Schizophrenia Research 179 (2017) 97–103

Contents lists available at ScienceDirect

Schizophrenia Research

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

The associations between multisensory temporal processing and symptoms of schizophrenia

Ryan A. Stevenson [a,b,*], Sohee Park [c], Channing Cochran [c], Lindsey G. McIntosh [c], Jean-Paul Noel [d], Morgan D. Barense [e,f], Susanne Ferber [e,f], Mark T. Wallace [d,g,h,i,j]

[a] The University of Western Ontario, Department of Psychology, London, ON, Canada
[b] The University of Western Ontario, Brain and Mind Institute, London, ON, Canada
[c] Vanderbilt University, Department of Psychology, Nashville, TN, USA
[d] Vanderbilt Brain Institute, Vanderbilt University, Nashville, TN, USA
[e] University of Toronto, Department of Psychology, Toronto, ON, Canada
[f] Rotman Research Institute, Toronto, ON, Canada
[g] Vanderbilt University Medical Center, Department of Hearing and Speech Sciences, Nashville, TN, USA
[h] Vanderbilt Kennedy Center, Vanderbilt University Medical Center, Nashville, TN, USA
[i] Vanderbilt University, Department of Psychology, Nashville, TN, USA
[j] Vanderbilt University Medical Center, Department of Psychiatry, Nashville, TN, USA

ARTICLE INFO

Article history:
Received 23 August 2016
Received in revised form 23 September 2016
Accepted 28 September 2016
Available online 13 October 2016

Keywords:
Multisensory integration
Schizophrenia
Temporal processing
Speech perception
Hallucinations
Audiovisual

ABSTRACT

Recent neurobiological accounts of schizophrenia have included an emphasis on changes in sensory processing. These sensory and perceptual deficits can have a cascading effect onto higher-level cognitive processes and clinical symptoms. One form of sensory dysfunction that has been consistently observed in schizophrenia is altered temporal processing. In this study, we investigated temporal processing within and across the auditory and visual modalities in individuals with schizophrenia (SCZ) and age-matched healthy controls. Individuals with SCZ showed auditory and visual temporal processing abnormalities, as well as multisensory temporal processing dysfunction that extended beyond that attributable to unisensory processing dysfunction. Most importantly, these multisensory temporal deficits were associated with the severity of hallucinations. This link between atypical multisensory temporal perception and clinical symptomatology suggests that clinical symptoms of schizophrenia may be at least partly a result of cascading effects from (multi)sensory disturbances. These results are discussed in terms of underlying neural bases and the possible implications for remediation.

1. Introduction

Hallucinations are a positive symptom in schizophrenia (SCZ) that can present as false perceptions in any sensory modality, but commonly take the form of perceived auditory voices. They are often conceptualized as false attribution of internal voices to an external source. As such, hallucinations in SCZ are frequently linked to the audiovisual speech-perception network, including areas of superior temporal and inferior frontal (i.e., Broca's) cortex (Jardri et al., 2011). One cognitive operation of this network is the integration of information across the auditory and visual systems, forming coherent percepts that comprise our conscious experience (Stevenson et al., 2014a). Speech is a powerful example of audiovisual integration, though integration extends to all manner of sensory inputs: we seamlessly bind together audible speech signals with their associated visual cues, affording substantial behavioral and perceptual benefits, ranging from faster response times (Raab, 1962) to improved speech perception (Sumby and Pollack, 1954) in healthy participants but not as much in SCZ patients. For example, seeing a speaker's visual articulation enhances speech perception under noisy conditions in healthy participants but less so in SCZ patients (Ross et al., 2007). Similarly, SCZ patients are less susceptible to the McGurk effect (Pearl et al., 2009), where the mouth movements an individual sees can alter what they believe to "hear" a speaker to be saying (McGurk and MacDonald, 1976), despite preserved unisensory abilities (Ross et al., 2007).

Impaired sensory integration is a hallmark neurological "soft sign" of SCZ (Heinrichs and Buchanan, 1988) that is often noted at the time of an individual's first psychotic episode and is correlated with SCZ symptomatology (Williams et al., 2010). Most germane to this report is the possible link between alterations in sensory integration and positive symptoms in SCZ, most notably hallucinations (Postmes et al., 2014). Exploring an integration-hallucination link is motivated by the overlap in the neural substrates for audiovisual integration and hallucinations, specifically in regions of the audiovisual speech-perception network.

---

[*] Corresponding author at: The University of Western Ontario, Department of Psychology, London, ON, Canada.
E-mail address: Rsteve28@uwo.ca (R.A. Stevenson).

http://dx.doi.org/10.1016/j.schres.2016.09.035

For example, SCZ is associated with structural (Kim et al., 2003) and functional changes within the superior temporal cortex (Surguladze et al., 2001; Szycik et al., 2009). This same area of cortex is heavily implicated in multisensory *temporal* processing (Stevenson et al., 2010). Furthermore, individuals with SCZ exhibit alterations in temporal processing (Carroll et al., 2008; Davalos et al., 2002; Elvevag et al., 2003; Foucher et al., 2007; Freedman, 1974; Giersch et al., 2009; Lalanne et al., 2012; Tysk, 1983a,b; Volz et al., 2001), and impaired audiovisual temporal precision in SCZ has been linked to inaccurately attributing auditory components of speech to temporally disparate visual speech signals (Martin et al., 2013).

Given the relationship between temporal processing and sensory integration (Stevenson et al., 2012b), and links between sensory integration and hallucination in SCZ, we hypothesize that impaired temporal perception in SCZ may be associated with hallucinations in SCZ. To investigate this, we first measured auditory, visual, and multisensory temporal perception in SCZ patients and a group of matched controls, verifying the presence of temporal dysfunction in SCZ and assessing if temporal-perception deficits were uniquely multisensory. Second, and of paramount importance, we measured the severity of hallucinations in SCZ participants with the *a priori* prediction that changes in multisensory temporal processing would be predictive of hallucinations. This finding would point to shared mechanistic substrates for changes in audiovisual temporal integration and the presence and severity of hallucinations.

2. Methods and materials

2.1. Overview

Participants completed four behavioral tasks: two unisensory timing tasks in which participants performed temporal order judgments (TOJ; “Which came first?”) with either auditory or visual stimuli, and two audiovisual timing tasks in which participants performed audiovisual simultaneity judgments (SJ; “Same time or different time?”), one with speech stimuli and one with simple flash-beep stimuli. Finally, participants completed standard metrics assessing SCZ symptomatology. Protocols were approved by Vanderbilt University Institutional Review Board and participants gave written informed consent to participate in the study.

2.2. Participants

Thirty-two participants competed the study, half who met the DSM-IV criteria for schizophrenia (SCZ; mean age = 42.3 ± 8.9 years, 8 females), and half healthy controls (HC; mean age = 41.9 ± 9.3 years, 10 females) matched for age ($t_{(30)} = 0.12$, $p = 0.91$) and gender ($\chi^2 = 0.51$, $p = 0.48$). SCZ symptoms were rated using the Brief Psychiatric Rating Scale (BPRS; mean = 15.4 ± 7.9), the Scale for Assessment of Positive Symptoms (SAPS; mean = 13.7 ± 11.7), and the Scale for Assessment of Negative Symptoms (SANS; mean = 32.2 ± 15.9), with hallucination severity derived from the SAPS global rating of hallucination scores (mean = 1.6 ± 1.6).

2.3. Stimuli and procedures

For all tasks, participants were asked to fixate towards a cross, and were actively monitored for compliance. Visual stimuli were presented on a screen approximately 60 cm from the participants. Auditory stimuli were presented through centrally aligned speakers. Tasks and trials were randomized in all cases. All responses were made via button press.

2.3.1. Unisensory timing tasks

For the unisensory auditory timing task, participants were presented with a pair of auditory beeps consisting of one high- and one low-pitch (1000 and 500 Hz) beep (duration = 7 ms), and performed a temporal order judgment task (TOJ; “Which came first?”). Individual unisensory-auditory beeps within each pair were separated by SOAs of 10, 20, 35, 50, 75, 100, 150, 200, and 250 ms. Twenty trials at each SOA were presented.

For the unisensory visual timing task, participants were presented with two white circles on a black background, one above and one below a fixation cross (duration = 10 ms) and performed a TOJ task. Individual unisensory-visual flashes within each pair were separated by SOAs of 10, 20, 30, 40, 60, 80, 100, and 150 ms. Twenty trials at each SOA were presented.

Temporal order judgment tasks were used with unisensory tasks as opposed to the SJ tasks used with multisensory stimuli based on previously collected data. When an SJ tasks was used with unisensory stimuli, most participants were near ceiling performance at detecting asynchronies even at the shortest SOAs.

2.3.2. Audiovisual timing tasks

In the audiovisual tasks, participants were presented with an auditory and a visual stimulus, and performed a simultaneity judgment task (SJ; “Were the auditory and visual stimuli presented at the same time?”). Two types of audiovisual stimuli were presented, each in a separate run. One set of stimuli were simple flash-beeps pairs. The visual flashes consisted of a white ring circumscribing the visual fixation cross on a black background presented for 10 ms. Auditory beep stimuli consisted of a 3500 Hz pure tone with a duration of 7 ms. For simple flash-beeps, SOAs included 0, ±10, ±20, ±50, ±80, and ±100 to 300 ms in 50 ms intervals (negative values indicate auditory-leading presentations, and positive values indicate visual-leading presentations). Twenty trials at each SOA were presented.

The second type of audiovisual stimuli were single syllable utterances, which were selected from a stimulus set that has been previously used successfully in studies of multisensory integration (Baum et al., 2015; Quinto et al., 2010; Stevenson et al., 2014b; Stevenson and Wallace, 2013). Stimuli consisted of two audiovisual clips of a female speaker uttering single instances of the syllables “ga” and “ba”. Visual stimuli were grayscale, and spanned 18.25 cm per side, and 2 s in duration, with each presentation containing the entire articulation of the syllable, including pre-articulatory gestures. For speech stimuli, SOAs included 0 to ±300 ms in 50 ms intervals and ±400 ms.

2.4. Analysis of behavioral tasks

In both the auditory and visual unisensory TOJ, individuals' mean responses were calculated at each SOA (Fig. 1A–B). A general linear model (GLM) was used to predict responses based on the categorical factor of diagnosis and the continuous factor of SOA. Additionally, each individual's mean responses were fit with a sigmoid curve, and the 75% threshold was extracted from this function (Fig. 1C) for both the visual and auditory tasks. Thresholds were then compared across groups, and subsequently used to predict multisensory temporal processing abilities. Twenty trials at each SOA were presented.

In both audiovisual SJ tasks, individuals' mean responses were calculated at each SOA (Fig. 1D–E). Individuals' mean responses from SJ tasks were used to calculate a temporal binding window (TBW) using a well-established method (Fister et al., 2016; Noel et al., 2016; Schlesinger et al., 2014; Stevenson et al., 2012a,b, 2014a,b, 2013; Stevenson and Wallace, 2013). Two psychometric sigmoid functions were fit to rates of perceived synchrony across SOAs; one to the audio-first (left) presentations and a second to the visual-first presentations (right). To account for non-zero points of subjective simultaneity (PSS), the SOA at which these two sigmoid functions crossed was extracted. If this point was greater or less than the next closest data point, two new sigmoid functions were fit splitting the data at the SOA at which the original sigmoid functions crossed. This process was continued in an iterative manner until the SOA at which best-fit sigmoid functions crossed fell between the two data points at which the data were split. Based off these final curves, the time interval between the 75% threshold of their left,