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

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

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 of 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.
For example, SCZ is associated with structural (Kim et al., 2003) and functional changes within the superior temporal cortex (Suruguladze 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 importance, we measured the severity of hallucinations in SCZ participants with 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 (χ² = 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 task 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 was 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,
auditory-leading curve and their right, visual-leading curve was calculated as the individual’s TBW (Fig. 1F).

All measures were compared between groups. Consecutive multiple hierarchical regressions were then conducted to investigate (a) the ability of unisensory temporal processing to account for impaired multisensory processing in SCZ, and (b) the ability of audiovisual temporal processing to predict hallucination symptomatology in SCZ.

3. Results

3.1. Unisensory temporal perception

Auditory temporal perception was indexed via an auditory TOJ task. Responses were averaged for each SOA for each individual. A mixed linear model (MLM) was then used to measure the impact of SOA and diagnosis on response accuracy (Fig. 1A). In this 2-factor MLM, both factors of diagnosis ($F_{(1,228.7)} = 84.57, p < 0.001$) and SOA ($F_{(1,26.2)} = 58.84, p < 0.001$) significantly contributed, but the two did not interact ($F_{(1,56.0)} = 0.06, p < 0.81$). To quantify this between-group difference, thresholds (75%-correct performance) for each individual were compared between diagnostic groups, with SCZ patients showing poorer auditory temporal acuity (Fig. 1C; $\text{mean}_{\text{SCZ}} = 140 \text{ ms}, \text{mean}_{\text{HC}} = 53 \text{ ms}, t_{(30)} = 2.70, p = 0.01, d = 0.99$).

Visual temporal perception was indexed via a visual TOJ task. Responses were averaged across SOAs for each individual. An MLM was then used to measure the impact of SOA and diagnosis on response accuracy (Fig. 1B). In this 2-factor MLM, both factors of diagnosis ($F_{(1,57.7)} = 28.51, p < 0.001$) and SOA ($F_{(1,214.5)} = 234.70, p < 0.001$) significantly contributed, but the two did not interact ($F_{(1,56.0)} = 0.06, p < 0.81$).

Fig. 1. Temporal perception. Individuals with (red) and without (black) schizophrenia completed a series of temporal tasks. In unisensory tasks, individuals with SCZ showed less temporal precision during auditory (Panel A) and visual (Panel B) perception. Thresholds for both can be seen in Panel C. Audiovisual temporal perception was also less temporally precise in schizophrenia with both simple flash-beep stimuli (Panel D) and speech stimuli (Panel E). Thresholds for both can be seen in Panel F. Error bars represent standard error. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
significantly contributed, but the two did not interact ($F_{1,124.5} = 0.08, p < 0.78$). To quantify this between-group difference, 75% thresholds were compared between diagnostic groups, with SCZ patients showing poorer visual acuity (Fig. 1C; mean_{SCZ} = 45 ms, mean_{HC} = 26 ms, $t_{30} = 2.22, p = 0.03, d = 0.81$).

3.2. Multisensory temporal perception

Audiovisual temporal acuity was tested using SJ tasks. Data from the SJ task using simple flash-beep stimuli were analyzed with a mixed-model, two-way, repeated-measures ANOVA (diagnosis × SOA). This analysis revealed significant effects of diagnosis ($F_{1,18} = 13.18, p = 0.001$, partial-$\eta^2 = 0.31$), SOA ($F_{1,18} = 68.33, p < 0.001$, partial-$\eta^2 = 0.70$), and an interaction between the two ($F_{1,18} = 7.94, p < 0.001$, partial-$\eta^2 = 0.21$) (Fig. 1D). Follow-up $t$-tests were conducted at each SOA (see Fig. 1D and Supplementary Table 1 for detailed statistics). Individuals’ temporal binding windows (TBW) were calculated and groups were compared (Fig. 1D; $t_{30} = 4.61, p = 0.694e^{-5}, d = 1.68$). Controls showed a mean TBW of 240 ms ± 114 ms, and SCZs exhibited a significantly enlarged mean TBW of 550 ms ± 243 ms. These results suggest that multisensory temporal acuity is less precise in individuals with SCZ (Fig. 1F).

SJ tasks using more complex audiovisual stimuli – specifically speech stimuli – were analyzed in an identical manner. A mixed-model, two-way, repeated-measures ANOVA showed significant results of added predictors are shown in Table 1. In a synopsis of these data, unisensory auditory, but not visual, TOJ performance was predictive of TBW width. In model 3, SCZ diagnosis also predicted a wider TBW beyond what was accounted for by unisensory deficits.

3.4. Multisensory temporal precision and hallucinations

Global hallucination severity measures in participants with SCZ were extracted from the SAPS. Initial correlations were conducted between hallucination severity and participants’ audiovisual temporal precision (i.e., TBW width). The widths of the TBW measured using both flash-beep stimuli ($R = −0.52, p = 0.038$) and speech stimuli ($R = −0.65, p = 0.006$) were significantly correlated with hallucination severity, in that the wider the TBW, the less severe their hallucinations were (Fig. 2). To control for the effects of age and gender, a hierarchical regression was conducted, revealing that (a) neither gender nor age accounted for a significant portion of variance ($p > 0.37$), and (b) audiovisual temporal precision was predictive of hallucination severity even when controlling for demographic variables ($R = 0.76, F_{\Delta 2,11} = 5.76, p = 0.019$). Detailed statistical results can be seen in Table 2.

Finally, an exploratory correlation analysis was conducted to detect any possible relationships between audiovisual temporal precision and overall positive and negative symptomatology. No relationships were observed for either flash-beep ($Rs = 0.23$ and 0.20, respectively) or speech ($Rs = 0.08$ and 0.12, respectively) stimuli (Table 2).

4. Discussion

This study provides a novel view into the relationships between impaired temporal processing, multisensory integration, and hallucinations in SCZ. Three main findings are evident in the data. First, this study confirms that individuals with SCZ show decreased temporal acuity in both auditory and visual perception, as well as in audiovisual temporal perception. Second, SCZ participants exhibit impairments in multisensory temporal acuity that extend beyond these unisensory changes, suggesting a level of specificity for these multisensory changes. Finally, and perhaps most importantly, multisensory temporal perception predicted one aspect of SCZ symptomatology, specifically the severity of hallucinations.

The ability to integrate sensory inputs into a unified perceptual whole is an essential cognitive process. The ability to link what is seen and heard, such as linking a voice one hears to the sight of the speaker’s mouth movements, creates a coherent representation of the external world. Changes in sensory integration are clinically significant features

Table 1

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Flash-beep stimuli</th>
<th>Speech stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td><strong>Step 2</strong></td>
<td><strong>Step 3</strong></td>
</tr>
<tr>
<td>Partial correlation (pr)</td>
<td>p-Value</td>
<td>Partial correlation (pr)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.06</td>
<td>0.29</td>
</tr>
<tr>
<td>Age</td>
<td>0.02</td>
<td>0.61</td>
</tr>
<tr>
<td>Age</td>
<td>0.06</td>
<td>0.61</td>
</tr>
<tr>
<td>Auditory threshold</td>
<td>0.52</td>
<td>0.61</td>
</tr>
<tr>
<td>Visual threshold</td>
<td>0.32</td>
<td>0.61</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>0.13</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Significant results of added predictors are shown in bold.

Please cite this article as: Stevenson, R.A., et al., The associations between multisensory temporal processing and symptoms of schizophrenia, Schizophr. Res. (2016), http://dx.doi.org/10.1016/j.schres.2016.09.035
Table 2

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Partial correlation ($r'$)</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.25</td>
<td>0.38</td>
</tr>
<tr>
<td>Age</td>
<td>0.20</td>
<td>0.48</td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.54</td>
<td>0.06</td>
</tr>
<tr>
<td>Age</td>
<td>0.05</td>
<td>0.87</td>
</tr>
<tr>
<td>TBW flash beep</td>
<td>-0.57</td>
<td>0.04</td>
</tr>
<tr>
<td>TBW speech</td>
<td>-0.43</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Significant results of added predictors are shown.

5. Conclusions

Our results support the hypothesis that sensory disturbances, specifically those in the temporal processing realm, contribute to hallucinations in SCZ. SCZ is associated with auditory and visual temporal...
dysfunction, with additional multisensory temporal dysfunction beyond that predicted by these unisensory deficits. These audiovisual temporal perceptual disturbances are also significantly predictive of clinical measures of hallucination severity, supporting the hypothesis that hallucinations may result from aberrant attribution, or integration, of internal auditory speech to an external speaker. These data are, to our knowledge, the first to demonstrate the cascading impact of sensory disturbances on higher-level, clinical symptomatology in SCZ. The highly overlapping neural architecture underlying temporal processing, multisensory integration, and speech perception, and that associated with hallucinations in SCZ, further support these findings. These findings also offer hope for the use of temporally-based sensory training methods as possible remediation tools in SCZ.

Role of funding

Funding did not impact the outcome of this research in any way.

Contributions

This study was designed by authors RS, SP, CC, LM, and MW and conducted by RS, CC, and LM. Data were analyzed by RS and interpreted by RS, SP, MB, SF, and MW. The manuscript was drafted by RS, and edited and approved by all authors.

Financial disclosures

Dr. Stevenson reported no biomedical financial interests or potential conflicts of interest.
Dr. Park reported no biomedical financial interests or potential conflicts of interest.
Ms. Cochran reported no biomedical financial interests or potential conflicts of interest.
Mr. McIntosh reported no biomedical financial interests or potential conflicts of interest.
Dr. Barense reported no biomedical financial interests or potential conflicts of interest.
Dr. Ferber reported no biomedical financial interests or potential conflicts of interest.
Dr. Wallace reported no biomedical financial interests or potential conflicts of interest.

Acknowledgements

R.S. was funded by a Banting Fellowship from the Canadian National Science and Engineering Research Council, the Autism Research Training Program funded by the Canadian Institutes for Health Research, and the National Institutes of Deafness and Communication Disorders (NICDC 1F32 DC011993). S.P. was funded by grants from the National Alliance for Research in Schizophrenia and Affective Disorders, and the National Institutes for Mental Health, Gertrude Conaway Vanderbilt Endowment. M.T.W. was funded by the National Institutes of Health grants CA183492 and HD083211, the Simons Foundation Autism Research Initiative and the Wallace Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.schres.2016.09.035.

References


