, FOV=224 mm, 5 mm slice thickness, 22 slices, no gap, 180 acquisitions). For the resting state scan, participants were instructed to close their eyes and remain awake.

Data Analysis

Resting state data were preprocessed, as described elsewhere in detail (Hoptman, 2010, Kelly, 2008, Margulies, 2007), using scripts provided by the 1000 Functional Connectomes Project (Biswal, 2010), available at http://www.nitrc.org/projects/fcon_1000/. Briefly, the first 10 volumes were discarded to eliminate T1 relaxation effects. Thereafter, images were motion-corrected using AFNI (Cox, 1986). Next, time series were smoothed using a 6 mm FWHM Gaussian kernel and spatially normalized to MNI space ($2 \times 2 \times 2$ mm³ resolution) using FSL (www.fmrib.ox.ac.uk/fsl). The MPRAGE image was segmented using FSL's FAST software to obtain image masks for white matter (WM) and cerebrospinal fluid (CSF) compartments, which were also projected into MNI space and used to extract the corresponding time series for the resting state scans. The WM and CSF time series were then averaged across voxels within their respective compartments. These time series, as well as the time series for the six motion parameters and the global signal were regressed out from the MNI-space EPI time series.

Four auditory cortex anatomically-based regions of interest (ROIs) were derived from the Harvard Oxford Cortical Structural Atlas that is part of FSL: left and right Heschl's gyrus (HG), and left and right planum temporale (PT). These regions were extracted from the atlas and were thresholded at 50% probability. They were then used to extract time series for the relevant ROIs from the residualized images described above. The resulting time series for each ROI were then regressed against the residualized images described above to compute the functional connectivity for each region.

In order to examine the specific role of primary auditory cortex, we also examined a core-mentalizing network ROIs identified in an activation likelihood estimation (ALE) meta-analyses for nonstory-based studies of ToM [See reference (Mar, 2011): Table 6]. These regions were available as supplementary material (<http://www.annualreviews.org/doi/suppl/10.1146/annurev-psych-120709-145406>). We downloaded the ROIs and resampled them to $2 \times 2 \times 2$ mm³ MNI space prior to data analysis. We included the top ten ROI's, as ranked by ALE size. In some cases, whole brain coverage was not possible, so computations were limited to voxels for which all subjects had data. The analyzed core-mentalizing ROI's are listed in Table 1.

Group-level analyses were conducted using FSL's ordinary least squares (OLS) model implemented in FLAME. The two-sample t-tests on rsFC maps between patients and normal controls were performed to examine the differences in rsFC between the two groups. This

statistical procedure produced thresholded z -statistic maps of clusters defined by a threshold of $Z=2.3$ and a corrected cluster threshold of $p=0.05$ using Gaussian Random Field theory (Worsley, 2001), and revealed brain regions showing significantly different rsFC between patients and healthy controls. These same corrections applied to the regression analyses between rsFC and sarcasm.

Because small amounts of movement from volume to volume can influence rsFC results (Power *et al.*, 2012), we computed framewise displacement (FD) for our data, which was used as covariates in all analyses. Four patients and three controls in the original cohort of 21 patients and 25 controls, had $FD>0.5$ on greater than 35 volumes (i.e., less than 4.8 min of useable data) and were eliminated from final analyses, yielding a reported sample of 17 patients and 22 controls (Supplemental Table 1). Groups did not differ in FD ($p<0.42$).

Results

Between Group Auditory task analysis

As predicted, highly significant differences in percent correct were seen between groups on a multivariate ANOVA across the three auditory tasks (Figure 1A: $F_{1,146}=118$, $p<0.001$), as well as significant group X task interaction ($F_{2,145}=6.8$, $p<0.001$), reflecting larger effect size group differences for sarcasm ($F_{1,146}=132.4$, $p<0.001$, $d=1.4$ sd), than for either tone-matching ($F_{1,146}=46.7$, $p<0.001$, $d=1.0$ sd) or AER differences ($F_{1,146}=57.7$, $p<0.001$, $d=1.1$ sd). For tone-matching, both patients and controls showed the expected improvement across levels, suggesting appropriate task engagement (Supplemental Table 2).

Deficits in overall accuracy in the sarcasm task reflected a reduction in both hits (i.e. correct detection of sarcastic utterances: $F_{1,146}=73.5$, $p<0.001$) and correct rejections (CR: i.e. correct detection of sincere utterances: $F_{1,146}=21.1$, $p<0.001$) (Figure 1A). Furthermore, signal detection analysis (Supplemental Table 2) of both sarcasm and tone-matching showed that both resulted from a reduction in sensitivity (sarcasm: $t_{139}=8.1$, $p<0.001$; tone-matching: $t_{146}=5.1$, $p<0.001$), with no significant difference in bias (sarcasm: $t_{139}=1.4$, $p=0.17$, tone-matching: $t_{146}=0.3$, $p=0.76$). Between-group percent correct differences for sarcasm ($F_{4,143}=57.7$, $p<0.001$), tone-matching ($F_{4,143}=20.7$, $p<0.001$) and AER ($F_{4,143}=29.2$, $p<0.001$) remained significant when controlling for age, gender and PSI, suggesting that they could not be solely accounted for by demographic variables or general cognitive ability.

Relationship among auditory measures—In the absence of covariates, sarcasm perception correlated significantly with both tone-matching performance ($r=0.56$, $n=148$, $p<0.001$) (Figure 1B) and AER ($r=0.70$, $n=148$, $p<0.001$) (Figure 1C) across groups. These correlations remained significant across group when controlling for PSI ($R=0.77$, $F_{3,144}=73.2$, $p<0.001$) or group membership ($R=0.80$, $F_{3,144}=87.4$, $p<0.001$) (Supplemental Table 3).

Moreover, independent correlations with sarcasm perception were seen within the schizophrenia group for tone-matching ($r=0.45$, $n=76$, $p<0.001$), AER ($r=0.56$, $n=76$, $p<0.001$) and PSI ($r=0.40$, $n=76$, $p<0.001$). In contrast, no significant correlation between sarcasm and tone-matching was observed in controls alone ($r=0.18$, $n=72$, $p=0.13$), although the correlations with PSI ($r=0.28$, $n=72$, $p=0.018$) and AER ($r=0.54$, $n=72$, $p<0.001$) remained significant.

Relationship with outcome and demographics clinical ratings—No significant correlations were seen between sarcasm perception and subject socioeconomic status (SES), duration of illness or CPZ equivalents. Significant correlations were seen between sarcasm

perception and general function measures GAF ($r=0.28$, $n=66$, $p=0.022$) and ILS ($r=0.38$, $n=73$, $p=0.001$).

Acoustic analysis

The psychophysical features (F0M, F0SD and intensity values) for the sarcastic and sincere stimuli were extracted using acoustic analysis (PRAAT) software (Table 2). Across all unique utterances in this task ($n=10$ pairs), F0M of sarcastic stimuli was significantly lower ($12\pm5\%$, $p>0.0001$) in sarcastic stimuli as compared to the corresponding sincere stimuli, while F0SD showed a trend towards being significantly lower ($21\pm28\%$, $p=0.065$). Other measures, such as intensity and intensity variability, were not significantly different.

To explore the influence of specific features on sarcasm perception (overall percent correct), we conducted a 3-way, group (patient/control) X intention (sincere/sarcastic) X stimulus (unique sentence/utterance) analysis across the 10 pairs of stimuli. As expected, patients showed worse overall performance ($F_{1,111}=102.2$, $p<0.00001$), as well as lower relative performance for sarcastic vs. sincere stimuli (group X intention: $F_{1,103}=15.7$, $p<0.0001$). Patients also showed differential response across stimuli vs. controls as reflected in a significant group X intention X stimulus ($F_{9,103}=3.2$, $p=0.002$).

In order to parse this interaction, stimuli were divided according to levels of F0M (Figure 2A) and F0SD (Figure 2B) based on the magnitude of the percent difference between sincere and sarcastic forms. Patients performed significantly below chance performance for stimuli with $<5\%$ difference in F0M between the sincere and sarcastic forms ($t_{51}=2.94$, $p=0.005$), suggesting that they heard stimuli with low levels of F0M difference as being actively sincere. Furthermore, significant group X F0M level ($F_{2,110}=4.4$, $p=0.015$) and group X F0SD level interactions ($F_{2,110}=8.8$, $p=0.0002$) was seen (Figure 2B).

Relationship of Functional Connectivity and Sarcasm

In order to determine potential neural substrates of sarcasm perception, an rsFC analysis was conducted. Seeds were placed in four auditory and ten core-mentalizing regions (Table 1). rsFC was then determined on a voxel-wise basis throughout brain, and regions that showed significant rsFC correlations to the seed relative to performance on the sarcasm task were identified. These regions were then used for across group correlational analysis. Separate analysis' were done for auditory and core seeds.

For auditory regions, a significant correlation was observed between sarcasm performance and rsFC between right HG and left precentral gyrus/medial frontal gyrus (Figure 3A, Supplemental Table 4). Clusters extended to the left postcentral gyrus (BA 3/4/1). A regression performed across all subjects showed significant correlation between rsFC and sarcasm performance, even when group was included as a factor ($r=0.37$, $n=39$, $p=0.022$, Figure 3B). The correlation was independently significant only within the patient (BA 6, $r=0.60$, $n=17$, $p=0.011$), but not control ($r=0.01$, $n=22$, $p=0.96$), group. The two correlation coefficients, moreover, differed significantly ($p=0.049$). No significant correlation regions relative to sarcasm were detected for the remaining auditory seeds (left HG or right/left PT).

For core-mentalizing regions, significant rsFC correlation regions were observed for 4 of the 10 seed locations (Supplemental Table 5). rsFC was primarily between the seed region and the precuneus/cuneus and surrounding cortex (Figure 3C). Core regions for which significant correlation patterns were observed included the right posterior superior temporal gyrus (R pSTG, Figure 3D), left posterior medial temporal gyrus (L pMTG), right anterior superior temporal sulcus (R aSTS), and right insula.

For both R pSTG (Figure 3D) and L pMTG, regression performed across all subjects showed a correlation with sarcasm that remained significant even after group was included as a factor, but which were independently significant only within the control, but not the patient groups (Supplemental Table 5). For R aSTS and insula, correlations were significant within the control group only.

Discussion

ToM and sarcasm perception depend upon interactions within large-scale brain networks involving sensory, as well as putative “core-mentalizing” brain regions identified in a recent meta-analysis (Mar, 2011). Dysfunction anywhere within these networks will produce behavioral deficits, with the pattern depending upon the nature and locus of the dysfunction. The present study confirms sarcasm detection deficits in schizophrenia, along with more basic auditory and emotion processing deficits, and relates these deficits to impairments within specific sensory/cognitive regions using both correlational analyses and rsfMRI. In patients, deficits in sarcasm detection correlate significantly with auditory dysfunction even following control for more general cognitive impairments, as reflected in PSI. Furthermore, in patients, but not controls, sarcasm detection performance correlates with functional connectivity between right auditory cortex, a region known to be involved in prosodic processing (Mitchell *et al.*, 2008) and left precentral gyrus, a region with a known role in emotion processing (Li *et al.*, 2012). In contrast, in controls, but not patients, correlations were only seen within core-mentalizing regions.

We have previously shown that inability to process mean pitch (F0M) and pitch variability (F0SD) contributes significantly to AER deficits in schizophrenia (Gold *et al.*, 2012). In this study, patients performed significantly below chance for stimuli in which the mean pitch (F0M) difference between sincere and sarcastic utterances was <5%, suggesting that they heard these stimuli as being actively sincere, even while controls heard them as primarily sarcastic (Figure 2A). Patients showed a similar inability to utilize pitch variability (F0SD) in discerning between sarcasm and sincere (Figure 2B). These findings thus suggest that impaired sensitivity to pitch change in schizophrenia contributes significantly to impairments in ToM, as well as AER, discrimination.

Controls showed a significant correlation between sarcasm detection and rsFC levels for four of the brain regions identified in the meta-analysis (Mar, 2011), but not to auditory regions. The rsFC targets of these core-mentalizing regions relative to sarcasm detection centered on the precuneus/cuneus (Figure 3C), consistent with a proposed role for this region as a processing core essential for mentalizing (Hagmann *et al.*, 2008). To our knowledge, this is the first study to show between-region rsFC correlations with sarcasm performance in control subjects, although reductions in task-related functional connectivity among R STG, insula and cuneus have been previously reported during a visual ToM task performance in schizophrenia (Das *et al.*, 2012a, Das *et al.*, 2012b). Although Das et al examined functional connectivity of core mentalizing regions, they did not evaluate the role of sensory regions using either behavioral or neuroimaging measures. In general, potential sensory contributions to cognitive impairments in schizophrenia have been understudied, primarily because researchers do not collect the appropriate measures in designing and conducting their studies.

In contrast, patients only had correlations between sarcasm performance and the rsFC of right auditory cortex, and not core-mentalizing regions. This suggests that for patients, function of the primary auditory cortex may be rate limiting for sarcasm detection and may determine the difference between chance-level and above chance level performance, while for controls, connectivity among higher brain regions may be rate limiting and may

differentiate between good and near-perfect performance. Taken together, the present results highlight the difficulties faced by patients in normal social situations. As shown by acoustic analysis of the stimuli, the degree of pitch difference typically used to communicate sarcasm is relatively small and only slightly above threshold for most healthy individuals (Supplemental Table 2). Thus, other factors involving mentalization determine the level of performance for controls. In contrast, for patients, the degree of pitch difference is at or below their detection threshold, so auditory discrimination itself becomes rate limiting. Similarly, for controls, integrity of rsFC between core-mentalizing regions becomes rate limiting, whereas for patients correlations are observed only for right auditory cortex.

Despite the robust findings, there are specific limitations to the present study. First, we did not include general ToM tasks. In other studies, several aspects of ToM, such as detection of lying, have been paradoxically intact despite deficits in detection of sarcasm (Kern *et al.*, 2009). These findings were interpreted as supporting a specific role for detection of psychophysical features, such as FOM and FOSD, as reported here. Nevertheless, deficits are observed across a large range of ToM tasks in schizophrenia, not all of which require pitch-based acoustic processing (Biedermann *et al.*, 2012). The reduced rsFC observed in core-mentalizing frontolimbic regions in patients suggests that tasks relying specifically on these regions should be sensitive to cortical processing impairments in schizophrenia.

Second, although we did not find any relationship with mean antipsychotic drug dose or illness duration, our sample consisted primarily of chronically medicated patients at a state psychiatric hospital and related supervised outpatient facilities, as well as excluding controls with Axis I disorders. Moreover, our measure of general cognition, while strongly supported by the literature (Bowie *et al.*, 2008, Kern *et al.*, 2011), does not measure other potentially important domains. Thus, we cannot fully separate the effects of medication or chronic cognitive or functional impairment on sarcasm detection or rsFC. However, it should be noted that rsFC is task-independent (i.e. obtained while subjects are not performing a specific behavioral paradigm), so that rsFC correlations are not confounded by the direct influence of motivation or performance on brain measures. Third, while we based our analyzed core-mentalizing ROI's on an independent meta-analysis, we note that our chosen regions are an approximation of the "true" core-mentalizing region, limiting the specificity of our findings. Finally, our acoustic analysis was limited by the limited number of stimuli in the task, and future studies should confirm these findings with a larger stimulus set.

In conclusion, schizophrenia patients show severe deficits in daily interaction that stem at least in part from simple communicatory disturbance. The insensitivity of patients to subtle pitch features is related to dysfunction at the level of auditory sensory cortex, and is detected with rsfMRI, which is an increasingly important method for delineating patterns of neural dysfunction underlying cognitive impairments in schizophrenia. Given the importance of functions such as AER or sarcasm detection to daily function, approaches to improving tonal detection ability, as well as connectivity between auditory and frontal-limbic brain regions, may be critical for sensory-based cognitive remediation and functional rehabilitation in schizophrenia. If replicated, the present results may be useful informing potential sensory based cognitive remediation (Fisher *et al.*, 2009, Kantrowitz *et al.*, 2009, Norton *et al.*, 2011) and brain stimulation based treatment development (Clark *et al.*, 2012, Demirtas-Tatlidede *et al.*, 2013).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Preparation of this manuscript was supported in part by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant Number UL1 RR024156 and the Dr. Joseph E. And Lillian Pisetsky Young Investigator Award for Clinical Research in Serious Mental Illness to JTK, R01 DA03383, P50 MH086385 and R37 MH49334 to DCJ, K01MH094689 and a NARSAD (Brain and Behavior Research Foundation) Young Investigator to DIL and R01MH064783 and R21MH084031 to MJH. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

We would also like to thank Diane Orbelo and Elliot Ross for the use of the APT.

Declaration of interest:

Dr. Kantrowitz reports having received consulting payments within the last 2 years from Quadrant Health, RTI Health solutions, the Saco Medical Group and AgencyRx. He has conducted clinical research supported by the NIMH, Roche-Genentech, EnVivo, Psychogenics, Sunovion, Novartis, Pfizer, Lilly, Jazz and GlaxoSmithKline. He owns a small number of shares of common stock in GlaxoSmithKline.

Dr. Javitt reports having received consulting payments within the last 2 years from Schering-Plough, Takeda, NPS Allelix, Solvay, Sepracor, AstraZeneca, Pfizer, Cypress, Merck, Sunovion, Lilly, and BMS. He has received research support from Pfizer and Roche. He holds intellectual property rights for use of NMDA modulators in treatment of neuropsychiatric disorders. He served on the advisory board of Promentis.

References

- Abu-Akel A, Shamay-Tsoory S. Neuroanatomical and neurochemical bases of theory of mind. *Neuropsychologia*. 2011; 49:2971–2984. [PubMed: 21803062]
- Banase R, Scherer K. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*. 1996; 70:614–636.
- Biedermann F, Frajo-Apor B, Hofer A. Theory of mind and its relevance in schizophrenia. *Current Opinion in Psychiatry*. 2012; 25:71–75. [PubMed: 22249083]
- Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine*. 1995; 34:537–541. [PubMed: 8524021]
- Biswal BB. Toward discovery science of human brain function. *PNAS : Proceedings of the National Academy of Sciences*. 2010; 107:4734–4739.
- Boersma P. Praat, a system for doing phonetics by computer. *Glott International*. 2001; 5:341–345.
- Bowie CR, Leung WW, Reichenberg A, McClure MM, Patterson TL, Heaton RK, Harvey PD. Predicting schizophrenia patients' real-world behavior with specific neuropsychological and functional capacity measures. *Biological Psychiatry*. 2008; 63:505–511. [PubMed: 17662256]
- Cheang HS, Pell MD. The sound of sarcasm. *Speech Communication*. 2008; 50:366–381.
- Cheang HS, Pell MD. Acoustic markers of sarcasm in Cantonese and English. *Journal of the Acoustical Society of America*. 2009; 126:1394–1405. [PubMed: 19739753]
- Clark VP, Coffman BA, Mayer AR, Weisend MP, Lane TD, Calhoun VD, Raybourn EM, Garcia CM, Wassermann EM. TDCS guided using fMRI significantly accelerates learning to identify concealed objects. *Neuroimage*. 2012; 59:117–128. [PubMed: 21094258]
- Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Assoc; 1988.
- Cox, RWx. Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*. 1986; 29:162–173. [PubMed: 8812068]
- Das P, Calhoun V, Malhi GS. Mentalizing in male schizophrenia patients is compromised by virtue of dysfunctional connectivity between task-positive and task-negative networks. *Schizophrenia Research*. 2012a; 140:51–58. [PubMed: 22795367]
- Das P, Lagopoulos J, Coulston CM, Henderson AF, Malhi GS. Mentalizing impairment in schizophrenia: a functional MRI study. *Schizophrenia Research*. 2012b; 134:158–164. [PubMed: 21943555]

- Demirtas-Tatlidede A, Vahabzadeh-Hagh AM, Pascual-Leone A. Can noninvasive brain stimulation enhance cognition in neuropsychiatric disorders? *Neuropharmacology*. 2013; 64:566–578. [PubMed: 22749945]
- Fett AK, Viechtbauer W, Dominguez MD, Penn DL, van Os J, Krabbendam L. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. *Neuroscience & Biobehavioral Reviews*. 2011; 35:573–588. [PubMed: 20620163]
- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. Structured clinical interview for Axis I DSM-IV disorders - Patient edition (SCID-I/P). New York: Biometrics Research Department, New York State Psychiatric Institute; 1994.
- Fisher M, Holland C, Merzenich MM, Vinogradov S. Using neuroplasticity-based auditory training to improve verbal memory in schizophrenia. *American Journal of Psychiatry*. 2009; 166:805–811. [PubMed: 19448187]
- Fong CT. The effects of emotional ambivalence on creativity. *Academy of Management Journal*. 2006; 49:1016–1030.
- Friston KJ. Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*. 1994; 2:56–78.
- Gibbs RW Jr. On the Psycholinguistics of Sarcasm. *Journal of experimental psychology: general*. 1986; 115:3–15.
- Gibbs RW Jr. Irony in Talk Among Friends. *Metaphor and Symbol*. 2000; 15:5–27.
- Gold R, Butler PD, Revheim N, Leitman DI, Hansen JA, Gur RC, Kantrowitz JT, Laukka P, Juslin PN, Silipo GS, Javitt DC. Auditory emotion recognition impairments in Schizophrenia: Relationship to acoustic features and cognition. *American Journal of Psychiatry*. 2012; 169:424–432. [PubMed: 22362394]
- Green M, Leitman D. Social cognition in schizophrenia. *Schizophrenia Bulletin*. 2008; 34:670–672. [PubMed: 18495642]
- Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ, Sporns O. Mapping the structural core of human cerebral cortex. *PLoS Biol*. 2008; 6:e159. [PubMed: 18597554]
- Hall RC. Global assessment of functioning. A modified scale. *Psychosomatics*. 1995; 36:267–275. [PubMed: 7638314]
- Hoptman MJ. Amplitude of low-frequency oscillations in schizophrenia: a resting state fMRI study. *Schizophrenia Research*. 2010; 117:13–20. [PubMed: 19854028]
- Jorgenson J. The functions of sarcastic irony in speech. *Journal of Pragmatics*. 1996; 26:613–634.
- Juslin PN, Laukka P. Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion*. 2001; 1:381–412. [PubMed: 12901399]
- Kantrowitz JT, Leitman DI, Lehrfeld JM, Laukka P, Juslin PN, Butler PD, Silipo G, Javitt DC. Reduction in tonal discriminations predicts receptive emotion processing deficits in schizophrenia and schizoaffective disorder. *Schizophrenia Bulletin*. (In Press).
- Kantrowitz JT, Revheim N, Pasternak R, Silipo G, Javitt DC. It's all in the cards: effect of stimulus manipulation on Wisconsin Card Sorting Test performance in schizophrenia. *Psychiatry Research*. 2009; 168:198–204. [PubMed: 19573928]
- Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*. 1987; 13:261–276. [PubMed: 3616518]
- Kelly AMC. Competition between functional brain networks mediates behavioral variability. *NeuroImage*. 2008; 39:527–537. [PubMed: 17919929]
- Kern RS, Gold JM, Dickinson D, Green MF, Nuechterlein KH, Baade LE, Keefe RS, Mesholam-Gately RI, Seidman LJ, Lee C, Sugar CA, Marder SR. The MCCB impairment profile for schizophrenia outpatients: results from the MATRICS psychometric and standardization study. *Schizophrenia Research*. 2011; 126:124–131. [PubMed: 21159492]
- Kern RS, Green MF, Fiske AP, Kee KS, Lee J, Sergi MJ, Horan WP, Subotnik KL, Sugar CA, Nuechterlein KH. Theory of mind deficits for processing counterfactual information in persons with chronic schizophrenia. *Psychological Medicine*. 2009; 39:645–654. [PubMed: 18694537]
- Kipps CM, Nestor PJ, Acosta-Cabronero J, Arnold R, Hodges JR. Understanding social dysfunction in the behavioural variant of frontotemporal dementia: the role of emotion and sarcasm processing. *Brain*. 2009; 132:592–603. [PubMed: 19126572]

- Leitman DI, Hoptman MJ, Foxe JJ, Saccente E, Wylie GR, Nierenberg J, Jalbrzikowski M, Lim KO, Javitt DC. The neural substrates of impaired prosodic detection in schizophrenia and its sensorial antecedents. *American Journal of Psychiatry*. 2007; 164:474–482. [PubMed: 17329473]
- Leitman DI, Laukka P, Juslin PN, Saccente E, Butler P, Javitt DC. Getting the Cue: Sensory Contributions to Auditory Emotion Recognition Impairments in Schizophrenia. *Schizophrenia Bulletin*. 2010; 36:545–556. [PubMed: 18791077]
- Leitman DI, Wolf DH, Laukka P, Ragland JD, Valdez JN, Turetsky BI, Gur RE, Gur RC. Not Pitch Perfect: Sensory Contributions to Affective Communication Impairment in Schizophrenia. *Biological Psychiatry*. 2011; 70:611–618. [PubMed: 21762876]
- Leitman DI, Ziwich R, Pasternak R, Javitt DC. Theory of Mind (ToM) and counterfactuality deficits in schizophrenia: misperception or misinterpretation? *Psychological Medicine*. 2006; 36:1075–1083. [PubMed: 16700967]
- Li HJ, Chan RC, Gong QY, Liu Y, Liu SM, Shum D, Ma ZL. Facial emotion processing in patients with schizophrenia and their non-psychotic siblings: a functional magnetic resonance imaging study. *Schizophrenia Research*. 2012; 134:143–150. [PubMed: 22113155]
- Loughead JW, Luborsky L, Weingarten CP, Krause ED, German RE, Kirk D, Gur RC. Brain activation during autobiographical relationship episode narratives: a core conflictual relationship theme approach. *Psychotherapy Research*. 2010; 20:321–336. [PubMed: 20099204]
- Mancuso F, Horan WP, Kern RS, Green MF. Social cognition in psychosis: multidimensional structure, clinical correlates, and relationship with functional outcome. *Schizophrenia Research*. 2011; 125:143–151. [PubMed: 21112743]
- Mar RA. The neural bases of social cognition and story comprehension. *Annu Rev Psychol*. 2011; 62:103–134. [PubMed: 21126178]
- Margulies DS. Mapping the functional connectivity of anterior cingulate cortex. *NeuroImage*. 2007; 37:579–588. [PubMed: 17604651]
- Materna S, Dicke PW, Thier P. The posterior superior temporal sulcus is involved in social communication not specific for the eyes. *Neuropsychologia*. 2008; 46:2759–2765. [PubMed: 18585742]
- Miron-Spektor E, Efrat-Treister D, Rafaeli A, Schwarz-Cohen O. Others' anger makes people work harder not smarter: the effect of observing anger and sarcasm on creative and analytic thinking. *Journal of Applied Psychology*. 2011; 96:1065–1075. [PubMed: 21574675]
- Mitchell RL, Ross ED. fMRI evidence for the effect of verbal complexity on lateralisation of the neural response associated with decoding prosodic emotion. *Neuropsychologia*. 2008; 46:2880–2887. [PubMed: 18635241]
- Norton DJ, McBain RK, Ongur D, Chen Y. Perceptual training strongly improves visual motion perception in schizophrenia. *Brain and Cognition*. 2011; 77:248–256. [PubMed: 21872380]
- Orbelo DM, Grim MA, Talbott RE, Ross ED. Impaired comprehension of affective prosody in elderly subjects is not predicted by age-related hearing loss or age-related cognitive decline. *Journal of Geriatric Psychiatry and Neurology*. 2005; 18:25–32. [PubMed: 15681625]
- Pedersen A, Koelkebeck K, Brandt M, Wee M, Kueppers KA, Kugel H, Kohl W, Bauer J, Ohrmann P. Theory of mind in patients with schizophrenia: Is mentalizing delayed? *Schizophrenia Research*. 2012; 137:224–229.
- Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*. 2012; 59:2142–2154. [PubMed: 22019881]
- Revheim N, Medalia A. The independent living scales as a measure of functional outcome for schizophrenia. *Psychiatric Services*. 2004; 55:1052–1054. [PubMed: 15345767]
- Rockwell P. Actors', partners', and observers' perceptions of sarcasm. *Perceptual & Motor Skills*. 2000; 91:665–668. [PubMed: 11065329]
- Rockwell P. Vocal features of conversational sarcasm: a comparison of methods. *Journal of Psycholinguistic Research*. 2007; 36:361–369. [PubMed: 17273931]
- Smith EE, White HL. Wit Creativity and Sarcasm. *Journal of Applied Psychology*. 1965; 49:131–134. [PubMed: 14318069]

- Sparks A, McDonald S, Lino B, O'Donnell M, Green MJ. Social cognition, empathy and functional outcome in schizophrenia. *Schizophrenia Research*. 2010; 122:172–178. [PubMed: 20609567]
- Turner JA, Chen H, Mathalon DH, Allen EA, Mayer AR, Abbott CC, Calhoun VD, Bustillo J. Reliability of the amplitude of low-frequency fluctuations in resting state fMRI in chronic schizophrenia. *Psychiatry Research*. 2012; 201:253–255. [PubMed: 22541511]
- Uchiyama H. Neural substrates of sarcasm: a functional magnetic-resonance imaging study. *Brain Research*. 2006; 1124:100–110. [PubMed: 17092490]
- Uchiyama HT, Saito DN, Tanabe HC, Harada T, Seki A, Ohno K, Koeda T, Sadato N. Distinction between the literal and intended meanings of sentences: a functional magnetic resonance imaging study of metaphor and sarcasm. *Cortex*. 2012; 48:563–583. [PubMed: 21333979]
- Vollm BA, Taylor AN, Richardson P, Corcoran R, Stirling J, McKie S, Deakin JF, Elliott R. Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage*. 2006; 29:90–98. [PubMed: 16122944]
- Wechsler, DA. Wechsler Adult Intelligence Scale-III. New York: Psychological Corporation; 1997.
- Winner E, Brownell H, Happe F, Blum A, Pincus D. Distinguishing lies from jokes: Theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain and Language*. 1998; 62:89–108. [PubMed: 9570881]
- Worsley, KJ. Statistical analysis of activation images. In: Jezzard, P.; Matthew, PM.; Smith, SM., editors. *Functional MRI: An introduction to methods*. Oxford, UK: Oxford University Press; 2001.

Figure 1.

(A) Bar graph (Mean $\pm$ SEM) of percent correct for auditory processing tasks.. Sarcasm: overall correct; Hits: correct detection of sarcastic utterances; CR: correct rejections, i.e. correct detection of sincere utterances; AER: Auditory Emotion Recognition; TMT: Tone-matching Task. * $p<0.05$ on independent samples t-test

(B,C) Scatter plots of sarcasm perception vs. AER (B) and TMT (C). Controls are in black circles and patients in white triangles. Significant within group correlations with TMT were seen for schizophrenia patients ($r=0.45$, $n=76$, $p<0.001$), but not controls ($r=0.18$, $n=72$, $p=0.13$).

Figure 2.

(A,B) Bar graph (Mean $\pm$ SEM) of percent correct of sarcasm, with stimuli categorized by the percent difference in FOM (A) or F0SD (B) between corresponding sincere and sarcastic statements. A group X FOM ($F_{2,110}=4.4$, $p=0.015$) and F0SD ($F_{2,110}=8.8$, $p=0.0002$) level interaction was found, suggesting patients had increasing difficulty with smaller percent differences. Dashed line represents chance performance rate (50%).

* $p<0.001$ between group on independent sample t-test.

^a $p<0.01$ vs. chance (50%) for patient group on one-sample t-test.

Figure 3.

(A) Correlation region of resting state functional connectivity (rsFC) for right Heschl's gyrus (RHG) and sarcasm perception in patients with schizophrenia. (B) Scatter plot of sarcasm perception vs. rsFC in RHG. (C) Correlation region of resting state functional connectivity (rsFC) for Left posterior Medial Temporal Gyrus (Lp MTG: red), Right anterior Superior Temporal Sulcus (Ra STS: blue), Right insula (green), and Right posterior Superior Temporal Gyrus (Rp STG: cyan) and sarcasm perception in controls. Correlations thresholded at $p < 0.05$, corrected for Gaussian Random Fields. (D) Scatter plot of sarcasm perception vs. rsFC in right posterior superior temporal gyrus (R pSTG). Controls are in black circles and patients in white triangles for figures 3B and D. Correlations remained significant using Spearman (non-parametric) correlations ($\rho=0.43$, $p=0.043$).

Table 1

Auditory and ToM core mentalizing seed regions

Seed Region	Brodmann Area (location)	ALE ($\times 10^3$)
Auditory		
R HG	41/42	--
L HG	41/42	--
R PT	22	--
L PT	22	--
Core-Mentalizing¹		
L MPFC	8	4.16
R aSTS	21	3.86
L IFG (orbitalis)	47	3.64
R Insula	47	3.09
L pMTG	22	3.05
R Precuneus	7	2.64
L aMTG	21	2.50
R pSTG	13	2.44
R amygdala	---	2.34
R IFG (opercularis)	44	2.23

¹Seeds identified in an activation likelihood estimation (ALE) meta-analyses of nonstory-based studies of ToM (Mar 2011)

L = left, R = right, p=posterior, a=anterior, HG=Heschel's Gyrus, PT=Planum Temporale, MTG = middle temporal gyrus, STS= superior temporal sulcus, STG=superior temporal gyrus, MPFG=medial prefrontal frontal gyrus, IFG=inferior frontal gyrus, rsFC=resting state functional connectivity

Table 2

Acoustic analysis of Sarcasm Task

	Sarcastic	Sincere	% Diff	t	df	p
F0M (Hz)	215±20	245±21	12.1±5%	6.5	9	0.0001
F0SD (Hz)	54±25	62±20	21.4±28%	2.1	9	0.065
Intensity Mean (dB)	66±3	66±3	0.1±0.1%	0.2	9	0.8
Intensity SD (dB)	10±2	10±2	0.1±0.1%	0.4	9	0.7