The Spanish Journal of Psychology (2020), **23**, e52, 1–7.

doi:10.1017/SJP.2020.53

Altered Auditory Self-recognition in People with Schizophrenia

Chaochao C. Pan and Aibao B. Zhou

Northwest Normal University of Anning District (China)

Abstract. Self-recognition is of great significance to our sense of self. To date, disturbances in the processing of visual self-recognition are well studied in people with schizophrenia, whereas relatively few studies have focused on the processing of self in other domains, such as auditory. An investigation of auditory self-recognition contributes to delineate changes related to self and the potential roots of the described psychopathological aspects connoting schizophrenia. By applying unimodal task and multisensory test, this study investigated auditory self-recognition in people with schizophrenia under unimodal and bimodal circumstances. Forty-six adults diagnosed with schizophrenia and thirty-two healthy controls were involved in this study. Results suggested that people with schizophrenia seemed to have significantly lower perceptual sensitivity in detecting self-voice, and also showed stricter judgment criteria in self-voice decision. Furthermore, in the presentation of stimuli that combined the stimulation of others' faces with one's own voice, people with schizophrenia mistakenly attributed the voices of others as their own. In conclusion, altered auditory self-recognition in people with schizophrenia was found.

Received 18 October 2019; Revised 23 October 2020; Accepted 23 October 2020

Keywords: auditory, bimodal, multisensory integration, self-disturbance

Schizophrenia (SZ) is a complicated mental disorder characterized by negative and positive symptoms, formal thought disorders, and significantly damaged social function (Kronbichler et al., 2018). Epidemiological studies showed that the disorder affects 1% of the population worldwide (Insel, 2010; Javitt & Sweet, 2015). Mental disability is closely associated with SZ due to protracted course of the disorder, which brought heavy burdens to society and families. It is commonly agreed that SZ is strictly associated with changes of self (Sass, 2014; Zhang et al., 2012). According to ipseity-disturbance model (IDM) proposed by Sass (2014), SZ is fundamentally a kind of self-disorder or ipseity-disturbance. The distortion or instability of the self is composed of two aspects: Hyper-reflexivity and diminishment of self-affection. Hyper-reflexivity refers to a sort of exaggerated self-consciousness in which normally background or tacit phenomena become focal points of attention. Diminished self-affection complements hyper-reflexivity and refers to degradation in the experience of self as an agent or effective subject of awareness (Hamm et al., 2018). IDM also suggests that seemingly different symptoms can be clear manifestations of the disturbed self (Nelson et al., 2014; van der Weiden et al., 2015). This skewed sense of self, which is linked to impairments of self-recognition (Ferroni et al., 2019; Gallese & Ferri, 2014), have been closely related to SZ pathophysiology and also seems to be associated with deficits in multisensory integration mechanisms (Ferroni et al., 2019).

For most people with SZ, self-disorder is elusive and difficult to verbalize (Sandsten et al., 2019). Various disturbances, such as altered stream of consciousness, depersonalization, or warped self-boundary, typically characterize this kind of disorder in phenomenological observations. It also has been emphasized that pre-psychotic disturbances to the self is a specific and constitutive trait of this disorder (Sandsten et al., 2019).

Based on the phenomenologically-oriented observations, researchers have gradually adopted experimental approaches to investigate self-disorders (Ferroni et al., 2019; Sass, 2014; She et al., 2017). An important contribution to this realm of research comes from the study of self-recognition, which is the developmental basis for

Correspondence concerning this article should be addressed to Chaochao Pan. Northwest Normal University of Anning District. School of Psychology. Lanzhou City, Gansu Province. 730070 Lanzhou (China).

E-mail: 18394797386@163.com

Conflicts of Interest: None.

Funding Statement: This work was supported by the Ethnic Research Project of the State People's Committee Foundation of China (C. Pan, grant number 2019-GMD-020); the Philosophy and social science planning of Gansu Province in 2019 (C. Pan, grant number 19YB074); and the Research funding project for Postgraduates of Northwest Normal University (C. Pan, grant number 2019KYZZ021006).

How to cite this article:
Pan, C. C., & Zhou, A. B. (2020). Altered auditory self-recognition in people with schizophrenia. *The Spanish Journal of Psychology, 23*. e52. Doi:10.1017/SJP.2020.53

self-identity and of great significance to the structured sense of self. Self-recognition has been mostly investigated by the studies of self-face recognition using unimodal task and multisensory illusion test. For example, some behavior studies applying visual search task, in which participants were instructed to detect a target (i.e., self-face) among distractors (i.e., distractor face images) in sets of two, four or eight faces, have investigated self-face recognition in vision domain. Results have found that altered visual self-recognition in people with SZ (Bortolon et al., 2016; Lee et al., 2007). Numerous neurophysiological studies have also found abnormalities in anatomical structures and functional activities related to face-processing (Lee et al., 2002; She et al., 2017). Interestingly, the multisensory paradigm of the enfacement illusion (EI) was introduced by Tsakiris (2008) to show how multisensory clues could affect self-face recognition (Ferroni et al., 2019; Sandsten et al., 2019). EI is a subjective experience of perceiving other's face as self-face under synchronous interpersonal multisensory stimulation. EI effect is quantitatively evaluated by the performance on self-face recognition arranged before and after the synchronous multisensory stimulation. Although at present under debate, people with SZ seem to show significantly altered self-other discrimination in multisensory paradigm of the EI, suggesting that deviations existed in their facial self-recognition.

Crucially, self-recognition is not only related to face but also to the voice, which represents a wealth of socially-relevant information, including paralinguistic clues to a speaker's identity and emotion (Chhabra et al., 2012). Deficits recognizing emotion from voice have been widely described, but relatively little attention has been given to investigating the ability to analyze speaker identity in people with SZ (Chhabra et al., 2012; Hoekert, et al., 2007). In a study by Chhabra et al. (2012), voice-discrimination was investigated based on acoustic characteristics, average fundamental frequency (F0), and resonance (Df), related to voice-identity perception (Baumann & Belin, 2010). In this study, participants heard unfamiliar voices of two speakers saying the same words in sequence, the participants were then instructed to rate the degree of perceived identity similarity on a seven-point rating scale ranging from *very similar* (1) to *very dissimilar* (7). The results showed that there was no significant difference in reliance on F0 between people with SZ and healthy controls in differentiating voices, suggesting that the ability to discriminate voice identity based on pitch clues may be relatively preserved in people with SZ. On the other hand, people with SZ were less dependent on Df to differentiate voices. Different response patterns indicated that there were some potentially important differences in voice-identity processing in SZ. Alba-Ferrara et al. (2012) examined the recognition of famous voices in people with SZ. The results showed that the people with SZ obtained lower hit rates and showed lower perceptual sensitivity in detecting famous voices. Recently, a line of study extended the multisensory integration procedure applied in the McGurk effect paradigm to voice-recognition (Peynircioğlu et al., 2017). The traditional McGurk effect firstly and mostly demonstrated in the field of speech. This effect is induced by the observation of motion picture uttering a plosive such as /ga/ at the same time of auditory perception hearing a different plosive such as /da/ (McGurk & MacDonald, 1976). Synchronous, but mismatched, audio-visual stimulation between two different syllables alters usual auditory perception, mishearing it toward the visual stimulation. Interestingly, a robust non-speech McGurk effect was also demonstrated in gender identification. In the study by Peynircioğlu et al. (2017), gender judgment in voice was investigated, using voice timbre and visual gender information. The results suggested that visual gender information overrode the voice gender identification, showing a robust non-speech McGurk effect.

Compared with visual self-recognition, relatively little attention has been given to examining auditory self-recognition. It is crucial to further investigate self-recognition in auditory, which contributes to delineate changes related to self and the potential roots of the described psychopathological aspects connoting SZ. Based on the above considerations, we decided to investigate auditory self-recognition under unimodal and bimodal circumstances. In the unimodal auditory self-recognition task, we used the voice-recognition paradigm that Alba-Ferrara et al. (2012) adopted. The McGurk effect paradigm was applied to bimodal self-recognition task. We hypothesized that people with SZ would have altered auditory self-recognition by being less sensitive to self-voice and hearing more their own voices as the voices of others.

Method

Participants

Seventy-eight adults participated in this study. Forty-six schizophrenic inpatients were recruited from Tianshui Psychiatric Hospital, Gansu Province, China. All patients met the diagnostic criteria for SZ in the DSM–IV, as determined by two experienced psychiatrists. During the experiment, all patients received atypical antipsychotic drugs and were a stable phase of recovery. The exclusion criteria were (a) having received electric shock treatment, (b) current drug abuse, (c) a history of brain injury, (d) comorbid neurological disorders, and (e) mental retardation. The clinical symptoms of SZ were evaluated using the Scale for the Assessment of Positive Symptoms (SAPS) and the Scale for the Assessment of Negative Symptoms (SANS). Additionally, thirty-two adults were recruited through advertisements to serve as healthy controls. They did not suffer from mental illness and had no family history of mental illness. All participants were recruited between May 2019 and September 2019. The people with SZ and healthy controls were matched in age, gender and educational level. The groups' demographic characteristics being compared using *t* tests and *chi-square* tests, and the groups were statistically equal. The specific demographic information was shown in Table 1.

All participants had normal vision and hearing. After description of the procedure, written informed consent was obtained from all participants. They were compensated with $3 for their participation. This study was approved by the Institutional Review Board of Tianshui Psychiatric Hospital and Northwest Normal University.

Material

Firstly, the meaningless monosyllable "a" was chosen as the auditory material. Secondly, each participant was uniformly trained to utter this syllable with standardized time, tone, loudness and rhythm. Finally, applying the Nikon–D7200 SLR camera to shoot video containing the entire pronunciation at a speed of 60 frames per second (1280 × 720 pixel resolution) before starting the experiment. The camera was 70 cm away from the participants during the shooting. The participants were asked to lean back against the white background and look directly at the camera in a neutral expression. After hearing the "start" instruction, participants spoke with clear lip movement. The pronunciation lasted about 1,700 ms, and then the natural end of the pronunciation was expressed by body language with closed lips. The participants' vocal materials were also collected at the same time as the video shooting, and the recording was collected with an aigo–R8611 recorder.

GoldWave software was applied to standardize the original audio materials. Audio tracks were opened with GoldWave and digitized with 16-bit sampling resolution and 44.1 kHz sampling rate, no adjustment applied to acoustic parameters, such as fundamental frequency and formant frequency, and the duration of the audio was uniformly adjusted to 1,700 ms. The processed audio was saved as stereo finally.

Adobe Premiere software was used to synthesize dynamic face and voice material according to experimental conditions. In synthesis, cancelling video soundtrack, and making the pronunciation of syllable coincided with the articulation of dynamic face. An unfamiliar other was randomly selected from video sets that consisted of all participants, which matched the self-face in age and gender. The synthesized videos lasted about 1,700 ms. According to experimental conditions, there were two types of videos: The participant's own face and voice (S_{face} + S_{voice}), other's face combined with their own voice (O_{face} + S_{voice}).

Procedure

The experiment was conducted in a silent room. All materials were presented on the computer using E-prime 2.0 software. The participants sat 60 cm away from the computer screen to complete two auditory self-recognition tasks. Prior to the formal experiment, participants carried out a practice experiment confirming that they understood the instructions correctly.

Participants completed two separate blocks investigating auditory self-recognition under unimodal and bimodal circumstances respectively. The order of the two blocks was balanced between participants and there was a 5 min break between blocks. In the unimodal auditory self-recognition task, participants were presented with two types of voice, self-voice and non-self voice. A non-self voice was randomly selected from voice sets that consisted of all participants' voices, which matched the self-voice in age and gender. There were 60 trials in this task, 30 trials of which were self-

Table 1. *Demographic Information*

	Schizophrenia patient sample (*N* = 48)	Healthy control sample (*N* = 32)	t/χ^2	*p*
Gender(male/female)	23/23	15/17	.07	.79
Age	32.65(8.89)	32.71(11.03)	.00	.98
Education (year)	8.61(1.74)	9.07(4.86)	.35	.56
Illness duration(year)	6.67(5.33)			
SAPS	34.91(23.67)			
SANS	28.98(21.16)			
Drug use (mg)	479.40(230.60)			

Note. The values in the brackets are standard deviations (*SD*); Antipsychotics drug use equivalent to Chlorpromazine dosage.

https://doi.org/10.1017/SJP.2020.53 Published online by Cambridge University Press

voice. The order of voice presentation was random. In each trial, a fixed red cross was presented in the center of the screen for 500 ms, followed by a voice for 3,000 ms. When the voice was presented, participants were required to judge whether the voice presented was their own or not quickly and accurately by pressing one of two keys (half of the participants responded "self" by pressing "J" with their right hand and "non-self" by pressing "F" with left hand, the other half of the participants responded "self" by pressing "F" on keyboard with their left hand and "non-self" by pressing "J" with right hand). An empty screen was presented as a buffer for 500ms after a response was submitted by the keyboard and the next trial began automatically.

In the bimodal self-recognition task, participants were presented with two types of videos, which were combinations of self-face with self-voice and of other's face with self-voice respectively. For video type, combination of self-face with self-voice was matched, while mismatched condition presented the combination of other's face with self-voice. The entire experiment consisted of 60 trials, 30 trials of which were matched type of video. Each trial began with a fixed red cross presented in the center of the screen for 500 ms. After the fixation, randomly followed by a video for 3,000 ms, participants were instructed to judge whether the speaker's voice in the video was themselves or not. The participants quickly and accurately responded by pressing one of two response keys (half of the participants responded "self" by pressing "J" with their right hand and "non-self" by pressing "F" with left hand, the other half of the participants responded "self" by pressing "F" on keyboard with their left hand and "non-self" by pressing "J" with right hand). An empty screen was presented for 500 ms after a response was made, and the next trial began automatically.

Data analysis

Sensitivity (d') and Response Bias (β) were analyzed as indicators of auditory self-recognition performance. To calculate the d', P(H) (the hitting rate) and P(FA) (false alarm rate) of each participant were calculated firstly, then the corresponding Z-value was obtained by checking the P-Z-O conversion table, and the $d'$was obtained by $d' = Z_H – Z_{FA}$. The smaller the d' is, the less the sensitivity is. Similarly, to calculate the β, P(H) and P(FA) of each participant were calculated, then the corresponding O-value was obtained by checking the P-Z-O table, and the β was obtained by $\beta = O_H/O_{FA}$. $\beta > 1$, indicating that the judgment criteria were strict.

Results

Independent t-test was performed separately to d' and β for unimodal auditory self-recognition. The results showed that there were significant difference between groups in d', $t(76) = 4.34$, $p < .01$, people with SZ (M = 2.43, SD = 1.70) were significantly less sensitive to self-voice than healthy controls (M = 3.80, SD = .66). Response bias showed a pattern similar to sensitivity, difference between groups in β approached significant, $t(76) = 1.90$, $p = .06$, people with SZ (M = 2.62, SD = 3.24) were more prudent than healthy controls (M = 1.41, SD = 1.68) in self-voice decision making.

Table 2 showed the mean values of d' and β for auditory self-recognition in bimodal circumstances. A 2 (Group: Schizophrenia patient, healthy control) × 2 (video type: Matched, mismatched) mixed ANOVA was performed on d' and β respectively. The results of the d' showed a significant interaction effect of group × video type, $F(1, 76) = 14.22$, $p < .001$, $\eta^2 = .20$. Post hoc analysis confirmed that in the people with SZ, sensitivity to self-voice was significantly higher in the matched than mismatched type of video, $F(1, 76) = 63.52$, $p < .001$, $\eta^2 = .53$. There was also a similar processing pattern for the healthy control, in which the sensitivity of the matched was significantly higher than mismatched type of video, $F(1, 76) = 4.13$, $p = .05$, $\eta^2 = .07$ (the result was showed in Figure 1). Additionally, the result showed a significant main effect of the video type, $F(1, 76) = 46.38$, $p < .001$, $\eta^2 = .45$, participants were more sensitive to self-voice in the combinations of self-face with self-voice than of other's face with self-voice. The main effect of group was also significant, $F(1, 76) = 39.37$, $p < .001$,

Table 2. *Descriptive Statistics of Two Groups in Bimodal Auditory Self-recognition Task*

Group	Video type	$M_{d'}(SD)$	$M_\beta(SD)$
Schizophrenia sample	S_{face}+S_{voice}	1.81(.98)	.88(1.45)
(N = 48)	O_{face}+S_{voice}	–.57(1.61)	.99(1.77)
Healthy control sample	S_{face}+S_{voice}	2.64(1.20)	1.92(2.95)
(N = 32)	O_{face}+S_{voice}	1.95(1.43)	2.48(3.39)

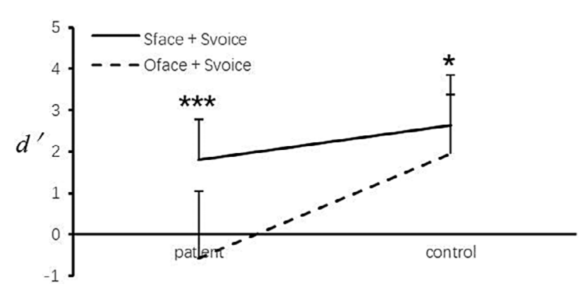

Figure 1. The Interaction Effect of Group × Video Type in d'

https://doi.org/10.1017/SJP.2020.53 Published online by Cambridge University Press

$\eta^2 = .41$, the sensitivity of the people with SZ to self-voice was significantly lower than that of the healthy control. The ANOVAs of the response bias showed a significant main effect of the group, $F(1, 76) = 5.40$, $p < .05$, $\eta^2 = .09$, and the main effect of the video type and the interaction effect of group × video type were not significant, $F(1, 76) = 1.03$, $p = .32$; $F(1, 76) = .48$, $p = .49$.

Discussion

The present study investigated the auditory self-recognition in people with SZ. To achieve this goal, the classical unimodal and bimodal voice-recognition paradigms were applied to test the potential changes of the auditory self. The result of unimodal showed that auditory self-recognition in people with SZ was altered. Specifically, the people with SZ seemed to have significantly lower perceptual sensitivity in detecting self-voice than healthy controls, and also showed stricter judgment criteria in self-voice decision. The result of bimodal provided further evidence for altered auditory self-recognition in people with SZ. Interestingly, we also found that in combinations of self-face with self-voice, sensitivity to self-voice was significantly higher than of other's face with self-voice, especially in people with SZ. This indicated that under the influence of other's face, people with SZ heard more their own voices as the voices of others. Overall, these results confirm that auditory self-recognition was altered in people with SZ.

Disturbances in the processing of visual self-recognition are well studied in people with SZ (Ferroni et al., 2019; Sandsten et al., 2019; She et al., 2017), whereas relatively few studies have focused on the processing of self in other domains, such as auditory. On the one side, it was well-known that hearing one's own voice recording from an outside source may sound dissimilar than hearing one's voice internally (i.e., there is distortion that occurs from bone conduction when perceiving one's own voice as it is produced during normal speech) (Hughes & Nicholson, 2010). We have no well way to properly solve the current difficulties on the basis of the current state of art. Furthermore, acoustic parameters relating to voice perception, such as fundamental frequency and formant frequency, must be considered in auditory self-recognition. In this study, we applied GoldWave software to prevent these parameters from confusing auditory self-recognition. The study of auditory self-recognition not only contributes to delineate changes related to self in non-visual domain but also provides further evidence for the notion that potential roots of the pathophysiological aspects relating to SZ are self-disorder.

The study of unimodal auditory self-recognition found significantly difference between groups with regard to perceptual sensitivity in detecting self-voice. Similar to other studies (Alba-Ferrara et al., 2012; Pinheiro et al., 2017; van der Weiden et al., 2015), the people with SZ were less sensitive to their own voice. Importantly, we also found that people with SZ were more prudent than healthy controls in self-voice decision making. This may be a distinct feature of auditory self-recognition in people with SZ. It is possible that the people with SZ have a stronger self-defense mechanism in auditory. As mentioned above, hearing one's own voice recording from an outside source may sound dissimilar than hearing one's voice internally. The discrepancies may result in mismatch between self-voice presented currently and auditory self-image, which activates the defense mechanism and makes a defensive negation to one's own voice. Additionally, it is important to note that, according to several fundamental studies (Chhabra et al., 2012; Javitt & Sweet, 2015), the perception of basically acoustic characteristics is degraded in people with SZ, which subsequently makes it difficult to recognize the internal psychological features such as emotion and identity carried by voice.

The results of bimodal confirmed the fact that altered auditory self-recognition in people with SZ. Additionally, we also demonstrated that in combinations of self-face with self-voice, sensitivity to self-voice was significantly higher than of other's face with self-voice, especially in people with SZ. This indicated that the processing of voice is more susceptible to the interference of mismatched face. It is the fact that auditory input lags behind visual input (about 150 ms), visual content can predict the auditory information (Arnal et al., 2009). When self-face combined with self-voice, visual predictability promotes self-voice recognition. If not, the error between face and voice have a negative effect on the performance of anditory self-recognition, and result in hearing more their own voices combined with other's face as the voices of others. Furthermore, we can't ignore the effect of psychotic symptoms/features in people with SZ, especially auditory verbal hallucinations (AVH), on auditory self-recognition. Previous studies have shown that AVH might be one's own speech. The activation of Broca's area, which can produce, but does not listen to speech, has been found in people with SZ having AVH (Asai et al., 2011). Therefore, these people might produce speech, but not think that it is they who actually spoke. As a result, they may hear their own voices as the voices of others.

Some limitations of present study should be highlighted. Firstly, recent studies have suggested that healthy general population might experience the psychotic symptoms (i.e., AVH like experiences), called schizotypal personality, indicating a continuum of the schizophrenia. The so-called schizotypal personality can be measured by schizotypal personality questionnaire. Only people with SZ in stable phase of recovery were involved in this study, further studies should investigate whether changes in anditory self-recognition observed from this study are a heritable endophenotypic marker of SZ from the perspective of schizophrenia spectrum disorder. Secondly, schizophrenia is a typical heterogeneous mental disorder. Various clinical subtypes, such as the paranoid-type schizophrenia, who have prominent positive symptoms, and the residual-type schizophrenia, with predominant negative symptoms, need to be investigated respectively in further studies. Lastly, there are differences between laboratory settings and actual environment in auditory self-recognition, and the differences should be considered.

In conclusion, we respectively investigated the auditory self-recognition in people with SZ under unimodal and bimodal circumstances. The results consistently showed that auditory self-recognition in people with SZ was altered. The present study extends the self-recognition of SZ to the auditory domain, which contributes to delineate changes related to self and the potential roots of the described psychopathological aspects connoting SZ.

References

Alba-Ferrara, L., Weis, S., Damjanovic, L., Rowett, M., & Hausmann, M. (2012). Voice identity recognition failure in patients with schizophrenia. *The Journal of Nervous and Mental Disease*, *200*, 784–790. http://doi.org/10.1097/NMD.0b013e318266f835

Arnal, L. H., Morillon, B., Kell, C. A., & Giraud, A.-L. (2009). Dual neural routing of visual facilitation in speech processing. *Journal of Neuroscience*, *29*, 13445–13453. http://doi.org/10.1523/JNEUROSCI.3194-09.2009

Asai, T., Sgimori, E., & Tanno, Y. (2011). *Hallucinations: Types, stages and treatments*. Nova Science Publishers.

Baumann, O., & Belin, P. (2010). Perceptual scaling of voice identity: Common dimensions for different vowels and speakers. *Psychological Research*, *74*, 110–120. http://doi.org/10.1007/s00426-008-0185-z

Bortolon, C., Capdevielle, D., Salesse, R. N., & Raffard, S. (2016). Further insight into self-face recognition in schizophrenia patients: Why ambiguity matters. *Journal of Behavior Therapy and Experimental Psychiatry*, *50*, 215–222. http://doi.org/10.1016/j.jbtep.2015.09.006

Chhabra, S., Badcock, J. C., Maybery, M. T., & Leung, D. (2012). Voice identity discrimination in schizophrenia. *Neuropsychologia*, *50*, 2730–2735. http://doi.org/10.1016/j.neuropsychologia.2012.08.006

Ferroni, F., Ardizzi, M., Sestito, M., Lucarini, V., Daniel, B. D., Paraboschi, F., Tonna, M., Marchesi, C., & Gallese, V. (2019). Shared multisensory experience affects others' boundary: The enfacement illusion in schizophrenia. *Schizophrenia Research*, *206*, 225–235. http://doi.org/10.1016/j.schres.2018.11.018

Gallese, V., & Ferri, F. (2014). Psychopathology of the bodily self and the brain: The case of schizophrenia. *Psychopathology*, *47*, 357–364. http://doi.org/10.1159/000365638

Hamm, J. A., Leonhardt, B. L., Ridenour, J., Lysaker, J. T., & Lysaker, P. H. (2018). Phenomenological and recovery models of the subjective experience of psychosis: Discrepancies and implications for treatment. *Psychosis*, *10*, 1–11. http://doi.org/10.1080/17522439.2018.1522540

Hoekert, M., Kahn, R. S., Pijnenborg, M., & Aleman, A. (2007). Impaired recognition and expression of emotional prosody in schizophrenia: Review and meta-analysis. *Schizophrenia Research*, *96*, 135–145. http://doi.org/10.1016/j.schres.2007.07.023

Hughes, S. M., & Nicholson, S. E. (2010). The processing of auditory and visual recognition of self-stimuli. *Consciousness and Cognition*, *19*, 1124–1134. http://doi.org/10.1016/j.concog.2010.03.001

Insel, T. R. (2010). Rethinking schizophrenia. *Nature*, *468*, 187–193. http://doi.org/10.1038/nature09552

Javitt, D. C., & Sweet, R. A. (2015). Auditory dysfunction in schizophrenia: Integrating clinical and basic features. *Nature Reviews Neuroscience*, *16*, 535–550. http://doi.org/10.1038/nrn4002

Kronbichler, L., Stelzig-Schöler, R., Pearce, B. G., Tschernegg, M., Said-Yürekli, S., Reich, L. A., Weber, S., Aichhorn, W., & Kronbichler, M. (2018). Schizophrenia and category-selectivity in the brain: Normal for faces but abnormal for houses. *Frontiers in Psychiatry*, *9*, 1–12. http://doi.org/10.3389/fpsyt.2018.00047

Lee, J., Kwon, J. S., Shin, Y.-W., Lee, K. J., & Park, S. (2007). Visual self-recognition in patients with schizophrenia. *Schizophrenia Research*, *94*, 215–220. http://doi.org/10.1016/j.schres.2007.03.032

Lee, C. U., Shenton, M. E., Salisbury, D. F., Kasai, K., Onitsuka, T., Dickey, C. C., Yurgelun-Todd, D., Kikinis, R., Jolesz, F. A., & McCarley, R.W. (2002). Fusiform gyrus volume reduction in first-episode schizophrenia: A magnetic resonance imaging study. *Archives of General Psychiatry*, *59*, 775–781. http://doi.org/10.1001/archpsyc.59.9.775

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 86–87. http://doi.org/10.1038/264746a0

Nelson, B., Whitford, T. J., Lavoie, S., & Sass, L. A. (2014). What are the neurocognitive correlates of basic self-disturbance in schizophrenia? Integrating phenomenology and neurocognition. Part 1 (Source monitoring deficits). *Schizophrenia Research*, *152*, 12–19. http://doi.org/10.1016/j.schres.2013.06.022

Peynircioğlu, Z. F., Brent, W., Tatz, J. R., & Wyatt, J. (2017). McGurk Effect in gender identification: Vision trumps audition in voice judgments. *The Journal of General Psychology*, *144*, 59–68. http://doi.org/10.1080/00221309.2016.1258388

Pinheiro, A. P., Rezaii, N., Rauber, A., Niznikiewicz, M. (2016). Is this my voice or yours? The role of emotion and acoustic quality in self-other voice discrimination in schizophrenia. *Cognitive Neuropsychiatry*, *21*, 335–353. http://doi.org/10.1080/13546805.2016.1208611

Sandsten, K. E., Nordgaard, J., Kjaer, T. W., Gallese, V., Ardizzi, M., Ferroni, F., Petersen, J., & Parnas, J. (2019). Altered self-recognition in patients with schizophrenia. *Schizophrenia Research*, *218*, 116–124. http://doi.org/10.1016/j.schres.2020.01.022

Sass, L. A. (2014). Self-disturbance and schizophrenia: Structure, specificity, pathogenesis (Current issues, new directions). *Schizophrenia Research*, *152*, 5–11. http://doi.org/10.1016/j.schres.2013.05.017

She, S., Zhang, B., Li, X., Zhang, X., Li, R., Li, J., Bi, T., & Zheng, Y. (2017). Face-related visual search deficits in first-episode schizophrenia. *Psychiatry Research*, *256*, 144–149. http://doi.org/10.1016/j.psychres.2017.06.021

Tsakiris, M. (2008). Looking for myself: Current multisensory input alters self-face recognition. *PLoS ONE*, *3*, e4040. http://doi.org/10.1371/journal.pone.0004040

van der Weiden, A., Prikken, M., & van Haren, N. E. M. (2015). Self–other integration and distinction in schizophrenia: A theoretical analysis and a review of the evidence. *Neuroscience & Biobehavioral Reviews*, *57*, 220–237. http://doi.org/10.1016/j.neubiorev.2015.09.004

Zhang, L., Zhu, H., Xu, M., Jia, H. X., & Liu, J. (2012). Selective impairment in recognizing the familiarity of self faces in schizophrenia. *Chinese Science Bulletin*, *57*, 918–923. http://doi.org/10.1007/s11434-012-5109-z