Short Communication

Basic auditory processing and emotion recognition in individuals at clinical high risk for psychosis[☆]

Michael S. Kraus [a],[*], Trina M. Walker [a], Diana Perkins [b], Richard S.E. Keefe [a]

[a] Department of Psychiatry & Behavioral Sciences, Duke University Medical Center, 10 Duke Medicine Circle, Durham, NC 27710, USA
[b] Department of Psychiatry, University of North Carolina, Chapel Hill, NC, USA

1. Introduction

The ability to understand others' emotions from subtle cues is a critical skill for navigating the complex social world, with increasing demands placed on this capability as children enter adolescence. Individuals at clinical high-risk (CHR) of developing psychosis exhibit social isolation with specific impairments in social cognition (Thompson et al., 2011) and emotion recognition (Addington et al., 2008; Amminger et al., 2011; Corcoran et al., 2015). The degree of social impairment and the severity of emotion recognition deficits are greater in individuals who later convert to psychosis than in those who do not (Cornblatt et al., 2011). Emotion recognition deficits thus represent promising targets for improving early identification and intervention strategies (Addington et al., 2008; Amminger et al., 2011; Cornblatt et al., 2011; Kee et al., 2003; Niemi et al., 2005). A significant challenge remains however, to identify a precursor of emotional recognition deficits for use as a salient diagnostic tool and therapeutic target.

Impairments in basic auditory processes may underlie the auditory emotion identification deficits and social functioning deficits associated with psychotic disorders and at-risk states. The auditory properties of speech convey emotional content (Banse and Scherer, 1996; Hammerschmidt and Jurgens, 2007), and the ability to identify emotion from these vocal cues is impaired in schizophrenia (Addington et al., 2008; Amminger et al., 2011; Bach et al., 2009; Corcoran et al., 2015;

Cornblatt et al., 2011; Dickey et al., 2010; Dickey et al., 2008; Kee et al., 2003; Leitman et al., 2007; Leitman et al., 2010; Niemi et al., 2005; Thompson et al., 2011), schizotypal personality disorder (Dickey et al., 2010; Dickey et al., 2008), and CHR subjects (Corcoran et al., 2015; Thompson et al., 2011). We recently reported that individuals with schizophrenia exhibit wide ranging basic auditory processing deficits, and that additional basic auditory processing abilities, including formant discrimination, sinusoidal amplitude modulation detection, and duration discrimination are associated with the ability to recognize emotion in voice (Kraus et al., 2019). In this study, we have aimed to determine whether CHR individuals exhibit similar widespread auditory processing deficits associated with impaired emotion recognition.

2. Methods and materials

2.1. Subjects

The study was approved by the Duke University institutional review board. Informed consent was obtained from all participants following full explanation of procedures. All subjects were between 12 and 30 years old and had intact hearing as demonstrated by audiometric thresholds at or below 25 dB at 500 Hz, and 20 dB at 1000 Hz, 2000 Hz, 3000 Hz or 4000 Hz. All subjects were screened to rule out a history of significant head injury, neurological or medical conditions that could

interfere with interpretation of cognitive/perception data, impaired intellectual functioning as determined by the investigator, treatment in the 8 weeks prior to assessment with any antipsychotic or mood stabilizer medication, or substance dependence with the last 6 months (excluding nicotine). CHR subjects met Psychosis-Risk Syndrome, Current Progression as per Criteria Of Psychosis-risk Syndromes (COPS) criteria, AND scored 3 or higher on P1 or P2 of the Scale of Prodromal Symptoms (SOPS). Healthy controls did not meet criteria for any COPS psychosis-risk syndrome or any current or past psychotic disorder or Cluster A personality disorder diagnoses, were not currently taking psychotropic medication and did not have any personal or familial (first degree relative) history of psychotic disorder.

2.2. Basic auditory battery

All of the auditory tasks were structured similarly, with two identical stimuli and one deviant to be detected.

2.2.1. Test of basic auditory capabilities (TBAC)

The TBAC is composed of 8 tests of auditory acuity chosen for their favorable psychometric properties and relatively specific factor loadings onto empirically determined auditory processing domains. The first three tests, *Pitch Discrimination*, *Intensity Discrimination* and *Duration Discrimination*, employ a 1 kHz, 250 ms tone as the standard and the target stimulus is varied along the dimension of interest by equal logarithmic steps. *Pulse Train* tests rhythmic change detection ability. *Embedded Tone* tests the ability to detect the presence of the fifth tone (duration varied to manipulate difficulty) in a nine tone sequence, which is absent in the standard. *Temporal Order* tests the ability to discriminate the order in which two tones are presented. *Syllable Sequence* is a speech analog to the temporal order subtest. *Nonword Recognition* assesses the ability to identify a nonsense syllable in noise.

2.2.2. Formant discrimination task

With consultation from Diane Kewley-Port, we adapted the formant discrimination task from Liu and Kewley-Port (Liu and Kewley-Port, 2004). Our modified version of the test assessed subjects' ability to distinguish first and second formant frequency shifts in two American English vowels ε ("eh") and Λ ("uh").

2.2.3. Sinusoidal amplitude modulation (SAM)

Individuals with schizophrenia have impaired physiological responses to SAM stimuli indicative of impaired phase locking due to reduced GABAergic modulation of pyramidal cell activity in auditory cortex (Teale et al., 2008). Phase locking is crucial to accurate encoding of fundamental frequency as well as speech formants (Johnson et al., 2008; Johnson et al., 2005). We therefore added a test of 60 Hz. SAM detection.

2.2.4. Pitch discrimination 6000 Hz

The frequency discrimination task that is included in the TBAC uses a standard tone of 1000 Hz – a frequency at which phase locking plays a significant role in frequency processing (Sek and Moore, 1995). As a control condition, we included another frequency discrimination task with a standard frequency of 6000 Hz – at which phase locking plays very little role in frequency processing compared to place coding signals.

2.2.5. Frequency modulation (FM)

All three tones had the same carrier wave frequency and differed only in that one tone was frequency modulated (depth manipulated to determine threshold). This experiment was run in separate blocks employing frequency modulations of 2 Hz (at which phase-locking plays a significant role in FM detection) and 10 Hz (at which FM detection depends primarily on place signaling).

2.3. Measures of auditory emotion processing

We assessed the ability to recognize emotion based on vocal affective cues in speech and non-speech affective bursts using two different tests.

2.3.1. Comprehensive affect testing system (CATS)

This battery was chosen because the American English accent of the actor seemed most appropriate for our demographic. The subtests included: emotion prosody discrimination, emotion prosody identification, and the identification of emotion from prosody in the presence of conflicting semantic meaning. In all tests, a male actor portrayed a specific emotion. Subjects were instructed to make their assessment of the emotion based solely on the sound of the actor's voice (Schaffer et al., 2006).

2.3.2. Montreal affective voices

The Montreal Affective Voices are freely available stimuli consisting of multiple male and female actors portraying various emotions through affective bursts (e.g. Chuckles, groans, sighs etc.). We developed a test of auditory emotion recognition in which subjects identified affective bursts using a 5-alternative forced-choice procedure (choosing from: anger, disgust, fear, happiness, and sadness).

2.4. Measures of cognition

We employed the following neurocognitive measures that were used in the North American Prodrome Study (NAPLS 2): WASI II Vocabulary, Matrix Reasoning, WRAT-4 Reading/Pronunciation, Hopkins Verbal Learning Test (HVLT) and BACS Symbol Coding.

2.5. Order of assessments

To minimize order effects, we divided the performance-based assessments into 2 blocks and counterbalanced the order of assessment separately in patients and controls.

2.6. Data analysis

An emotion recognition composite was constructed by first calculating Z-scores for the CATS Prosody and MAV outcome variables, and then calculating T-scores for the sum of these Z-scores. Group differences were assessed by between-groups t-tests and computation of Cohen's d. All correlations were performed in the CHR group and computed as Pearson's correlation coefficients.

3. Results

3.1. Demographics

29 individuals at CHR for psychosis and 57 healthy controls were included in the study. The CHR individuals and healthy controls did not differ significantly on demographic variables (Table 1).

3.2. Emotion recognition, cognition

Individuals at CHR for psychosis did not differ in performance from healthy controls on either the CATS or MAV assessments or an Emotion Recognition Composite constructed from the scores on the individual batteries (Table 2).

CHR individuals scored significantly lower on BACS Symbol Coding, but did not differ from healthy controls on any other cognitive measures (Table 3).

3.3. Basic auditory processing

A previous factor analysis of the basic auditory tests in this battery

Table 1

Demographics.

	CHR			Controls				
	N	Mean	S.D.	N	Mean	S.D.	t	p
Age	29	21.59	3.52	57	22.72	4.03	−1.34	0.18
Education (years)	29	14.41	2.37	57	14.58	2.49	−0.30	0.76
Maternal Education	29	15.10	2.60	54	15.22	3.04	−0.19	0.85
Paternal Education	29	16.00	2.78	53	15.25	3.20	1.11	0.27
Gender								
Male	9	31.03%		18	31.58%			
Female	20	68.97%		39	68.42%			
Race								
White	21	72.41%		33	57.89%			
African-American	4	13.79%		12	21.05%			
Other	4	13.79%		12	21.05%			

Table 2

Group Means on Emotion Recognition Measures. For all measures, higher scores indicated better performance. Cohen's d is expressed as CHR performance compared to that of controls. Thus, a positive value for Cohen's d indicates the CHR group performed better. CATS: Comprehensive Affect Testing System Prosody Scale, MAV: Montreal Affective Voices, Emotion Composite: Composite score of CATS and MAV.

Measure	CHR		Controls		
	Mean	SD	Mean	SD	Cohen's d
CATS	91.18	4.39	90.65	6.07	0.10
MAV	48.67	3.75	47.61	3.69	0.28
Emotion Comp.	51.59	9.08	48.73	10.38	0.29

Table 3

Group Means on Cognitive Measures. All measures are reported as raw scores, except for WASI FSIQ-2. Cohen's d is expressed as CHR performance compared to that of controls. Thus, a positive value for Cohen's d indicates the CHR group performed better.

	CHR		Controls		
Measure	Mean	SD	Mean	SD	Cohen's d
BACS symbol coding	60.21*	12.15	66.09	12.06	−0.49
HVLT	28.61	4.90	29.54	3.85	−0.21
WRAT	64.64	4.21	63.15	5.69	0.30
WASI Vocab	42.89	4.52	42.63	5.29	0.05
WASI Matrix Reasoning	22.93	2.69	23.31	2.49	−0.15
WASI FSIQ-2	113.96	11.69	114.52	13.06	−0.04

* $p < .05$.

indicated a single factor solution and thus all analyses were performed at the level of the individual tests (Kraus et al., 2019). CHR Individuals performed significantly worse on the TBAC subtest of duration discrimination (Table 4). Although the effect size of this difference was d = 0.61, it did not survive Bonferroni correction for multiple comparisons. The groups did not differ significantly on any of the other measures of auditory discrimination.

3.4. Correlation of basic auditory processing with emotion recognition and cognition

Within the CHR sample, the emotion recognition composite score was significantly correlated with the basic auditory skills of intensity discrimination (0.59), syllable order (0.44) and syllable recognition (0.56) (Table 5). Cognitive measures were also significantly correlated to several basic auditory skills (Table 5) and to the emotion recognition composite (Table 6).

Table 4

Group Means on Basic Auditory Skills. SAM 60: Sinusoidal Amplitude Modulation Detection at 60 Hz, FM 2 Hz: Frequency Modulation Detection at 2 Hz, FM 10 Hz: Frequency Modulation Detection at 10 Hz. For SAM 60 Hz, Pitch Discrimination (6000 Hz), FM 2 Hz and FM 10 Hz scores are reported as thresholds for discrimination and thus lower scores indicate better performance. These tests have been italicized in the table below. For all other tests, scores reflect number of correct responses, and thus higher scores are better. For the measures reported as thresholds, the value of Cohen's d was reversed so that for all tests, a positive Cohen's d indicates better performance in the CHR group compared to the control group.

Measure	CHR		Controls		
	Mean	SD	Mean	SD	Cohen's d
Frequency Disc.	72.31	25.07	61.61	29.76	0.39
Intensity Disc.	58.08	28.31	58.35	27.90	−0.01
Duration Disc.	48.65*	30.35	65.93	25.90	−0.61
Pulse-Train	46.81	29.25	55.56	29.89	−0.30
Embedded Tone	60.46	32.39	58.07	30.06	0.08
Temporal Order	44.69	32.14	48.11	30.61	−0.11
Syllable Order	35.89	26.83	42.56	30.71	−0.23
Syllable Recognition	68.46	26.81	61.09	24.83	0.29
Formant Disc.	49.47	9.61	50.18	10.27	−0.07
SAM60	−31.44	7.30	−31.16	5.84	0.04
Pitch Disc. (6000 Hz)	57.33	95.32	121.61	207.02	0.34
FM 2 Hz	1.43	0.57	1.96	1.94	0.13
FM 10 Hz	0.41	0.68	0.32	0.15	−0.35

* $p < .05$.

4. Discussion

Previous research has indicated widespread impairment of basic auditory skills in schizophrenia that is associated with difficulty understanding emotion from the vocal characteristics of a speaker (Kraus et al., 2019). We investigated whether these impairments are present in individuals at high-risk for psychosis. We found no indication of impaired emotion recognition and minimal impairment of basic auditory skills in high-risk individuals that was confined to a deficit in duration discrimination. Within the CHR group, duration discrimination was not associated with emotion recognition performance. While there were additional group differences of small effect sizes that may have become significant with a larger sample, several of these differences were in the direction of superior performance in the CHR group. This finding of intact auditory emotion recognition and minimally impaired basic auditory skills is in line with that from Corcoran et al. (2015) who did not find deficits in auditory emotion recognition or tone matching in a sample of 27 CHR individuals who did not convert to psychosis within 2 years, but did find these deficits in 2 CHR individuals who did convert to psychosis within 2 years. Follow-up with our CHR sample was not within the scope of this study and thus we are unable to report on conversion rate or compare converters to non-converters on our measures. However, our results suggest that auditory discrimination and

Table 5

Correlations within the CHR sample of basic auditory skills with emotion recognition and cognitive measures. SAM 60: Sinusoidal Amplitude Modulation Detection at 60 Hz, FM 2 Hz: Frequency Modulation Detection at 2 Hz, FM 10 Hz: Frequency Modulation Detection at 10 Hz. SAM 60 Hz, Pitch Discrimination (6000 Hz), FM 2 Hz and FM 10 Hz scores are thresholds for discrimination and thus negative correlations with emotion recognition and cognitive measures indicate that better performance on these tests of auditory discrimination was associated with better performance on the emotion recognition or cognitive measure. These tests have been italicized in the table below.

	CATS	MAV	Emotion Composite	BACS Symbol Codiing	HVLT	WRAT	WASI Vocab	WASI Matrix Reasoning	WASI IQ Estimate
Frequency Disc.	0.03	0.31	0.23	0.32	0.04	0.38*	0.50**	0.29	0.47*
Intensity Disc.	0.22	0.68**	0.59**	0.32	0.36	0.56**	0.67**	0.42*	0.74**
Duration Disc.	0.04	0.17	0.14	0.04	0.04	−0.07	0.04	0.12	0.19
Pulse-Train Disc.	−0.03	0.12	0.06	0.08	−0.11	0.00	−0.13	0.03	−0.03
Embedded Tone	0.15	0.17	0.20	−0.11	−0.09	0.21	0.09	0.26	0.19
Temporal Order	0.17	0.37	0.35	0.37	0.22	0.41*	0.51**	0.30	0.50**
Syllable Order	0.26	0.44*	0.44*	0.35	0.40*	0.24	0.52**	−0.24	0.28
Syllable Recognition	0.32	0.56**	0.56**	0.15	0.50*	0.64**	0.71**	0.18	0.55**
Formant Disc.	−0.02	0.36	0.23	0.25	0.34	0.37	0.52**	0.45*	0.56**
SAM 60	−0.21	−0.33	−0.34	−0.24	−0.32	−0.27	−0.08	0.01	−0.12
Pitch Disc. (6000 Hz)	−0.08	−0.16	−0.15	−0.46*	−0.06	−0.30	−0.35	−0.25	−0.41*
FM 2 Hz	0.05	−0.32	−0.19	−0.06	−0.31	−0.22	−0.37	−0.04	−0.20
FM 10 Hz	0.04	−0.18	−0.10	−0.05	−0.24	−0.10	−0.05	−0.15	0.05

* $p < .05$.
** $p < .01$.

Table 6

Correlations within the CHR sample between cognitive and emotion recognition measures.

	BACS Symbol Codiing	HVLT	WRAT	WASI Vocab	WASI Matrix Reaoning	WASI FSIQ-2
CATS	−0.17	0.14	0.48**	0.04	−0.05	0.02
MAV	0.16	0.56**	0.53**	0.44*	0.21	0.40*
Emotion composite	0.02	0.47*	0.64**	0.33	0.12	0.29

* $p < .05$.
** $p < .01$.

emotion recognition impairment are not characteristic of high-risk status and thus may emerge near psychosis onset only in individuals who later go on to convert.

While the auditory tasks chosen for this study had previously demonstrated sensitivity to auditory impairments in individuals with schizophrenia, these tasks did not register impairments in our CHR group. It is possible that more sensitive tests could pick up subtle auditory processing deficits in individuals at CHR. It should also be noted that our CHR group exhibited impairment on BACS Symbol Coding performance, but were not impaired on any other cognitive outcomes. This differs from previous findings of more widespread cognitive deficits in CHR individuals (Seidman et al., 2016). Thus, it is possible that our results, based on a small sample of CHR individuals, may not reflect the extent of auditory processing deficits that would be seen in a larger sample.

CRediT authorship contribution statement

Michael Kraus Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Visualization, Project Administration. **Trina Walker**: Methodology, Investigation, Writing – Review & Editing. **Diana Perkins**: Conceptualization, Methodology, Investigation, Writing – Review & Editing, Supervision. **Richard Keefe**: Conceptualization, Methodology, Investigation, Resources, Writing – Review & Editing, Supervision, Funding acquisition.

Declaration of competing interest

Michael Kraus has no conflicts to report.
Trina Walker reports no conflicts of interest.
Diana Perkins reports no conflicts of interest.
Richard S.E. Keefe is the owner of VeraSci, a for-profit company that provides clinical trials support and related services for over 100 business entities, primarily pharmaceutical companies.

References

Addington, J., Penn, D., Woods, S.W., Addington, D., Perkins, D.O., 2008. Facial affect recognition in individuals at clinical high risk for psychosis. Br. J. Psychiatry 192 (1), 67–68.

Amminger, G.P., Schäfer, M.R., Papageorgiou, K., et al., 2011. Emotion recognition in individuals at clinical high-risk for schizophrenia. Schizophr. Bull. 38 (5), 1030–1039.

Bach, D.R., Buxtorf, K., Grandjean, D., Strik, W.K., 2009. The influence of emotion clarity on emotional prosody identification in paranoid schizophrenia. Psychol. Med. 39 (06), 927–938.

Banse, R., Scherer, K.R., 1996. Acoustic profiles in vocal emotion expression. J. Pers. Soc. Psychol. 70 (3), 614–636.

Corcoran, C.M., Keilp, J.G., Kayser, J., et al., 2015. Emotion recognition deficits as predictors of transition in individuals at clinical high risk for schizophrenia: a neurodevelopmental perspective. Psychol. Med. 45 (14), 2959–2973.

Cornblatt, B.A., Carrión, R.E., Addington, J., et al., 2011. Risk factors for psychosis: impaired social and role functioning. Schizophr. Bull. 38 (6), 1247–1257.

Dickey, C.C., Morocz, I.A., Niznikiewicz, M.A., et al., 2008. Auditory processing abnormalities in schizotypal personality disorder: an fMRI experiment using tones of deviant pitch and duration. Schizophr. Res. 103 (1–3), 26–39.

Dickey, C.C., Morocz, I.A., Minney, D., et al., 2010. Factors in sensory processing of prosody in schizotypal personality disorder: an fMRI experiment. Schizophr. Res. 121 (1–3), 75–89.

Hammerschmidt, K., Jurgens, U., 2007. Acoustical correlates of affective prosody. J. Voice 21 (5), 531–540.

Johnson, K.L., Nicol, T.G., Kraus, N., 2005. Brain stem response to speech: a biological marker of auditory processing. Ear Hear. 26 (5), 424–434.

Johnson, K.L., Nicol, T., Zecker, S.G., Bradlow, A.R., Skoe, E., Kraus, N., 2008. Brainstem encoding of voiced consonant–vowel stop syllables. Clin. Neurophysiol. 119 (11), 2623–2635.

Kee, K.S., Green, M.F., Mintz, J., Brekke, J.S., 2003. Is emotion processing a predictor of functional outcome in schizophrenia? Schizophr. Bull. 29 (3), 487–497.

Kraus, M.S., Walker, T.M., Jarskog, L.F., Millet, R.A., Keefe, R.S.E., 2019. Basic auditory processing deficits and their association with auditory emotion recognition in schizophrenia. Schizophr. Res. 204, 155–161.

Leitman, D.I., Hoptman, M.J., Foxe, J.J., et al., 2007. The neural substrates of impaired prosodic detection in schizophrenia and its sensorial antecedents. Am. J. Psychiatry 164 (3), 474–482.

Leitman, D.I., Laukka, P., Juslin, P.N., Saccente, E., Butler, P., Javitt, D.C., 2010. Getting the cue: sensory contributions to auditory emotion recognition impairments in schizophrenia. Schizophr. Bull. 36 (3), 545–556.

Liu, C., Kewley-Port, D., 2004. Vowel formant discrimination for high-fidelity speech. J. Acoust. Soc. Am. 116 (2), 1224–1233.

Niemi, L.T., Suvisaari, J.M., Haukka, J.K., Lönnqvist, J.K., 2005. Childhood predictors of future psychiatric morbidity in offspring of mothers with psychotic disorder. Br. J. Psychiatry 186 (2), 108–114.

Schaffer, S.G., Froming, K.B., Gregory, A.L., Levy, C.M., Ekman, P., 2006. Emotion Processing: The Comprehensive Affective Testing System User's Manual. Psychology Software Inc.

Seidman, L.J., Shapiro, D.I., Stone, W.S., et al., 2016. Association of neurocognition with transition to psychosis: baseline functioning in the second phase of the North American Prodrome Longitudinal Study. JAMA Psychiatry. 73 (12), 1239–1248.

Sek, A., Moore, B.C., 1995. Frequency discrimination as a function of frequency, measured in several ways. J. Acoust. Soc. Am. 97 (4), 2479–2486.

Teale, P., Collins, D., Maharajh, K., Rojas, D.C., Kronberg, E., Reite, M., 2008. Cortical source estimates of gamma band amplitude and phase are different in schizophrenia. NeuroImage 42 (4), 1481–1489.

Thompson, A.D., Bartholomeusz, C., Yung, A.R., 2011. Social cognition deficits and the 'ultra high risk' for psychosis population: a review of literature. Early Interv. Psychiatry 5 (3), 192–202.