Mismatch negativity impairment is associated with deficits in identifying real-world environmental sounds in schizophrenia

Permalink
https://escholarship.org/uc/item/3xn591sq

Journal
SCHIZOPHRENIA RESEARCH, 191

ISSN
0920-9964

Authors
Joshi, YB
Breitenstein, B
Tarasenko, M
et al.

Publication Date
2018

DOI
10.1016/j.schres.2017.05.020

Peer reviewed
Mismatch negativity impairment is associated with deficits in identifying real-world environmental sounds in schizophrenia

Yash B. Joshi, Barbara Breitenstein, Melissa Tarasenko, Michael L. Thomas, Wei-Li Chang, Joyce Sprocka, Richard F. Sharp, Gregory A. Light

Background: Patients with schizophrenia (SZ) have impairments in processing auditory information that have been linked to deficits in cognitive and psychosocial functioning. Dysfunction in auditory sensory processing in SZ has been indexed by mismatch negativity (MMN), an event-related potential evoked by a rare, deviant stimulus embedded within a sequence of identical standard stimuli. Although MMN deficits in SZ have been studied extensively, relatively little is known about how these deficits relate to accurately identifying real-world, ecologically-salient sounds.

Methods: MMN was assessed in SZ patients (n = 21) and non-psychiatric comparison subjects (NCS; n = 16). Participants were also assessed in their ability to identify common environmental sounds using a subset of 80 sound clips from the International Affective Digitized Sounds 2nd Ed collection.

Results: SZ patients made significantly more errors in environmental sound identification (p < 0.001, d = 0.86) and showed significantly reduced MMN amplitude deficits in MMN compared to NCS (p < 0.01, d = 0.97). In SZ patients, MMN deficits were associated with significantly greater environmental sound identification errors (r = 0.61, p < 0.01).

Conclusions: Impairments in early auditory information processing in schizophrenia account for significant proportions of variance in the ability to identify real-world, functionally relevant environmental sounds. This study supports the view that interventions targeting deficits in low-level auditory sensory processing may also impact more complex cognitive brain processes relevant to psychosocial disability.

1. Introduction

Auditory object perception reflects the ability to identify sounds in one’s environment. This involves transforming acoustic parameters, such as pitch, harmonicity, onset, timing, timbre, and spatial location into meaningful percept based on previous experiences, all of which require accurate auditory perception. In patients with schizophrenia (SZ), auditory perception is profoundly impaired, contributing to bottom-up deficits associated with cognitive dysfunction, and, ultimately, social functioning (Thomas et al., 2017; Vinogradov and Nagarajan, 2017; Force et al., 2008; Kiang et al., 2009; Leitman et al., 2005, 2010; Javitt, 2009; Braff and Light, 2004; Rissling et al., 2014; Kirihara et al., 2012).

Early auditory perceptual impairment can be assessed electrophysiologically by mismatch negativity (MMN), a well-established event-related potential (ERP) evoked by a rare unexpected stimuli occurring within a sequence of frequent standard stimuli. MMN indexes auditory sensory discrimination and is a predominantly automatic process that can be measured in the absence of directed attention (Kasai et al., 1999; Wacongne, 2016; Symonds et al., 2016; Umbricht et al., 1999, 1998; Rissling et al., 2013). MMN deficits are well documented in SZ and have been associated with cognitive, social cognitive, and functional impairments (Light and Braff, 2005a, 2005b; Kawakubo et al., 2006, 2007; Light et al., 2007; Rissling et al., 2013, 2014; Light et al., 2015; Perez et al., 2017; Wynn et al., 2010; Thomas et al., 2017).

Although the association between MMN and errors in processing verbal stimuli (tone matching, synthetically derived emotional stimuli, reading ability) have been well-characterized (e.g., Leitman et al., 2005; Leitman et al., 2010), less is known about non-verbal auditory stimuli. In one of the few studies of environmental sound identification, SZ patients were less accurate than controls in identifying sounds across

1 denotes shared first authorship.

Corresponding author at: Department of Psychiatry, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0804, United States.

E-mail address: glight@ucsd.edu (G.A. Light).
a range of semantic categories (Tüscher et al., 2005). It is not yet known, however, if these deficits in higher-order auditory mnemonic processing are associated with, or perhaps even a consequence of impairments in early auditory information processing. We hypothesized that SZ patients would have deficits in environmental sound identification and these deficits would correlate with MMN. In addition to impairments in the ability to accurately recognize the content of auditory stimuli, patients with SZ have deficits in emotional expression and perception, including interpretation of emotional tones of voices (Hoekert et al., 2007; Kohler et al., 2004; Kohler and Martin, 2006; Kring and Moran, 2008; Mandal et al., 1998; Trèmeau, 2006; Kantrowitz et al., 2015). Therefore, we secondarily aimed to determine whether SZ patients show abnormal ratings of emotional valence and arousal to real-world environmental stimuli.

2. Materials and method

2.1. Participants

Participants included SZ patients (n = 21) and non-psychiatric comparison subjects (NCS; n = 16). SZ patients were recruited from community residential facilities and via physician referral; comparison subjects were recruited through newspaper advertisements and flyers posted in the community. All subjects gave written informed consent via methods approved by the UCSD Institutional Review Board (protocol #071128). Exclusion criteria for patients and controls included any current or past neurological insult or a positive urine toxicology screen. Audiometer testing was utilized to ensure that all participants had normal hearing in both ears and could detect 45-dB SPL tones at 500, 1000 and 6000 Hz. Patients met criteria for SZ based on the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID; First et al., 1996). Patients who met criteria for Axis I disorders other than SZ were excluded from the study. Screening interviews were conducted with control subjects to rule out past or present Axis I or II diagnoses (SCID-NP; First et al., 2002; SCID-II; First et al., 1997). Demographic and clinical characteristics of the sample are presented in Table 1, and patients’ psychotropic medications are presented in Table 1, and patients’ psychotropic medications are presented in Table 1, and patients’ psychotropic medications are presented in Table 1.

There were no significant sex differences across groups (χ² = 0.19, p > 0.05); however, patients were older (t = 3.80, p = 0.001) and completed significantly fewer years of education (t = 4.53, p < 0.001) than the NCS.

2.2. Emotion ratings and sound identification task

A subset of 80 six-second sound clips was selected from the Interna-
tional Affective Digitized Sounds, 2nd Ed. (IADS-2) collection (Bradley and Lang, 2007) for presentation to participants. The sounds were classified into the following categories based on data gathered from the extensive IADS-2 normative sample: 1) positive valence/high arousal (e.g. crowd cheering, boy laughing), 2) negative valence/high arousal (e.g. plane crash, baby crying), 3) positive valence/low arousal (e.g. seagull, whistling), and 4) neutral (e.g. lawnmower, sneeze). The IADS-2 battery does not contain negative/low arousal sounds. Twenty exemplars from each category were delivered through headphones in a randomized order using E-Prime Version 2.0 (PST Technologies, Pittsburgh, PA). Participants were instructed to listen carefully to each sound and then rate how they felt while listening to the sounds on a computer using a 1–9 self-assessment manikin (SAM; a non-verbal pictorial assessment of valence and arousal using a Likert-type scale; Bradley and Lang, 1994) that ranged from very unhappy to very happy and from very calm to very excited.

After rating all 80 stimuli, the sounds were presented again and participants were asked to identify the source of the sounds in an open response format. Incorrect answers were also qualitatively evaluated by researcher consensus as “incorrect but perceptually similar” (e.g., bees buzzing for bacon sizzling, rain for applause), “incorrect and perceptually dissimilar” (e.g., birds chirping for vacuum cleaner, car engine for people at a restaurant), or “completely unknown.” If the participant responded with “completely unknown”, they were prompted two additional times to identify the sound. Total duration of the rating and identification tasks was approximately 40 min.

2.3. EEG recordings

Electroencephalograph (EEG) data acquisition, stimulation, and processing was performed in accordance with previously published methods (Light and Braff, 2005a, 2005b; Rissling et al., 2012). EEG was recorded from 64 channels arranged in a cap. A reference electrode was placed at the nose tip, in addition to a ground electrode at Fpz. Four additional electrodes were placed above and below the left eye as well as at the outer canthi of both eyes in order to monitor blinks and eye movements. EEG was digitally referenced off-line to linked mastoids (TP9/TP10). All impedances were kept below 4 kΩ. Signals were digitized at a rate of 1 kHz with system acquisition filter settings at 0.5–100 Hz. Subjects were presented with binaural tones (1 kHz, 85 dB sound pressure level, with 1 ms rise/fall) with a fixed stimulus onset-to-onset asynchrony of 500 ms using E-Prime software. Standard (90% probability; 50 ms duration) and deviant (10% probability; 100 ms duration) tones were presented in pseudorandom order using foam insert earphones. During EEG recording, subjects were instructed to watch a silent, non-valenced cartoon video. EEG acquisition was terminated when 225 deviant trials free of gross artifacts (± 100 μV at frontocentral electrodes) were collected. Subsequent data processing was performed offline using automated procedures. Continuous recordings were mathematically corrected for eye movement artifact employing independent component analyses. Data were then divided into epochs relative to the onset of stimuli (−100 to 500 ms) and centered at the mean of the prestimulus baseline. Epochs containing ± 50 μV in frontal recording sites were automatically rejected. MMN waveforms were generated by subtracting the ERP waveforms in response to standard tones from the waveforms elicited by the deviant tones. The resultant difference waves were low-pass filtered at 20 Hz (zerophase shift, 24 dB/octave rolloff) to remove any residual high-frequency artifact consistent with established methods (Jahshan et al., 2012; Kiang et al., 2009; Light and Braff, 2005a, 2005b; Light et al., 2007). Search windows for peak MMN were 135–205 ms, respectively. Mean amplitudes in the 25 ms surrounding the identified peaks were then automatically calculated for each electrode channel.

2.4. Statistical analyses

Inferential tests of group differences were based on t-tests and chi-squared statistics. Well-defined MMNs were present in all subjects’ data and were verified by inspection of butterfly plots and topographical maps. MMN analyses utilized data from channel Fz. Relationships between MMN and behavioral data were examined with Spearman rank correlation coefficients. A significance level of α = 0.05 was used for all analyses.

| Table 1 Demographic and clinical information. (Means ± standard deviation given where applicable). |
|-------------------------|-------------------------|-------------------------|
|                         | NCS                     | SZ                      |
| Age, years*             | 35.8 (8.64)             | 45.5 (7.00)             |
| Sex (5 female)          | 50.0%                   | 42.9%                   |
| Years of education*     | 14.88 (2.68)            | 11.48 (1.89)            |
| Number of hospitalizations | –                      | 7.50 (5.74)            |
| SAPS scale scores       | –                      | 8.52 ± 4.27            |
| SANS scale scores       | –                      | 15.67 ± 3.67            |

* Patients differed significantly from normal controls, p < 0.01.
3. Results

3.1. Mismatch negativity

Consistent with previous studies, SZ patients exhibited significantly reduced MMN amplitudes compared to NCS at Fz ($t(35) = 2.92, p < 0.01, d = 0.97$). As shown in Fig. 1, NCS had a mean amplitude (SD) of $-2.43 \mu V (0.95)$, SZ patients $-1.50 \mu V (0.97)$.

3.2. Sound identification paradigm

SZ patients were significantly impaired in sound identification as compared to controls, as shown in Table 2. NCS were significantly better at identifying sounds overall ($t(35) = 4.44, p < 0.001, d = 0.86$) and within each sound category compared to SZ. Both groups showed similar error patterns, with “incorrect but perceptually similar” as the most common error type, followed by “completely unknown” and “incorrect and perceptually dissimilar”. However, schizophrenia patients demonstrated significantly more incorrect perceptually similar errors, incorrect perceptually dissimilar errors, and completely unknown errors than NCS.

3.3. Affective sound rating paradigm

No significant differences were found between SZ patients and HCS in their ratings of arousal or valence in any of the 4 sound categories (negative valence/high arousal, AN; positive valence/high arousal, AP; neutral, N; positive valence/low arousal, UP), as shown in Fig. 2 (all $p’s > 0.13$).

3.4. Correlation analyses

Correlations between MMN correct sound identification, and all three identification error types (“completely unknown sound”, “incorrect but perceptually similar”, “incorrect and perceptually dissimilar”) were examined. MMN deficits in SZ patients were correlated with sound identification, but no correlation between MMN was found in sound identification accuracy in NCS. As shown in Fig. 3, larger amplitude MMN was associated with greater sound identification accuracy ($r = 0.61, p < 0.01$), a finding that persisted even when controlling for age ($r = 0.58, p < 0.01$). MMN deficits significantly correlated with two out of three error types: “completely unknown” ($r = 0.65, p < 0.01$) and “incorrect but perceptually similar” ($r = 0.33, p < 0.05$). MMN deficits did not correlate with the third error type “incorrect and perceptually dissimilar” ($r = 0.13, p > 0.05$).

Table 2

Mean correct and different error categories out of the entire sound pool (80 total sounds) ± standard deviation (SD) for non-psychiatric comparison subjects (NCS) and schizophrenia patients (SZ).

<table>
<thead>
<tr>
<th>Error Category</th>
<th>NCS Mean (SD)</th>
<th>SZ Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct (%)</td>
<td>94.3% (3.29)</td>
<td>84.7% (6.43)</td>
</tr>
<tr>
<td>Incorrect, perceptually similar</td>
<td>5.4% (3.29)</td>
<td>10.3% (4.31)</td>
</tr>
<tr>
<td>Incorrect, perceptually dissimilar</td>
<td>0.15% (0.43)</td>
<td>2.15% (2.96)</td>
</tr>
<tr>
<td>Completely unknown</td>
<td>0.15% (0.43)</td>
<td>2.9% (3.36)</td>
</tr>
</tbody>
</table>

* Patients differed significantly from normal controls, $p < 0.01$.  
** Patients differed significantly from normal controls, $p < 0.05$.

4. Discussion

Results of this study show that deficits in mismatch negativity (MMN)—a candidate biomarker for early detection, disease progression and treatment development in schizophrenia—are associated with impaired ability to identify real-world, functionally relevant environmental sounds. These findings suggest that environmental sound misidentification is a core sensory deficit in SZ patients (Tüschler et al., 2005), and perhaps a consequence of impairments in the basic auditory network functioning (e.g., Thomas et al., 2017; Javitt, 2009; Rissling et al., 2014; Takahashi et al., 2013; Kirihara et al., 2012).

Results further indicated that patients with schizophrenia show no differences in their ratings of both the arousal and valence of environmental sound stimuli in comparison to HCSs. This finding has previously been observed in other studies that used visually (e.g., film clips, pictures, faces) and gustatory (e.g., consuming flavored liquids) stimuli, as well as behavioral tasks (e.g., maintaining facial expressions, social interactions; for a review, see Cohen and Minor, 2010; Kring and Moran, 2008). Indeed, whereas patients with schizophrenia consistently report low levels of positive affect and high levels of negative affect on trait measures, when directly exposed to affective stimuli, they do not differ significantly from healthy subjects in their emotional responses. The data in the present study supports these prior findings, and suggests that affective and semantic identification of environmental sounds are dissociable processes. That notwithstanding, future studies are needed to examine how the emotional processing of stimuli varies across modalities and clarify the neural substrates of impaired sound identification.

The results of the present study should be considered in light of several limitations. First, the primary finding of an association between MMN and environmental sound identification does not necessarily
implies that a causal link between impairments in early auditory processing and sound identification. While these findings are generally consistent with our recent structural equation modeling study of schizophrenia, larger sample sizes are required disentangle the presumed causal relationships (cf. Thomas et al., 2017). As with many cross-sectional studies of schizophrenia patients, the extent to which psychotropic medications influence the results is not known. Although antipsychotic medications do not appear to substantially impact MMN amplitudes (e.g., Leitman et al., 2010; Light et al., 2012; Swerdlow et al., 2016b), it is possible that patients’ medications may have normalized their affective responses to environmental sounds. Longitudinal randomized controlled studies are needed to characterize medication effects on behavioral and neurophysiologic responses to both ecologically relevant and simple tone-probe stimuli.

The association of sensory processing with higher-order cognition, including accurate identification of environmental sounds, supports the emerging strategy of targeting abnormalities in low-level auditory network dysfunction, which is increasingly viewed as a promising component of next-generation treatments (e.g., Fisher et al., 2009; Vinogradov and Nagarajan, 2017; Light and Swerdlow, 2015; Tarasenko et al., 2014, Thomas et al., 2017; Kautzrott et al., 2015). The use of neurophysiologic biomarkers of early auditory processing (e.g., MMN, gamma oscillations) to predict and monitor response to these treatments (e.g., Perez et al., 2017; Swerdlow et al., 2016a, 2016b; Light et al., in press) will accelerate the pace of translating findings from experimental medicine paradigms to clinical practice.

Conflicts of Interest

Dr. Light has served as a consultant for Astellas, Boehringer-Ingelheim, Heptares, Merck, NeuroSig, and Takeda unrelated to this project. The remaining authors report no financial relationships with commercial interests. None of the authors have relevant financial interests in the manuscript. Funding sources had no role in the conduct of this research.

Acknowledgements

This work was supported by the Sidney R. Baer, Jr. Foundation, Brain and Behavior Research Foundation, Department of Veteran Affairs (VISN 22 Mental Illness Research, Education, and Clinical Center; MIRECC) and National Institute of Mental Health (MH079777). Dr. Joshi is supported by a training grant from the National Institute of Mental Health, MH1101072 and by the VA MIRECC PALA Pilot grant. Dr. Thomas is supported by a National Institute of Mental Health Career Development Award (K23 MH102420).

References


