

Research Paper

A systematic review of the neural correlates of multisensory integration in schizophrenia

Cornelia Gröhn, Elin Norgren, Lars Eriksson *

Department of Social and Psychological Studies, Karlstad University, Karlstad, Sweden

ARTICLE INFO

Keywords:

Schizophrenia
Multisensory integration
fMRI
EEG
Neural correlates
Multimodal perception

ABSTRACT

Multisensory integration (MSI), in which sensory signals from different modalities are unified, is necessary for our comprehensive perception of and effective adaptation to the objects and events around us. However, individuals with schizophrenia suffer from impairments in MSI, which could explain typical symptoms like hallucination and reality distortion. Because the neural correlates of aberrant MSI in schizophrenia help us understand the physiognomy of this psychiatric disorder, we performed a systematic review of the current research on this subject. The literature search concerned investigated MSI in diagnosed schizophrenia patients compared to healthy controls using brain imaging. Seventeen of 317 identified studies were finally included. To assess risk of bias, the Newcastle-Ottawa quality assessment was used, and the review was written according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA). The results indicated that multisensory processes in schizophrenia are associated with aberrant, mainly reduced, neural activity in several brain regions, as measured by event-related potentials, oscillations, activity and connectivity. The conclusion is that a fronto-temporal region, comprising the frontal inferior gyrus, middle temporal gyrus and superior temporal gyrus/sulcus, along with the fusiform gyrus and dorsal visual stream in the occipital-parietal lobe are possible key regions of deficient MSI in schizophrenia.

1. Introduction

We constantly encounter an abundance of sensory information that has to be successfully organized for us to be able to make sense of it. The converging processes of sensory modalities (e.g., auditory, visual, tactile modalities) required to generate a meaningful and coherent perception underlie the concept of multisensory integration (MSI; Talsma et al., 2010). For example, during a conversation with a friend in a busy restaurant, you will perceptually process your friend's voice (i.e., auditory stimuli) and articulations (i.e., visual stimuli) concurrently and rapidly to integrate them. This integration will increase your speech perception because MSI enhances perceptual acuity and improves detection, discrimination and response speed (Wallace et al., 2020). MSI does not only help us avoid cognitive overload and create meaning in the constant sensory information flood (Jensen et al., 2019), but also plays a crucial role for our daily functioning and well-being through guiding our responses to the complex outer world (Tseng et al., 2015).

Different types of research paradigms have been used to study the effects of MSI on behavior. Some focus on multisensory illusions that

show how information from different sensory modalities can fuse together into one percept. One example is the *McGurk effect* in which a video of a person saying one phoneme (e.g., 'ga') is dubbed with a recording of another phoneme (e.g., 'ba') and resulting in the perceived illusion of a mixed phoneme (e.g., 'da') (McGurk & MacDonald, 1976). Other studies focus on the performance enhancement multisensory stimuli enable. One example is the redundant signals effect (RSE), which shows that responses are faster and more precise with stimuli presented in multiple sensory modalities compared to a single modality (Hershenson, 1962; Kinchla, 1974). Groundbreaking electrophysiology studies of neurons in the superior colliculus discovered important principles for MSI (Meredith and Stein, 1983, 1986; Stein and Stanford, 2008). According to the principle of inverse effectiveness, multisensory enhancement is greater if the unisensory signals are of low intensity. In addition, multisensory facilitation is maximal when stimuli from different modalities are presented at the same time (temporal rule) at around the same place (spatial rule), and decreases with increased inter-stimulus onset (Stone et al., 2014).

* Corresponding author at: Department of Social and Psychological Studies, Karlstad University, SE-651 88 Karlstad, Sweden.

E-mail address: lars.m.eriksson@kau.se (L. Eriksson).

1.1. Neural correlates of multisensory integration

Numerous brain-imaging studies of human and non-human animals (e.g., primates, rodents and cats) have provided evidence for MSI in, and anatomical pathways between, several different locations in the brain. These locations include both higher-order and lower-order cortical areas that are multisensory in nature. Important higher-order association areas are the temporal (i.e., superior temporal sulcus), parietal (i.e., intra parietal sulcus), and frontal (i.e., premotor cortex, prefrontal cortex) cortical regions (Cappe et al., 2012a; Ghazanfar and Schroeder, 2006). For example, the superior temporal sulcus (STS) contains neurons with multisensory properties (i.e., bimodal, trimodal or sub-threshold neurons) that respond to stimuli in auditory, visual and somatosensory modalities (Baylis et al., 1987; Desimone and Gross, 1979), and has anatomical connections with visual and auditory cortical areas as well as the prefrontal cortex (Cappe et al., 2012a). Multiple neuroimaging studies on humans have shown that STS is responsive to multisensory stimuli, indicating that it possibly plays a general role in perception of speech and biological motion (see Calvert, 2001, for a review). Functional magnetic resonance imaging (fMRI) studies on humans indicate early interactions between lower-level sensory cortical areas in multisensory processing (see Cappe et al., 2012a, for a review), and the superior colliculus is an important subcortical site for multisensory integration (Meredith and Stein, 1983, 1986; Stein and Stanford, 2008). Other subcortical structures like the striatum, cerebellum, amygdala and thalamus are also involved in multisensory processes (Engel et al., 2012).

However, the neural foundations of multisensory processes, constituting oscillations, networks and functional connectivity, are still not well-understood (Keil and Senkowski, 2018). The classical view suggests that integration occurs by convergence (Engel et al., 2012), and this means that integration is a feedforward and hierarchical process in which sensory information is first processed in primary sensory cortices and then integrated in higher order association areas and specialized subcortical regions (Meredith, 2002; Stein and Meredith, 1993). Integration by convergence cannot alone explain multisensory processing, in that, for example, we know cross-modal interactions take place already in primary cortices (Engel et al., 2012; Ghazanfar and Schroeder, 2006; Kayser and Logothetis, 2007). In their review, Keil and Senkowski (2018) present an integrative framework for the role of neural oscillations in multisensory integration built on recent studies. They propose that different frequency band-power (e.g., alpha-, beta-, gamma-, delta-, and theta-band) and functional connectivity networks are associated with different multisensory processes. For example, bottom-up processes, top-down processes and predictions across sensory modalities modulate integration, leaving MSI highly flexible and context-dependent. Nevertheless, our understanding of the neural correlates is still rather poor.

1.2. MSI and schizophrenia

Schizophrenia is a psychiatric condition characterized by both positive and negative symptoms. Hallucination, delusion and disorganization are referred to as positive symptoms while negative symptoms are a loss of premorbid functions such as loss of motivation and avolition (Bodén, 2016). The associated cognitive impairments in schizophrenia are hypothesized to result from atypical neural communication (van den Heuvel and Fornito, 2014), and several fMRI studies have identified altered anatomical neural connectivity in various brain regions (Allen et al., 2008; Crossley et al., 2016). A consequence of these neurocognitive alterations is the disability to acquire optimal adaptation for normal functioning with deficits in sensory integration (Heinrichs and Buchanan, 1988; Steinmann et al., 2019).

Many studies confirm that deficits in integration of sensory information in schizophrenia are apparent in visual, auditory and tactile modalities (e.g., de Gelder et al., 2003; Ferri et al., 2014; Vlcek et al.,

2014). The impaired sensory perception is suggested to be related to passivity symptoms like hallucinations (Surguladze et al., 2001). The internal forward model is a widely accepted explanatory model of passivity symptoms in schizophrenia (Chen et al., 2015). It suggests that the characteristic passivity symptoms in schizophrenia derive from an inability to make accurate predictions about the perceptual or sensory outcome of their own intentional actions (Frith, 2005). Postmes et al. (2014) also argue that passivity symptoms in schizophrenia arise due to perceptual incoherence. They theorize that hallucination represents a coping mechanism to resolve incoherent multisensory experiences. Since sensory disturbances are apparent across several modalities and hallucinations are a multimodal phenomenon, Wallace et al. (2020) state that the understanding of these deficits will benefit from a multisensory perspective.

A growing number of studies have investigated MSI in schizophrenia with different paradigms and several of them present evidence of significant behavioral consequences of impaired audiovisual integration (e.g., de Jong et al., 2009; Ross et al., 2007; Stevenson et al., 2017). Facilitation effects on reaction times do not occur in schizophrenia patients in the same way as they do in healthy controls (Williams et al., 2010). In addition, patients experience the McGurk effect more rarely, which indicates less integration interference (Pearl et al., 2009). Findings manifest abnormalities in MSI with widened temporal binding window for individuals with schizophrenia (e.g., Foucher et al., 2007; Zhou et al., 2018), and impaired MSI appears to be most evident in speech-related audiovisual stimuli and worsened in a noisy environment (Wallace and Stevenson, 2014).

However, the results have not been consistent. The results of de Boer-Schellekens et al. (2014) indicate no deficits in MSI during tasks only engaging lower-level functions, supporting the notion that diminished sensitivity to visual temporal order inflicts on MSI rather than indicating deficits in MSI alone. de Gelder et al. (2003) found no difference in performance on simple MSI tasks between individuals with schizophrenia and healthy controls, but impairments in MSI were evident in schizophrenia during tasks involving speech. Williams et al. (2010) found impaired MSI using a basic audiovisual paradigm consisting of a simple detection task without any speech components. Still, some studies show no impairments in MSI (e.g., Aine et al., 2017).

In the meta-analysis of Tseng et al. (2015), it was concluded that impairments in MSI for non-emotional stimuli are evident in schizophrenia. Two studies reported increased multisensory facilitation effects with stimuli containing emotion-triggering aspects, for example a sad face or the sound of laughter (de Gelder et al., 2005; Van den Stock et al., 2011), while others report decreased integration of emotional stimuli (e.g., de Jong et al., 2009). Even though there are contradictory results, Tseng et al. (2015) lean towards suggesting impaired MSI for emotional stimuli as well.

1.3. Aberrant neural correlates in schizophrenia

In schizophrenia, there is a range of different brain regions that display abnormalities (Goghari et al., 2010; Shenton et al., 2001). When studied with magnetic resonance imaging (MRI), there is evidence for ventricular enlargement (e.g., DeLisi et al., 1997; Lieberman et al., 1996); peculiarities in medial temporal lobe structures including amygdala and hippocampus (e.g., Gur et al., 2007; Surguladze et al., 2006; Williams et al., 2007); volume reduction in superior temporal gyrus (STG; e.g., Barta et al., 1990; Kim et al., 2003); parietal lobe atypicality (e.g., Arce et al., 2006); frontal lobe abnormalities (e.g., MacDonald et al., 2005; Taylor et al., 2007); and subcortical brain region deformity such as shape distortion in corpus callosum and volume abnormalities in basal ganglia (e.g., Ebdrup et al., 2010; Frumin et al., 2002; Gur et al., 1998). Additionally, deficits in the visual dorsal stream have been established repeatedly in the schizophrenia population (Koychev et al., 2011).

It has been proposed that it is relevant to look at these neurocognitive

impairments from a perspective of larger scale neural circuits and cortical networks rather than just focus on local neural areas or specific brain regions (Onitsuka et al., 2013). New models are needed to explain the neurological variations in schizophrenia based on higher function, neural connectivity and brain rhythms rather than just anatomy and basic functions. One established model concerning this is the so-called dysconnectivity hypothesis that suggests abnormal interaction between different brain regions, especially a fragmentation of fronto-temporal connection, to be the reason for positive symptoms in schizophrenia (e.g., Friston et al., 2016).

Research on schizophrenia patients with electroencephalography (EEG) reveals aberrant oscillatory activity in several frequencies, including theta/delta bands (Boutros et al., 2008) and alpha/gamma bands (White et al., 2010). For instance, a decrease in synchronized gamma and beta power has been found when exposing the participants to different stimuli (Uhlhaas and Singer, 2013). When investigating neural oscillations in individuals with schizophrenia presented with auditory stimuli, several studies have found deficits in steady-state evoked potentials in gamma frequencies but also to some extent at lower frequency bands (e.g., Krishnan et al., 2009; Kwon et al., 1999). Deficits found in evoked oscillations when schizophrenia subjects processed visual information have been proposed to be a sign of reduced ability to arrange incoming sensory information accurately to oscillatory activity (e.g., Uhlhaas and Singer, 2010).

Kopell et al. (2011) showed that beta-receptor rhythms may mediate how healthy individuals integrate information from different sensory modalities. Together with the suggestion that these beta-receptor rhythms are weakened in the schizophrenic brain (Pittman-Polletta et al., 2015), a possible connection between oscillation alterations in schizophrenia and MSI is put forth. Furthermore, impairment in oscillations could possibly provide explanations for suggested problems with functional connectivity of cortical networks in schizophrenia (Uhlhaas and Singer, 2010).

1.4. Investigation rationale

MSI is associated with specific neural correlates, brain regions and oscillation rhythms (e.g., Cappe et al., 2012b; Keil and Senkowski, 2018). Contemporary behavioral studies indicate impairment in MSI for the schizophrenia population (Tseng et al., 2015), and multisensory impairments are evident in brain-imaging studies (e.g., Jardri et al., 2009; Silbersweig et al., 1995). It is relevant to continue to map disorder-related brain activity while processing multisensory stimuli in order to understand how the pathogenesis of schizophrenia hinders the afflicted from processing incoming sensory information in a coherent and cohesive way (Mäntylä et al., 2018). A greater understanding of how sensory processes are impacted in the brain of the schizophrenia population could be of use in treatment, and several studies suggest that new ways to help diagnose schizophrenia in the future are to be found in different types of brain imaging (e.g., Boutros et al., 2008; Kambeitz et al., 2015; Salvador et al., 2019). Because different brain-imaging techniques can shed further light on how MSI is expressed in schizophrenia, it is relevant to compile the existing literature in this field. It can also be considered that the most valuable studies are those directly comparing how the neural correlates during MSI differ between the population with schizophrenia and the healthy population. Accordingly, the aim of the present study was to perform a systematic review of this research with such direct comparisons of neural activity and connectivity. To our knowledge, there is no previous systematic review addressing this.

2. Method

2.1. Eligibility criteria

The target population was individuals diagnosed with schizophrenia

or schizoaffective disorder (SP) according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) or the International Statistical Classification of Diseases and Related Health Problems (ICD). The control group consisted of healthy individuals without current acute, severe or chronic disease. These healthy controls (HC) constitute the comparative reference on multisensory tests.

We selected a broad definition of multisensory tests based on MSI research to be able to include as many relevant articles as possible. If the study used an acceptable MSI paradigm, where two different sensory modalities were examined at the same time, it was included for further full-article assessment. All different sensory modalities were included (e.g., visual, auditory, olfactory, tactile, and taste). The outcome had to be measured with some type of brain-imaging technique (e.g., EEG, magnetoencephalography (MEG), fMRI).

Studies were included only if peer-reviewed, original articles in English, and accessible to the authors. Studies were excluded if they included a target population with mixed diagnoses, for example schizophrenia and/or schizoaffective disorder combined with other participants with bipolar or psychotic disorders.

2.2. Search strategy and selection process

The review followed the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA), and as suggested by PRISMA (Moher et al., 2009, 2015) the selection phase in this systematic review consisted of four different stages (Fig. 1). First, articles were identified through a database search. Second, the articles with duplicates removed were screened according to the selection criteria concerning title plus abstract. Third, the eligibility of each full-text article was rated according to predetermined criteria. Finally, approved articles were included in the systematic review.

The database search was performed in several different databases: PsychInfo, PubMed and Web of Science. Keywords from relevant articles and reviews were chosen as search terms. Search terms used were (“multisensory” OR “multi-sensory” OR “multimodal” OR “crossmodal”) AND (“integration”) AND (“neural” OR “cortical” OR “ERP” OR “fMRI” OR “EEG” OR “brain” OR “PET” OR “MEG” OR “oscillations”) AND (“schizophrenia” OR “schizo*” OR “schizoaffective” OR “psychosis”). Only peer-reviewed, original articles in English were included and records were searched from the earliest available until January 2020. In total, 317 records were identified through the database search. Two of the authors made the selection of articles, and disagreements were resolved by discussion between all three authors.

Identification was followed by screening in which duplicates were removed and titles plus abstracts were screened to check whether they met inclusion criteria. The reference lists from included articles were also searched and their relevant abstracts were screened for inclusion. Articles that did not meet the inclusion criteria were excluded (e.g., only unisensory or not clearly multisensory, no MSI task, systematic reviews and meta-analysis, animal studies, no diagnosis, no control group, did not measure outcome with brain-imaging techniques or inaccessible). Twenty-nine articles were assessed for eligibility. Two reviewers checked the eligibility of these full-text articles and inclusions of articles were based on consensus. Twelve additional articles were finally excluded. Three articles did not use DSM or ICD for diagnosis and two articles only included first time psychosis. Seven articles were excluded for several reasons, including the actual MSI task not studied with brain imaging (though other tasks were, $n = 1$); stimuli not presented at the same time and therefore not classified as MSI task ($n = 2$); primarily concerned other cognitive functions than MSI leading to difficulties in interpreting the results ($n = 2$); not a primary source ($n = 1$); and not describing any multisensory results ($n = 1$). Thus, 17 articles were included in the systematic review (see Fig. 1).

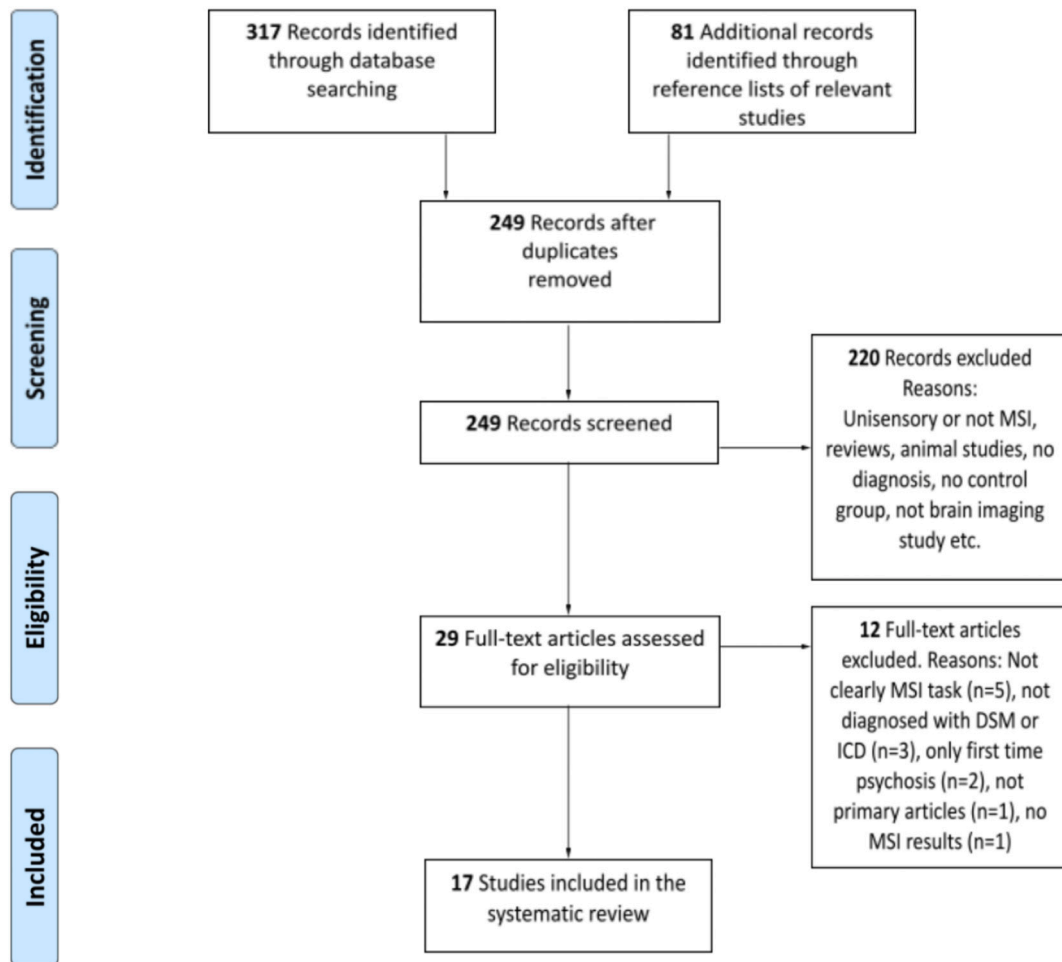


Fig. 1. Flow of information through the different phases of the systematic review.

2.3. Quality assessment

The Newcastle-Ottawa Quality Assessment Scale (NOS; Wells et al., 2012) was used to assess the risk of bias in the results of the 17 included articles (see Appendix 1). The NOS comprises eight points or items grouped into three different sections, which are comparability, selection, and exposure. Each item was rated for a maximal score of one star and the maximal total score was eight. Like other systematic reviews, articles rated less than five out of eight were excluded (e.g., de Dieuleveult et al., 2017). The NOS was chosen since it is often used for systematic reviews and it is considered credible by the authors. Two authors performed the quality assessment of the selected studies, and there were no disagreements.

2.4. Data extraction and analysis of results

The collected studies were primarily grouped based on what brain-imaging technique had been used. Key characteristics in the data were described according to first author, title, publication year, number of participants, cognitive tests, participant characteristics, experiments, brain-imaging analysis, brain-imaging results, and performance when appropriate (see Appendix 2). Two authors made the data extraction and there were no disagreements.

For the analysis of results, the studies were grouped into social or non-social based on type of experiment and into fMRI and EEG/MEG based on brain-imaging technique. The results of how the neural correlates during MSI differ between the schizophrenia population and the healthy population were summarized for each brain-imaging technique.

Since the experimental designs differed, a table (see Table 1) was created that presents the studies according to different conditions compared, brain-imaging results, performance results, unisensory processing, and brain regions.

3. Results

3.1. Quality assessment

The included studies all proved good quality according to the NOS (see Appendix 1). Ten studies had eight out of eight stars, four studies had seven stars and three studies had six stars. Thus, none of the studies had less than five stars and all 17 studies remained included.

3.2. Experimental design

The studies were sorted according to brain-imaging technique, fMRI ($n = 7$) and EEG ($n = 10$), and type of experiment, non-social ($n = 7$) and social ($n = 10$). The included studies compared unisensory to multisensory conditions ($n = 5$) or different multisensory conditions with one another (e.g., congruent vs incongruent, $n = 5$) or a combination of these ($n = 7$). All included studies used audiovisual stimuli only (see Table 1).

Experiments were considered non-social if they did not include a task concerning multisensory stimuli that consisted of speech, faces or other social stimuli ($n = 7$). Of the seven studies labeled as non-social, two used a simple multisensory detection task (Hanlon et al., 2016; Wynn et al., 2014). One study only included the instruction to “look and listen” during presentation of multisensory stimuli (Braus et al., 2002). Four

Table 1

Overview of results. Articles are only presented once for each head category (exception brain regions), and articles that fit into more than one head category are sorted according to their primary results. Participants that are included in several studies (when performing different analysis of the same data) are only counted once in the table when equivalent results are reported.

Results overview		Number of articles		Number of participants (SP; HC)		Number of articles		Number of participants (SP; HC)	
		EEG (<i>n</i> = 10)	fMRI (<i>n</i> = 7)	EEG (<i>n</i> = 189; 191)	fMRI (<i>n</i> = 93; 93)	Non-social (<i>n</i> = 7)	Social (<i>n</i> = 10)	Non-social (<i>n</i> = 147; 147)	Social (<i>n</i> = 136; 137)
Conditions	Unisensory vs multisensory	3	2	65; 63	45; 45	4	1	94; 90	17; 18
	Different multisensory	3	2	53; 53	15; 15	0	5	0	68; 68
	A combination of the above	4	3	71; 75	33; 33	3	4	53; 57	51; 51
Results in neural activity	Sig. group differences	7	6	141; 146	76; 76	5	8	113; 117	104; 105
	No sig. group differences	2	1	48; 45	17; 17	1	2	33;30	32; 32
	Enhanced facilitation in SP	1	0	14; 15	0	1	0	14;15	0
Results in performance ^b	Sig. group differences	1	2	17; 17	33; 33	0	3	0	50; 50
	No sig. group differences	5	3	101; 99	33; 34	4	3	93;90	53; 54
	Enhanced facilitation in SP	2 ^a	0	46; 57	0	2 ^a	0	46;57	0
Unisensory processing	Deficits in SP	7	1	136; 138	17; 17	4	4	86; 87	67; 68
	Not reported	3	6	53; 53	76; 76	3	6	60;60	69; 69
Brain regions	Frontal/temporal	4	7	102; 113	93; 93	3	8	91; 102	104; 104
	Parietal/occipital	5	3	100; 104	60; 60	5	3	113; 116	47; 48

^a Also found lower response accuracy in SP.

^b Not all studies reported performance results.

studies involved audiovisual discrimination tasks (Balz et al., 2016; Sanfratello et al., 2018; Stone et al., 2011, 2014), two of which included manipulations of synchronization of audiovisual stimuli (Sanfratello et al., 2018; Stone et al., 2014). As mentioned in the introduction, it has been shown that visual deficits in SP are associated with deficits in the dorsal stream (i.e., “the where pathway” in the parietal lobe). By placing a multisensory stimulus in the peripheral view, which is the near condition in these experiments, the dorsal stream is engaged and the multisensory processing in the pathway can be recorded (Sanfratello et al., 2018; Stone et al., 2011, 2014). Sanfratello et al. (2018) and Stone et al. (2014) performed different analyses of the same dataset. In other words, they did not perform independent experiments but instead analyzed the same data.

Experiments were considered social if they included a task concerning social multisensory stimuli, like faces, voices or speech perception (*n* = 10). Three of the studies labeled as social used the McGurk paradigm or a version of it with some sort of discrimination task (Liu et al., 2016; Roa Romero et al., 2016a,b) or no specific task (Stekelenburg et al., 2013). One study had an emotional discrimination task based on face-voice stimuli (Müller et al., 2014). Two studies used an audiovisual speech perception paradigm including detection of semantic categories (Szyck et al., 2009, 2013). Three studies investigated multisensory stimuli in the form of speech and gesture of which one used a content judgement task (Wroblewski et al., 2020) and two had a simple task for attentional purposes only (Straube et al., 2013, 2014). Szyck et al. (2009, 2013) derive from the same dataset, as well as the two studies of Straube et al. (2013, 2014).

3.3. Participants

The total number of participants was evenly distributed between the group from the target population and the control group (SP: *n* = 282 and HC: *n* = 284).

Thirteen articles reported patient mean ages between 33.12 and 39.15 years (Balz et al., 2016; Hanlon et al., 2016; Müller et al., 2014; Roa Romero et al., 2016a,b; Sanfratello et al., 2018; Stekelenburg et al., 2013; Stone et al., 2014; Straube et al., 2013, 2014; Szyck et al., 2009,

2013; Wroblewski et al., 2020). One article had a mean age of 25.1 (Braus et al., 2002) and three reported mean ages between 42.21 and 48.8 (Liu et al., 2016; Stone et al., 2011; Wynn et al., 2014). Thus, mean age in the included studies ranged from 25.1 to 48.8.

Five of the studies also included participants diagnosed with schizoaffective disorder (Hanlon et al., 2016; Sanfratello et al., 2018; Stekelenburg et al., 2013; Stone et al., 2014; Wroblewski et al., 2020). In three of these, the exact number of participants with schizoaffective disorder was not reported. However, in the two studies where the distribution is mentioned (Hanlon et al., 2016; Stekelenburg et al., 2013) the participants diagnosed with schizoaffective disorder are few in relation to the participants diagnosed with schizophrenia (SP-AFF: *n* = 3 and SP: *n* = 32).

Fifteen of the studies reported that patients were medicated with antipsychotic medication (Balz et al., 2016; Hanlon et al., 2016; Liu et al., 2016; Müller et al., 2014; Roa Romero et al., 2016a,b; Stekelenburg et al., 2013; Stone et al., 2011, 2014; Straube et al., 2013, 2014; Szyck et al., 2009, 2013; Wroblewski et al., 2020; Wynn et al., 2014).

Eight studies reported mean duration of disorder between 8.24 and 14.33 years (Balz et al., 2016; Liu et al., 2016; Müller et al., 2014; Roa Romero et al., 2016a,b; Stekelenburg et al., 2013; Szyck et al., 2009, 2013). One article focused on individuals with recent onset diagnosed schizophrenia, meaning they were first-episode, neuroleptic-naive and experienced their first psychiatric hospitalization during the experiment (Braus et al., 2002). Five studies specified that the participants were outpatients (Balz et al., 2016; Müller et al., 2014; Roa Romero et al., 2016a,b; Wynn et al., 2014).

3.4. Performance, response times and accuracy

Four studies reported multisensory performance deficits in SP due to lower response accuracy for SP than HC in some conditions (Roa Romero et al., 2016a; Stone et al., 2011) or across conditions (Stone et al., 2014; Wroblewski et al., 2020). Two studies reported longer RTs for SP than HC (Roa Romero et al., 2016a; Straube et al., 2013). Two articles reported greater multisensory facilitation in SP compared to HC when comparing visual unisensory trials with multisensory trials (Stone et al.,

2011, 2014). Eight articles did not find any significant group differences or no indications of group differences in response times (Liu et al., 2016; Müller et al., 2014; Wynn et al., 2014), accuracy (Hanlon et al., 2016; Szyzik et al., 2009, 2013) or perceived illusions (Balz et al., 2016; Roa Romero et al., 2016b). Four articles did not report any performance data (Braus et al., 2002; Sanfratello et al., 2018; Stekelenburg et al., 2013; Straube et al., 2014).

3.5. Correlations with clinical measures

Several studies examined correlations between EEG, MEG or fMRI results and clinical measures. The clinical measures included Positive and Negative Syndrome Scale (PANSS), Scale for the Assessment of Negative Symptoms (SANS), Scale for the Assessment of Positive Symptoms (SAPS), Brief Psychiatric Rating Scale (BPRS), evaluation of cognitive domains with Measurement and Treatment Research to Improve Cognition in Schizophrenia Consensus Cognitive Battery (MATRICS), and medication dose. The outlines of these correlations are presented below. For PANSS, SANS, SAPS and BPRS a mixture of positive, negative and no correlations were found. Two studies found negative correlations, between positive symptoms and multisensory amplitude in the left intraparietal sulcus (IPS; Sanfratello et al., 2018) and between positive symptoms and connectivity to the left inferior frontal gyrus (IFG; Straube et al., 2014). In line with these, Müller et al. (2014) reported negative correlations between P100 amplitude and total BPRS scores. For all three, decreased amplitude or connectivity was associated with more severe symptoms. Wroblewski et al. (2020) found a negative correlation between middle temporal gyrus to superior temporal sulcus (MTG-STs) connectivity and the SANS attention subscale, indicating that decreased connectivity was associated with more severe symptoms. Furthermore, SANS avolition showed positive correlation with RT to auditory, visual and audiovisual stimuli, respectively, and negative correlation with hit rate (HR) to audiovisual stimuli. That is, more severe apathy was associated with longer RTs and lower HRs (Wynn et al., 2014). Several studies did not find any significant correlations with PANSS (Hanlon et al., 2016; Roa Romero et al., 2016a,b; Stone et al., 2011), PANSS Negative (Sanfratello et al., 2018; Straube et al., 2014) or BPRS (Braus et al., 2002; Stekelenburg et al., 2013; Wynn et al., 2014). (A number of non-significant correlations are left out here since it is out of scope for this systematic review.) Stone et al. (2014) reported that the relationship between gamma-band power and MATRICS scores was different for SP compared to HC. For example, correlations between multisensory gamma-band power and MATRICS scores were only found in HC.

Some of the included publications investigated the relationship between their results and medication dosage with the majority reporting no significant correlations (Balz et al., 2016; Liu et al., 2016; Müller et al., 2014; Roa Romero et al., 2016a; Sanfratello et al., 2018; Stekelenburg et al., 2013; Szyzik et al., 2013). The exceptions are positive correlation between dosage and response rate in Stone et al. (2011), and positive correlation between dosage and gamma-band power in Stone et al. (2014).

3.6. EEG and MEG findings

Ten studies recorded neural activity with EEG and MEG, including various experimental designs and aims. Group differences between SP and HC in amplitudes and latencies were evident in a majority of studies. It was mainly reported that SP had reduced amplitudes and/or aberrant latencies compared to HC, indicating deficits in MSI (Balz et al., 2016; Sanfratello et al., 2018; Stone et al., 2011). One study found aberrant neural activity in early event-related potential (ERP) amplitudes and latencies, while late ERPs were only reduced to visual but not multisensory stimuli (Liu et al., 2016). Applying the principles of sub- and superadditivity (i.e., multisensory ERP minus the sum of unisensory ERPs), Stone et al. (2011) found enhanced facilitation in SP while two

studies did not find any group differences (Liu et al., 2016; Wynn et al., 2014). Two studies found no group differences in global field power amplitude, a measure only reflecting effects simultaneously present in a large number of electrodes (Roa Romero et al., 2016a,b).

Oscillatory activity (i.e., theta, gamma and beta-band power) was found to be reduced in Balz et al. (2016) and Roa Romero et al. (2016a), and alpha-band power more strongly suppressed in Roa Romero et al. (2016b). In one study, both increases and decreases were found, mainly concerning greater gamma-band power in SP (Stone et al., 2014). However, this was due to failure in suppressing oscillatory power in all but one condition (i.e., gamma-band power was stronger than both baseline and HC in only one condition). This was the so-called near condition, where stimuli presented in the peripheral view were meant to activate the visual dorsal stream.

When comparing congruent to incongruent conditions of McGurk illusion trials or unisensory to multisensory trials, various interaction effects were found. SP had the opposite neural activity pattern compared to HC (Roa Romero et al., 2016b; Stekelenburg et al., 2013) or diminished differences between conditions (Liu et al., 2016; Roa Romero et al., 2016a). For example, HC displayed larger amplitudes in the congruent trials whereas SP displayed larger amplitudes in the incongruent trials, indicating that SP could not benefit from congruent multisensory information. Müller et al. (2014), on the other hand, only found reduced amplitudes for SP during emotional incongruent trials but no group differences during emotional congruent trials. This indicates that congruent sounds can affect responses and lead to neural responses similar to HC for emotional face-voice stimuli. However, the results from Müller et al. (2014) do not support the notion of SP having altered audiovisual integration since no difference was found between incongruent and congruent stimuli within the group.

Stekelenburg et al. (2013) focused on phonological predictions based on temporal and content information. The authors found that visual temporal information about sound onset did not have the suppression and/or speed-up effect on N100 for SP as it had for HC. These findings corresponded to group differences in multisensory integration. Roa Romero et al. (2016a) investigated crossmodal prediction error (PE) processing when recording theta-band oscillations and found interaction effects indicating that SP had impairments in resolution and in updating of violated predictions. These two studies indicate impairments in prediction processes during MSI in SP.

Seven out of ten articles reported deficits in unisensory processing (Liu et al., 2016; Müller et al., 2014; Wynn et al., 2014; Sanfratello et al., 2018; Stekelenburg et al., 2013; Stone et al., 2011, 2014), primarily in visual processing. Two of these did not find deficits in multisensory integration in SP (Müller et al., 2014; Stone et al., 2011).

Three studies investigated multisensory deficits in the visual dorsal stream, where two studies derive from the same dataset (Sanfratello et al., 2018; Stone et al., 2014). Results indicate deficits in the visual dorsal stream during multisensory processing in SP (Sanfratello et al., 2018; Stone et al., 2014). In line with these results, Liu et al. (2016) reported reduced amplitudes in response to face stimuli in multisensory conditions, indicating that impairments in face processing are also evident in face-voice integration.

Five studies located group differences in multisensory processing in parietal and occipital regions (Balz et al., 2016; Liu et al., 2016; Müller et al., 2014; Sanfratello et al., 2018; Stone et al., 2011), and four studies located differences in frontal/temporal regions (Roa Romero et al., 2016a,b; Stekelenburg et al., 2013; Stone et al., 2014). Source estimations using a linear distributed inverse solution based on a Local Autoregressive Average (LAURA) in Stekelenburg et al. (2013) revealed several abnormalities (i.e., under activation or displaced activation) in neural networks of audiovisual integration. These networks involved auditory cortex, superior temporal gyrus, middle temporal gyrus, and inferior frontal gyrus.

3.7. fMRI findings

There are some findings of subcortical abnormalities (Braus et al., 2002; Szycik et al., 2009) and alterations in cortical areas were discovered in all of the fMRI-studies. The neural abnormalities in the cortical areas were predominantly focused around the STS (Straube et al., 2013, 2014; Szycik et al., 2013; Wroblewski et al., 2020); the STG (Hanlon et al., 2016; Szycik et al., 2009); the IFG (Straube et al., 2013, 2014; Szycik et al., 2009, 2013; Wroblewski et al., 2020); the fusiform gyrus (FG; Szycik et al., 2009; Wroblewski et al., 2020); and the visual dorsal stream (Braus et al., 2002; Hanlon et al., 2016). Two articles found reduced connectivity in both STS and IFG for SP during audiovisual trials (Straube et al., 2014; Szycik et al., 2013), whereas one found reduced connectivity only in the MTG/STS region (Wroblewski et al., 2020) and another suggested dysfunction only in the left IFG (Straube et al., 2013). Yet, it is important to add that the decreased connectivity in the left IFG was mainly found in incongruent trials in Szycik et al. (2013), a condition that was not part of the experiment made by Wroblewski et al. (2020). This could possibly be connected to the results with dysfunctions in abstract, but not concrete, multisensory information processing found in STS, bilateral IFG (Straube et al., 2014), and the left IFG (Straube et al., 2013).

Overall, there are tendencies of lower responsiveness and less activation in the SP group (Braus et al., 2002; Hanlon et al., 2016; Straube et al., 2013, 2014; Szycik et al., 2009, 2013). In three studies, however, the impairments are not as distinct as expected, with many similarities between HC and SP (Straube et al., 2013, 2014; Wroblewski et al., 2020).

Two articles report that impairments in audiovisual integration originate from dysconnectivity in both the STS and the IFG, especially when SP is exposed to incongruent stimuli (Szycik et al., 2013) or when integrating stimuli with abstract content (Straube et al., 2014). Moreover, abnormal MSI processing is explained to derive from deficits in inhibition (Hanlon et al., 2016); impairments in the verbal pathway (Wroblewski et al., 2020); defected speech motor system in the right hemisphere and decreased lateralization of speech functions to the left hemisphere (Szycik et al., 2009); alterations in cortical as well as subcortical activation patterns, mainly focused around high-order frontoparietal cortex and thalamus (Braus et al., 2002); and a failure to activate the left IFG and posterior MTG (Straube et al., 2013). In addition, Straube et al. (2014) suggest that dysconnectivity of left STS and prefrontal cortex could be associated with failure to utilize the hemispheric functions adequately in a context-dependent manner.

3.8. Main findings

The main findings are summarized and presented in Table 2. In short, reduced neural activity and additional aberrant neural response-patterns were evident in SP during MSI tasks in both fMRI and EEG/MEG studies, and distributed across several brain regions and networks. Furthermore, no difference in outcome was ascertained between social and non-social tasks (see Table 1).

Summary of reduced activity in neural correlates during MSI tasks:

- Reduced amplitudes in occipital/parietal ERPs during MSI tasks were revealed when using EEG or MEG (Balz et al., 2016; Liu et al., 2016; Sanfratello et al., 2018; Stone et al., 2011).
- Reduced oscillatory activity was found in gamma/beta-band power in occipital/parietal regions (Balz et al., 2016); in theta-band power in frontal regions (Roa Romero et al., 2016a); and stronger suppression in alpha-band oscillations was observed (Roa Romero et al., 2016b).
- Reduced fMRI-activation was seen both as an overall pattern (Szycik et al., 2009) and in specific brain regions (Braus et al., 2002; Hanlon et al., 2016), for example bilaterally in secondary auditory regions, the default-mode network and the dorsal visual pathway, and in the

Table 2

Main findings in brain-imaging results showing only results reflecting group differences and/or interaction effects in activity, connectivity, oscillations, amplitudes and latencies during multisensory-task recordings. The number of studies is indicated by *n*.

	fMRI	EEG
Reduced activity	<p><i>Brain regions with reduced activity/absence of activation:</i> Dorsolateral prefrontal cortex (<i>n</i> = 1), thalamus (<i>n</i> = 1), STG (<i>n</i> = 1), MTG (<i>n</i> = 1), STS (<i>n</i> = 1), and IFG (<i>n</i> = 1) and overall weaker activation (<i>n</i> = 1)</p> <p><i>Networks with reduced activity/absence of activation:</i> Dorsal visual stream (<i>n</i> = 1), default mode circuit (<i>n</i> = 1)</p> <p><i>Reduced connectivity:</i> STS-MTG (<i>n</i> = 1), STS in general (<i>n</i> = 2), IFG in general (<i>n</i> = 2)</p>	<p><i>Oscillations:</i> Reduced frontal theta-band (<i>n</i> = 1) and occipital gamma-band and beta-band power (<i>n</i> = 1)</p> <p><i>Amplitudes:</i> Reduced early amplitudes in parietal/occipital regions (<i>n</i> = 3) Reduced amplitude in early evoked potentials in temporal-occipital and parietal occipital regions (<i>n</i> = 1) Reduced amplitudes in parietal (i.e., IPS SPG and IPG) and temporal (i.e., STS and STG) regions (<i>n</i> = 1)</p>
Aberrant activity	<p><i>Contrasting patterns to HC between conditions^a:</i> In cuneus (<i>n</i> = 1), insula (<i>n</i> = 1) cingulate gyrus (<i>n</i> = 2), STG (<i>n</i> = 1), STS (<i>n</i> = 1), hippocampus (<i>n</i> = 1), precuneus (<i>n</i> = 1), caudate nucleus (<i>n</i> = 1), FG (<i>n</i> = 1), middle occipital gyrus (<i>n</i> = 1), parahippocampal gyrus (<i>n</i> = 1), transverse temporal gyrus (<i>n</i> = 1), MTG (<i>n</i> = 2), and IFG (<i>n</i> = 2)</p> <p><i>Brain regions with other aberrant activation:</i> Thalamus (<i>n</i> = 1), precuneus (<i>n</i> = 1, MTG (<i>n</i> = 1), STS (<i>n</i> = 1), STG (<i>n</i> = 1), FG (<i>n</i> = 2), nucleus accumbens (<i>n</i> = 1)</p> <p><i>Aberrant connectivity:</i> STS-IFG (<i>n</i> = 1) STS-frontal cortex (<i>n</i> = 1)</p> <p><i>Contrasting patterns to HC between conditions^a:</i> STS-connectivity in general (<i>n</i> = 1) IFG-paracentral lobule connectivity (<i>n</i> = 1) STS-precuneus/cuneus connectivity (<i>n</i> = 1)</p> <p>Diminished IFG-connectivity differences between conditions^a (<i>n</i> = 1)</p>	<p><i>Contrasting patterns to HC between conditions^a:</i> Alpha-band power more strongly suppressed in auditory and/or frontal areas (<i>n</i> = 1) Stronger power and failure to suppress frontal gamma-band power (<i>n</i> = 1) Delayed peak latency in parietal regions (<i>n</i> = 2)</p> <p><i>Diminished differences between conditions^a:</i> In frontal theta-band power (<i>n</i> = 1) In fronto-central, medio-central and parietal amplitudes (<i>n</i> = 3)</p> <p>Under or displaced activation in auditory cortex, STG, MTG and IFG (<i>n</i> = 1)</p>

^a Unisensory vs multisensory, incongruent vs congruent, congruent vs illusory.

right thalamus and prefrontal cortex. In addition, reduced activation was visible during MSI tasks with an abstract content (Straube et al., 2013).

- Reduced connectivity was found in all conditions between STS and frontal cortex (Straube et al., 2014), as well as for the STS and MTG (Wroblewski et al., 2020), and during congruent stimuli for left IFG to a range of different brain regions (Szycik et al., 2013).

Summary of aberrant neural activity during MSI tasks:

- Aberrant neural processes (i.e., amplitudes, oscillations, activity and connectivity) were made evident by various interaction effects where SP often demonstrated contrasting response patterns compared to

HC (Müller et al., 2014; Roa Romero et al., 2016a,b; Stekelenburg et al., 2013; Szyck et al., 2009, 2013).

- In SP, differences between conditions (congruent vs incongruent or illusory) were either diminished (e.g., Stekelenburg et al., 2013) or SP had the opposite neural activity pattern to HC (e.g., Roa Romero et al., 2016a,b; Szyck et al., 2009, 2013), demonstrating aberrant processing of MSI stimuli.
- As mentioned above, SP mainly demonstrated reduced neural responses to conditions meant to elicit multisensory integration processes (e.g., Roa Romero et al., 2016a,b; Stekelenburg et al., 2013; Szyck et al., 2009, 2013). However, in some studies SP only had deficits in the incongruent (Müller et al., 2014) or unisensory (Stone et al., 2011) conditions, indicating normal integration or even enhanced multisensory facilitation.
- This additional aberrant neural activity was associated with both parietal/occipital regions (Liu et al., 2016; Müller et al., 2014) and fronto-temporal regions (Roa Romero et al., 2016a,b; Stekelenburg et al., 2013; Stone et al., 2011). More specifically, it involved the posterior STS (Szyck et al., 2013), left posterior STS (Straube et al., 2014), the FG (Szyck et al., 2009; Wroblewski et al., 2020), and auditory cortex, STG, MTG and IFG (e.g., Stekelenburg et al., 2013).

4. Discussion

The research on MSI in schizophrenia recorded with EEG, MEG and fMRI is still a relatively unexplored, but growing, research topic. In this review, seventeen studies were included and their results lead us to two main insights. First, the activity in neural correlates during MSI is aberrant and mainly reduced in SP compared to HC. Second, these differences are distributed across several brain regions and networks associated with sensory and multisensory processes. However, we could not find any clear connections between experimental design, brain-imaging technique and results in neural activity, making it difficult to draw any further conclusions about how neural and performance outcomes may be modified by experimental design and brain-imaging technique.

4.1. Aberrant and reduced neural responses in SP

In comparing SP to HC, thirteen articles reported significantly reduced neural responses in SP whereas only three reported no differences and one enhanced facilitation in SP. The differences were either due to reduced activity, connectivity, oscillations and ERPs and/or various interaction effects in a majority of the studies. These results are mostly interpreted as altered audiovisual integration at the neural level in SP, which might reflect deficits or anomalies in MSI on performance level (e.g., Balz et al., 2016; Roa Romero et al., 2016b; Sanfratello et al., 2018; Szyck et al., 2009). It is also suggested that SP lack the facilitating or enhancement effect of congruent multisensory information seen in HC (e.g., Roa Romero et al., 2016a,b; Stekelenburg et al., 2013; Szyck et al., 2009, 2013). This means that healthy individuals have different neural responses to congruent compared to incongruent audiovisual information, which is helpful when interpreting and responding to the outside world. The lack of difference in neural response in SP might indicate difficulties both in integrating congruent stimuli and in differentiating between incongruent stimuli. Consequently, the perceptual benefits of MSI are impaired for SP. Aberrant neural activity has been associated with performance deficits in MSI in previous research (Magnée et al., 2009; Sass et al., 2013; for reviews see also Keil and Senkowski, 2018, Tseng et al., 2015, and Wallace et al., 2020). The inability to respond appropriately to audiovisual information could be due to impairments in prediction processing, such as deficits in making predictions about sounds from visual information (Stekelenburg et al., 2013), and/or impairments in prediction error processing (Roa Romero et al., 2016a). This is in line with the integrative framework that Keil and Senkowski (2018) present, in which they theorize that SP have deficits

in generating intersensory predictions and evaluating stimulus congruence as well as deficits in the resolution of incongruence. The internal forward model system illustrates how impairments in prediction error processing can cause positive symptoms like hallucinations in SP. In healthy individuals, because of sensory prediction the response to the sound of their own voice is attenuated in auditory cortex when speaking. This attenuation is not evident in SP, especially for those who report auditory hallucinations (Frith, 2005).

However, the dataset in this review points in several directions and some studies did not find any group differences in neural activity. Stone et al. (2011) even found that multisensory facilitation was enhanced in the SP group. In Liu et al. (2016) aberrant neural activity was found only in early ERPs while late ERPs were only reduced to visual but not multisensory stimuli. Müller et al. (2014) compared the emotional rating of congruent to incongruent face-voice stimuli and found deficits only in the incongruent condition. They concluded that their findings demonstrate that early ERPs (i.e., P100 amplitude) can be affected by congruent sound, but only when the stimuli have emotional valence and not when neutral. Furthermore, Wynn et al. (2014) did not find any difference between groups in ERPs, and Straube et al. (2013, 2014) and Wroblewski et al., 2020 found that SP did not differ from HC in brain activity or connectivity during MSI tasks involving co-verbal gestures with concrete content. Taken together, the studies suggest aberrant and reduced neural activity reflecting deficits in multisensory integration. However, at least under some circumstances, SP seem to have intact MSI with the ability to benefit from audiovisual cues, and maybe even compensate for deficits in visual processing during congruent multisensory stimuli.

4.2. Brain regions and networks associated with deficits in MSI in SP

There are multiple brain regions, both cortical and subcortical, found to be impaired in the different articles included in this review. Our results, however, indicate that the brain regions with aberrant processing during MSI in SP are the STS, the STG, the FG, the MTG and the IFG. Moreover, deficits were seen in the visual dorsal stream, with impairments mainly visible in the IFG and the STS. This is comparable to previous findings, as aberrant processing in SP has been found previously in the IFG (Arce et al., 2006), the MTG (Surguladze et al., 2006), the FG (Onitsuka et al., 2003) and the STG (Barta et al., 1990). The STS is not frequently reported as a region normally associated with the neural deficits of the SP population. However, there are studies showing a dysfunctional pattern in most parts of the brain, and describing the deficits in larger regions (i.e., temporal lobe/cortex) makes it harder to distinguish if the STS is affected or not (Shenton et al., 2001). On the other hand, the STS is a region that has been found to be very important for MSI processing (Desimone and Gross, 1979). There is a possibility that the STS function is mainly reduced in SP when responding to MSI stimuli, which could explain why this is not a region typically found to be impaired.

There are indications that the dorsal visual stream is associated with dysfunction in SP during MSI in the present review (e.g., Sanfratello et al., 2018). Reduced activity in the extrastriate visual cortex, which is part of the dorsal visual stream, has previously been seen in the SP population during perceptual organization of visual information (Sil-verstein and Keane, 2011). Other articles also indicate deficits in the dorsal visual stream for SP (e.g., Butler and Javitt, 2005), seemingly involved during MSI (Kaposvari et al., 2015). This could suggest that these dysfunctions are due to aberrant unisensory processing affecting MSI rather than aberrant multisensory processing in itself. However, in Stone et al. (2014), deficits were seen in the dorsal visual stream for unisensory visual stimuli besides the multisensory deficits, and the authors analyzed whether the unisensory deficits could be directly mapped onto the changes in MSI response. They could not, which indicates that the aberrant processing in the visual dorsal stream during MSI does not emerge directly from deficits in unisensory processing.

In some articles in this systematic review the impairments are located in particular in the left IFG (e.g., [Szyzik et al., 2013](#)), a region also known as Broca's area. One study finds deficits only in the left STS ([Straube et al., 2014](#)), which is part of Wernicke's area. Both Broca's area and Wernicke's area are strongly associated with speech functions and language ability ([Kolb and Whishaw, 2015](#)). The articles in our systematic review showing deficits in these areas all use social multisensory stimuli accompanied with speech. Furthermore, as mentioned in the introduction, the STS bilaterally is thought to be involved in social functions such as both speech perception and biological motion ([Calvert, 2001](#)). This agrees with our findings, since the disturbances in the STS are only visible in the studies using gestures or lip movements as part of the multisensory stimuli. This could indicate that these deficits are more strongly associated with speech perception or language comprehension, or perhaps social visual information, rather than multisensory processing. On the other hand, also mentioned in the introduction, the STS is an important region for multisensory processing. In addition, the left IFG, together with its right homolog, has proved to be of importance specifically for audiovisual integration of speech in healthy individuals ([Curcic-Blake et al., 2013](#); [Ojanen et al., 2005](#); [Pekkola et al., 2005](#)). This makes it reasonable to believe that our results do indicate reduced MSI.

Several of the articles find deficits in the IFG bilaterally (e.g., [Szyzik et al., 2009](#)). The right IFG also takes a great part in functions involving attention and inhibition ([Hampshire et al., 2010](#)), and both attention and inhibition are known to be impaired in SP (e.g., [Arce et al., 2006](#); [Nuechterlein et al., 2015](#)). However, when using experiments where the paradigm is designed to test automatic audiovisual integration processes, not dependent on attention, deficits are still visible for SP ([de Gelder et al., 2003](#)). Regarding inhibition, [Hampshire et al. \(2010\)](#) argue that the right IFG is more likely to be part of a network that tunes in on task-relevant stimuli rather than alone responsible for inhibitory control. They further conclude that the right IFG has a far greater role than just inhibition. Our systematic review opens up to the possibility that this greater role could involve multisensory processing.

Three areas additionally observed to be involved in MSI are the FG, the MTG and the STG ([Pehrs et al., 2014](#); [Surguladze et al., 2006](#)), which all have been indicated to be impaired in this systematic review. All of these are also found to be associated with proneness to hallucinations and the production of them ([Kim et al., 2003](#); [Kunzelmann et al., 2019](#); [Zhang et al., 2017](#)). As previously mentioned, the link between multisensory disturbances and symptoms like hallucinations or distorted reality is commonplace ([Frith, 2005](#)). The positive symptoms in schizophrenia are often explained with the dysconnectivity hypothesis (e.g., [Friston et al., 2016](#)), and dysconnectivity was found in several of the articles. They included dysconnectivity both in and between multiple brain regions throughout all of the lobes, including dysconnectivity in the parietal cortex. Moreover, decreased parietal connectivity has been demonstrated to lead to decreased MSI facilitation ([Brang et al., 2013](#)). Possibly, there is a link between the dysconnectivity typical of the positive symptoms in SP and the evident dysconnectivity seen with decreases in multisensory facilitation. The correlations with clinical measures found in this systematic review further indicate this link as positive symptoms were negatively correlated to MSI amplitudes in one article ([Sanfratello et al., 2018](#)) and to connectivity in IFG in another ([Straube et al., 2014](#)). Furthermore, reduced neural activity was associated with more severe symptoms in several studies (e.g., [Müller et al., 2014](#); [Wroblewski et al., 2020](#)).

4.3. Strengths and limitations

This systematic review followed the PRISMA guidelines in the aspiration to ensure a fair and reliable selection process and to improve the quality of the systematic review ([Moher et al., 2009, 2015](#)). The literature search was made in the three databases PubMed, PsycInfo and Web of Science with multiple search terms and a broad definition of MSI, and all possible brain-imaging techniques were included. Furthermore, all

included articles had to meet pre-determined quality standards according to the NOS ([Wells et al., 2012](#)). Two reviewers checked the eligibility of the 29 full-text articles that emerged from the first screening, and the ultimate inclusion of 17 articles was based on consensus. Even though these measures were taken to find as many relevant articles as possible, some articles investigating MSI in schizophrenia with brain-imaging techniques might not be included. This could be due to search limitations such as time constraints, overlooked relevant search terms, and/or articles published after the search period. It could also be due to researchers not labeling their experiments as MSI-related even though they might be. It is possible that the unintended exclusion of relevant search terms could have affected our findings. One possible way to address this in future reviews could be to add search terms like “sensory”, “illusion” and “speech” so that articles not labeled as MSI-related could be included in the initial search process. The definition of MSI chosen by the authors, also affecting the inclusion process, could be categorized as too wide or too narrow compared to other definitions. Since the identification of articles in the three databases was divided between the two reviewers, there is a risk of bias in the selection process. If the reviewers had performed the identification and screening of articles in all three databases individually to make sure the same articles were included for assessment of eligibility, it would have decreased the risk of bias and the risk of missing relevant articles.

Since this review focuses on an emerging research field, it is based on a rather limited sample with 17 included articles. Furthermore, because the dataset is the same for some of the articles the number of included participants is even smaller than what perhaps could be expected (SP: $n = 282$; HC: $n = 284$). There are no optimal numbers of included articles in a systematic review and the number of relevant articles is highly dependent on the size of the research field. However, with 17 included studies the risk of overlooking relevant articles is that just one article could affect the results and conclusions of the systematic review, whether they point in the same direction or not. Furthermore, a limited number of participants increase the risk of making a type II error (i.e., not finding a difference that would be visible in a larger sample). This risk is even greater with greater variance in group characteristics. Even though all groups from the included articles were evaluated within the NOS ([Wells et al., 2012](#)), and considered equal, there is still a possibility for some variance. For example, one study included neuroleptic-naïve first-episode schizophrenic patients while others included patients with illness durations up to 16 years.

A difficulty in this systematic review has been to compile and interpret the included studies that are quite heterogeneous. The experimental designs differ greatly due to different focuses in aim. For example, some of the studies focus on abstract thinking ability rather than MSI per se, some compare unisensory to multisensory conditions and others incongruent versus congruent. Therefore, there is a possibility that the aberrant neural activity is better explained by differences in experimental design, for example complexity of the task or if the task was social or non-social. However, we could not find any connections between results in neural activity and type of task. In our results, both social tasks and non-social tasks were associated with differences in neural processing between SP and HC. This further strengthens the assumption that MSI processing is aberrant in SP, since it cannot be explained with the tasks having social elements. Still, it is possible that there are connections in more detailed measures between different tasks and activity or connectivity, but that this is not evident in the present review. Furthermore, it could be that the study sample is too small and too versatile to find underlying experimental factors behind the outcomes. There is a possibility that these differences will become evident when more research is available. For example, underlying factors could be how MSI is defined and computed, differences in stimulus characteristics and/or analytic methods, and how conditions are compared.

A further limitation is that schizophrenia is a heterogeneous disorder ([Mohr et al., 2004](#)). For example, not all individuals diagnosed with schizophrenia show catatonic behavior, even though it is a symptom of

the disease. Therefore, it is a possibility that the inconsistencies showing in this systematic review are a consequence of that heterogeneity. It could be that not all individuals with schizophrenia have impairments in MSI, at least not to the same extent. Furthermore, some of the individuals with schizophrenia could have found ways to compensate. Due to the differences in design, we cannot draw any conclusions concerning under what specific conditions MSI is affected in schizophrenia or its behavioral manifestation. However, the combined experimental designs in these studies mirror how people come across MSI in their everyday life, increasing external validity and the possibility to generalize results. The variety in group characteristics and the range of different types of design are both representative of reality, clarifying that the differences we can see from this systematic review are evident for several types of conditions and across groups.

The aim of this systematic review was to compile and analyze evidence of MSI impairments in schizophrenia using performance measures. However, although aberrant activity was a prominent finding in this review, it was not clearly reflected in the performance measures. Five out of the thirteen articles that reported performance measures found impairments in response accuracy, RTs or both (Roa Romero et al., 2016a; Stone et al., 2011, 2014; Straube et al., 2013; Wroblewski et al., 2020). In six studies, differences in neural activity were found, but no difference in RTs, accuracy or perceived illusions (Balz et al., 2016; Hanlon et al., 2016; Liu et al., 2016; Müller et al., 2014; Roa Romero et al., 2016b; Szycik et al., 2009, 2013). (Note that Szycik et al., 2013, report the same performance results as Szycik et al., 2009, and that these two are counted here as a single study.) Even though behavioral performance was not the focus of this review, the authors believe it is important to reflect upon this result. It is possible to argue that it affects the validity and importance of the findings negatively. What conclusions can we draw from differences in neural activity between SP and HC when these are not reflected in expected behavioral differences?

In this systematic review, all different sensory modalities were included in the search process, but only audiovisual studies were included. This indicates that there is a majority of studies on neural correlates during audiovisual integration, and that the integration of other senses might be somewhat neglected in the empirical research. As mentioned in the introduction, it is well known that schizophrenia is associated with impairments in several sensory modalities. To understand the impairments in MSI fully, the associated neural correlates and clinical consequences, it is important to investigate other sensory modalities (e.g., tactile, olfactory, taste). Additionally, it is important to investigate impairments in multisensory integration of bodily signals, somatosensory input, especially since somatosensory impairments have been theorized to explain the “self-disorders” in schizophrenia (Postmes et al., 2014). Self-disorder is the difficulty to distinguish between self and other, which may cause passivity symptoms characteristic of schizophrenia (Chen et al., 2015). Multisensory impairments probably affect both environmental and bodily sensory signals in schizophrenia.

Appendix 1. NOS quality assessment

Quality assessment criteria									
Selection (4 stars)					Comparability (2 stars)		Exposure (2 stars)		
Study	Is the population definition adequate?	Representativeness of the population	Selection of controls	Definition of controls	Study controls for schizophrenia	Study controls for additional factor	Same method of ascertainment for cases and controls	Non-response rate	Quality score (max. 8 stars)
	Schizophrenia or schizoaffective disorder according to DSM or ICD	Representative of the schizophrenia population	Same community	Healthy individuals with no current or	Diagnosed schizophrenia compared to healthy controls	Gender, age, education/ cognitive functioning (e.g., IQ)	Methods are identical	Same rate for both groups	

(continued on next page)

4.4. Conclusions and further directions

This is the first systematic review examining how the neural correlates during multisensory integration differ between individuals with schizophrenia and the healthy population. We found aberrant and reduced neural activity measured with EEG, MEG and fMRI, presumably reflecting deficits in multisensory integration in schizophrenia. This was evident in several brain regions involving multisensory integration, mainly in the temporal cortex (i.e., STS/STG, MTG and FG) as well as in frontal (i.e., IFG) and occipito-parietal (i.e., the dorsal stream) regions. However, results indicate that MSI could be intact during some conditions. Furthermore, difference in neural activity was not always reflected in performance deficits.

Further research could investigate the possibility that MSI is reduced in schizophrenia patients only during some conditions to find out the underlying mechanism(s) triggered by those specific conditions. Since this systematic review did not see any clear differences in neural correlates in schizophrenia patients when performing an audiovisual task categorized as social instead of non-social, it would be interesting to investigate this further. Moreover, clarifying the role of other sensory modalities like the tactile modality, and the integration of somatosensory input in schizophrenic self-disorders seem justified. The relationship between deficits in unisensory and multisensory processing is unclear and future research ought to continue to determine how they are related. If some individuals are able to compensate for deficits, it would be beneficial for others who suffer impairments in MSI if future research would investigate this topic further.

CRedit authorship contribution statement

Cornelia Gröhn: Conceptualization, Formal analysis, Investigation, Writing – Original Draft.

Elin Norgren: Conceptualization, Formal analysis, Investigation, Writing – Original Draft.

Lars Eriksson: Conceptualization, Methodology, Writing – Review & Editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

A version of the present study was presented in a Master's thesis (Norgren and Gröhn, 2020), with only abstract published, at Karlstad University, Karlstad, Sweden.

(continued)

Study	Quality assessment criteria								Quality score (max. 8 stars)
	Selection (4 stars)				Comparability (2 stars)		Exposure (2 stars)		
	Is the population definition adequate?	Representativeness of the population	Selection of controls	Definition of controls	Study controls for schizophrenia	Study controls for additional factor	Same method of ascertainment for cases and controls	Non-response rate	
	Schizophrenia or schizoaffective disorder according to DSM or ICD	Representative of the schizophrenia population	Same community	Healthy individuals with no current or chronic major disease	Diagnosed schizophrenia compared to healthy controls	Gender, age, education/cognitive functioning (e.g., IQ)	Methods are identical	Same rate for both groups	
				chronic major disease					
Balz et al., 2016	*	*	*	*	*	*	*	*	8
Braus et al., 2002	*	–	*	*	*	*	*	*	7
Hanlon et al., 2016	*	*	*	*	*	*	*	*	8
Liu et al., 2016	*	*	*	*	*	–	*	*	7
Müller et al., 2014	*	*	*	*	*	*	*	*	8
Roa Romero et al., 2016a	*	*	*	*	*	*	*	*	8
Roa Romero et al., 2016b	*	*	*	*	*	*	*	*	8
Sanfratello et al., 2018	*	*	*	*	*	*	*	*	8
Stekelenburg et al., 2013	*	*	*	*	*	*	*	*	8
Stone et al., 2011	*	*	*	*	*	*	*	*	8
Stone et al., 2014	*	*	*	–	*	–	*	*	6
Straube et al., 2013	*	*	–	*	*	–	*	*	6
Straube et al., 2014	*	*	–	*	*	–	*	*	6
Szycik et al., 2009	*	*	*	*	*	–	*	*	7
Szycik et al., 2009	*	*	*	*	*	–	*	*	7
Wroblewski et al., 2020	*	*	*	*	*	*	*	*	8
Wynn et al., 2014	*	*	*	*	*	*	*	*	8

Appendix 2. Included articles

Title, first author (year of publication)	Study characteristics					Experiment	Brain imaging analysis	Main results and major conclusions	
	Participants (number; female/male; age interval; mean (SD))	Clinical and cognitive assessment (Test; mean total score)		Controlled/matched for	Tasks and instructions				
	Schizophrenia/ Schizoaffective	Healthy controls	Subjects	Control group	Conditions	Brain imaging technique			
Balz et al. (2016)	15; 3/12; -; 33.87 (7.23)	15; 3/12; -; 36.13 (7.91)	DSM-IV-TR, PANSS; 74.07 (8.81), BACS; 248.27 (41.65)	SCID non-patient edition, BACS; 259.60 (36.89)	Age, gender, education, handedness, normal hearing, normal or corrected-normal vision, no neurological disorders, alcohol or substance abuse. 40%	Sound-induced flash illusion (SIFI) Instructions: Report the number of perceived visual stimuli by pressing a button with the index, middle,	1) AOV1 2) AOV2 3) A1V1 4) A2V0 5) A2V1 6) A2V2 Numbers denote the number of auditory (A) and visual (V) inputs. A = tone. V = flash.	EEG; ERP “ERPs were analyzed in a time interval from 0 to 400 ms following stimulus onset. Non-parametric tests, separately between the different percepts,	“Analysis of ERP data revealed multisensory integration effects in SCZ compared to HC. The study complements recent reports of altered

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics						Brain imaging analysis	Main results and major conclusions		
	Participants (number; female/male; age interval; mean (SD))		Clinical and cognitive assessment (Test; mean total score)		Controlled/ matched for	Experiment				
	Schizophrenia/ Schizoaffective	Healthy controls	Subjects	Control group	Tasks and instructions	Conditions			Brain imaging technique	
					underwent a multi-drug screening, all negative results.	or ring finger of their right hand to indicate whether they perceived 0, 1, or 2 flashes, respectively.	between groups, and for the Perception by Group interaction. To test differences between the processing of multisensory and unisensory stimuli, visual trials, as well as auditory trials were analyzed. Non-parametric tests between conditions, between groups and for the Condition by Group interaction. To test if there are differences between the processing of multisensory and unisensory stimuli, an analogous analysis was carried out for unisensory visual stimuli with the within-subject factor Condition and the between-subjects factor Group."	multisensory integration in the McGurk illusion, suggesting processing differences between SCZ and HC in various prominent multisensory illusion paradigms."		
Braus et al. (2002)	12; 6/6; -; 25.1 (4.8); school education, 10.6 ± 1.8 years,	11; 5/6; -; 29.4 (6.2); school education, 12.0 ± 1.4 years.	DSM-IV; ICD-10; BPRS 49.9 (5.7). First psychiatric hospitalization.	-	Age, gender, education, handedness, no substance abuse. For HC: no history of significant medical or psychiatric illnesses For SP: No history of neurologic disorders, no abnormalities showing in neurologic examination. None of the participants had ever received neuroleptic medication before being enrolled in the study. All patients were neuroleptic-naive, having their first schizophrenic episode with the predominance of delusions,	Audiovisual integration/ detection task. Instructions: "look and listen".	1) Simulation - simultaneous presentation of a visual stimulus (flickering checkerboard) and auditory stimulus (drumbeats) 2) Fixation	fMRI	"General linear model, 2-stage analysis procedure contrast. Data were analyzed by modelling the different conditions as boxcar functions convoluted with the hemodynamic response function. The contrast was +1 for AV stimulation and -1 for the fixation condition. Data were analyzed by including the contrast images of all subjects of each group into a second-level group analysis. One-sample <i>t</i> -tests were	"The study supports the concept of widespread cortical and subcortical deficits in schizophrenia. The findings suggest abnormal functioning early in the information processing and in high-order association cortices already at illness onset, before the administration of medication or the most confounding effects of illness duration. The main regions have been implicated in visual motion

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics				Controlled/ matched for	Experiment			Brain imaging analysis	Main results and major conclusions
	Participants (number; female/ male; age interval; mean (SD))	Clinical and cognitive assessment (Test; mean total score)				Tasks and instructions	Conditions	Brain imaging technique		
	Schizophrenia/ Healthy Schizoaffective controls	Subjects	Control group							
					hallucinations, and distrust. Urine screening for drug usage was negative. All stabilized patients were re-diagnosed 6 months after the initial examination, and the diagnosis was stable in all cases.			performed to test for in-group correspondences and 2-sample t-tests to identify areas that are less active in SP compared with HC."	perception and discrimination as well as in attention to sensorial events and perceptual synthesis."	
Hanlon et al. (2016)	33; 6/27; -; 38.0 (13.66); (12.61 ± 1.77 years of education). Does not say how many participants were diagnosed with SP or SP-AFF.	34; 6/28; -; 37.50 (13.35); (12.74 ± 1.83 years of education)	DSM-IV; SCID; PANSS 58.67 (13.71); WTAR- 55.08 (12.96); Calgary Depression Scale 3.30 (3.87); FTNT 0.55(1.00); Olanzapine equivalent 11.77 (7.48); SAS 1.30 (1.59); AIMS 1.18 (2.16); BAS 0.24 (0.50); UPSA	WTAR 58.14 (7.36) as an estimate of pre-morbid intelligence	Age, gender, education, no history of neurological disorder, head trauma with loss of consciousness greater than 5 min, no active substance dependence or abuse (except for nicotine) within the past year, no lifetime history of dependence or use within the last 12 months of hallucinogens, amphetamines or cocaine, or intellectual disability. All abstained from smoking and caffeine intake for at least 1 h before. For HC: no history of a psychotic disorder in a first-degree relative, a current or past psychiatric disorder (with the exception of one lifetime depressive episode), depression or antidepressant use within the last 6 months, or lifetime antidepressant use of more than 1 year. For SP: all patients were taking stable doses of antipsychotic medications. All were drug tested.	Audiovisual detection task Instructions: Participants pressed a button with their right index finger to indicate target detection	1) visual blue box and auditory tone presented simultaneously (2000 Hz). 2) White fixation cross on black background and quiet	fMRI Group × Time voxel-wise linear mixed-effect ANCOVA was used to examine group differences. Time: 0 to 20s post-stimulus onset.	"During a simple multisensory detection task, SP exhibited abnormal PSU after normal task-related activation in secondary unisensory areas, as well as abnormal task-induced deactivation in the aDMN, and the two hemodynamic abnormalities were related. Current results are consistent with the view that separate neural processes underlie the two phases of the HRF and that they are differentially affected in SP."	
Liu et al. (2016)	17; 1/16; -; 16.9 (9.66)	18; 1/17; -; 44.6 (7.94)	SCID, Axis I and II. WAIS full scale; 92.1 (11.15) PANSS; 80.50 (30.79).	SCID, Axis I and II. WAIS full scale; 102 (14.40)	Age, gender, handedness, normal hearing, vision, no history	Face-voice discrimination task Instructions: Press a	1) Visual stimuli of 36 human and 3 primate neutral faces. 2) Auditory stimuli	EEG; ERP "The amplitude and latency of P100, N170 and P270 were	"There were no group differences in the multisensory	

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics						Brain imaging analysis	Main results and major conclusions	
	Participants (number; female/male; age interval; mean (SD))		Clinical and cognitive assessment (Test; mean total score)		Controlled/ matched for	Experiment			
	Schizophrenia/ Schizoaffective	Healthy controls	Subjects	Control group	Tasks and instructions	Conditions			Brain imaging technique
			Onset age, 26.2 (9.25) Illness duration, 16.7 (9.19) Antipsychotic medication, 17		of neurological illness, no history of DSM-IV diagnosis of drug or alcohol abuse, parental socioeconomic status, verbal IQ above 75, no hearing, vision or upper body impairment. For HC no history of psychiatric disorder in oneself or first-degree relatives.	response button to a monkey face, press a button to a monkey face-voice	of 36 neutral human sounds and 3 primate digitized voices. 3) Audiovisual stimuli of 36 congruent human face-voice stimuli and 9 primate combinations.	separately submitted MANOVA with group as between-subjects factor, and condition and electrode as within-subjects factors. The amplitude and latency of fronto-central N100 and P200 were separately submitted to MANOVA with group as between-subjects factor, and condition, region as within-subjects factors. To assess the differences in multisensory integration between groups the difference ERP waveforms were obtained where the 'sum' ERPs were subtracted from the 'multimodal' ERPs for each group. Then, between-group comparisons of the difference ERP waves were conducted with time point-by-time point independent samples <i>t</i> -tests."	effects as tested with the time point analysis. Differences emerged in the traditional ERP analyses: while the N250 was more negative in the audiovisual relative to the visual condition in both groups, the P400 was reduced to faces but not to face-voice information in the schizophrenia group. In group comparisons, the face processing difficulty existed both to face alone and face-voice stimuli and extended to a lack of facilitatory effects of face-voice stimuli on P270 latency in the schizophrenia group."
Müller et al. (2014)	15; 4/11; -; 35.1 (9.26)	15; 3/12; -; 40.8 (10.67)	SCID, BRPS; 46.93 (9.68), SAPS; 29.80 (21.6), SANS; 37.80 (13.3), Illness duration, 14.33 (9.12) Antipsychotic medication, 12	Age, gender, education, handedness, normal or corrected-normal vision, tested negative on a urine drug screening. For HC: no history of neurological or psychiatric disorder, including substance abuse, or any history of an Axis I disorder in a first-degree relative. For SP: No comorbid psychiatric or neurological	Audiovisual emotion discrimination task. Instructions: Participants were instructed to ignore the sound and blurred face and to just rate the expression of the clear face. The task was to respond as fast and accurately as possible as soon as the eight-point rating scale was displayed. The rating	Visual stimuli (V): a face demonstrating happy (H), fear (F) or neutral (N). Auditory stimuli (A): a voice demonstrating happy (H), fear (F) and neutral (N) 1) HV + HA 2) HV + FA 3) HV + NA 4) FV + HA 5) FV + FA 6) FV + NA 7) NV + HA 8) NV + FA 9) NV + NA	EEG; ERP	"ERP sounds: MANOVA with the independent variables emotional expression of the sound and group and the dependent variables amplitudes of P1, N1 and P2 at Cz was calculated to test for group differences. ERPs visual: MANOVA with the independent variables facial emotional expression, side, caudality and group and the	"In summary, this study demonstrates deficits in early visual face processing in schizophrenia, which may be counteracted or attenuated by concurrently presented emotionally congruent sounds."

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics						Brain imaging analysis	Main results and major conclusions		
	Participants (number; female/male; age interval; mean (SD))		Clinical and cognitive assessment (Test; mean total score)		Controlled/ matched for	Experiment				
	Schizophrenia/ Schizoaffective	Healthy controls	Subjects	Control group	Tasks and instructions	Conditions			Brain imaging technique	
Roa Romero et al. (2016a)	17; 5/12; -; 35.24 (7.73)	17; 4/13; -; 36.00 (8.29)	DSM-IV and ICD-10, PANSS; 72.41 (5.80), BACS; 245.65 (43.31), illness duration 8.24, (4.47) Antipsychotic medication, 17	SCID, BACS; 273.47 (37.37)	Age, gender, education, handedness, normal hearing and normal or corrected to normal vision, no record of neurological disorder, no alcohol or substance abuse, 45% underwent multidrug-screening test.	Crossmodal prediction error processing, discrimination task. Instructions: The participants' tasks were to respond to the occasional auditory syllable /fa/ and to an occasional change of the fixation cross, which turned into a white circle.	1) High predictive and congruent (e.g., /Pa/ + /Pa/) 2) high-predictive and incongruent (e.g., /Pa/ + /La/) 3) low-predictive and congruent (e.g., /Ga/ + /Ga/) 4) low-predictive and incongruent (e.g., /Ga/ + /Ta/)	EEG; ERP and oscillatory activity	“Analysis of event-related activity focused on global field power (GFP). The analysis of oscillatory responses focused on oscillatory power to high- and low-predictive congruent and incongruent trials. We focused the analysis on group differences with respect to the PE, which was expressed as the difference between neural responses to incongruent vs. congruent trials. The differences between congruent and incongruent syllables were entered into running repeated-measures ANOVAs with the factors Group and Predictability. Running ANOVAs for GFP data were computed for each sample point in a 0 to 500 ms interval following auditory syllable onset. Cortical sources of main effects or	“The study demonstrates a crossmodal PE deficit in SP. In auditory areas we observed similar processing of audiovisual stimulus incongruence in SP and HC. The key novel finding is that SP lack a crossmodal PE processing-related enhancement of frontal theta-band oscillations. The reduced frontal theta-band power presumably reflects a top-down crossmodal processing deficit”.

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics						Brain imaging analysis	Main results and major conclusions	
	Participants (number; female/male; age interval; mean (SD))		Clinical and cognitive assessment (Test; mean total score)		Controlled/ matched for	Experiment			
	Schizophrenia/ Schizoaffective	Healthy controls	Subjects	Control group	Tasks and instructions	Conditions			Brain imaging technique
Roa Romero et al. (2016b)	21; 4/10; -; 35.57(6.55)	21; 4/10; -; 36.79 (7.18)	DSM-IV-TR, ICD-10, PANSS; 74.50 (7.43), BACS; 248.71 (43.76). Illness duration, 9.50 (5.63) Antipsychotic medication, 14	SCID non-patient edition, BACS; 266.50 (39.70)	Age, gender, education, handedness, normal hearing, corrected vision, no neurological disorder, alcohol or substance abuse. Multi drug screening; none positive (45% of participants). For HC: No psychiatric disorder. For SP: No other psychiatric disorder.	McGurk Illusion Instructions: Indicate by button press with the index, middle, ring or small finger of their right hand whether they had perceived the syllable /pa/, /ga/, /ka/, "or something else" respectively	1) Congruent syllables (/pa/, /pa/). 2) Incongruent syllables (/pa/, /ka/). 3) McGurk illusory syllables (/ga/, /pa/)	EEG; ERP and oscillatory activity	interactions in the analysis were visualized by a local autoregressive average (LAURA) analysis. Oscillatory power was analyzed for a region of interest (ROI) encompassing 10 fronto-central electrodes. Running ANOVAs with the same factors as in the GFP analysis were computed for each sample point and frequency in a 0 to 1000 ms interval following auditory syllable onset" "The analysis of EEG data focused on the comparison of ERPs and oscillatory responses to McGurk illusion and congruent trials. Examined ERP amplitudes and oscillatory power at a medio-central (ROI), comprising 16 channels. For ERPs, GFP, as well as oscillatory power running 2 × 2 ANOVAs with the factors Group and Condition were conducted for each sample point. ERPs were analyzed in a time window from 0 to 500 ms and oscillatory responses from 0 to 850 ms following auditory syllable onset. Significant main effects or interactions were followed-up by <i>t</i> -tests".

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics					Brain imaging analysis	Main results and major conclusions			
	Participants (number; female/ male; age interval; mean (SD))		Clinical and cognitive assessment (Test; mean total score)		Controlled/ matched for			Experiment		
	Schizophrenia/ Schizoaffective	Healthy controls	Subjects	Control group	Tasks and instructions			Conditions	Brain imaging technique	
Sanfratello et al. (2018)	53; 9/44; -; 39.5 (13.6)	56; 17/39; -; 36.3 (12.0)	SCID-IP, PANSS Positive 14.8 (4.9), PANSS Negative 14.9 (5.1), UPSA, IQ; 101.2 (21)	SCID-NP, IQ; 116.5 (13)	Age, gender, education, no history of neurological disorder, significant head trauma, no current diagnosis of substance abuse. For HC: no history of psychiatric disorder and no first-degree relatives with a history of a psychotic disorder.	Audiovisual discrimination task. Instructions: Participants were asked to indicate whether the stimulus was near to them or far from them by pressing one of two buttons on a response device with either their right index finger or right middle finger, respectively	1) Near visual stimulus. 2) Far visual stimulus. 3) Near auditory stimulus. 4) Far auditory stimulus. 5) Near audiovisual synchronous stimulus. 6) Far audiovisual synchronous stimulus. 7) Near audiovisual asynchronous stimulus. 8) Far audiovisual asynchronous stimuli	MEG; ERP	“The ROIs investigated were bilateral: occipital pole, insula, STS, STG, SPG, IPG, and IPS. A 3-way mixed ANOVA was conducted, for both amplitude of activation and for peak amplitude latency, with group entered as a between-subjects factor, and hemisphere, location of stimulus, and presentation of stimuli entered as within-subjects factors. A hierarchical regression analysis was performed with RT as the independent variable; amplitude was entered first, followed by group, and the interaction term last. A second hierarchical regression analysis was performed for latency, with RT as the independent variable; latency was entered first, followed by group, and the interaction term last.”	“Our work here provides evidence of significant brain differences in SP vs HC in parietal areas, e.g. areas involved in multisensory integration, particularly IPS. Our results provide additional evidence of a dorsal visual stream impairment in SP.”
Stekelenburg et al. (2013)	18; 1/17; -; 38 (9)	18; 1/17; -; 39 (8.1)	DSM-IV-TR, M. I.N.I, BPRS; 41.4 (11.5). Illness duration, 16.2 (5.6) Antipsychotic medication, 18		Age, gender, education, handedness, normal hearing and normal/corrected vision, no neurological illness, no alcohol or drug dependence or abuse within last year, or long duration of past abuse, no history of electroconvulsive treatment, no medications	Audiovisual integration based on temporal and content information. Instructions: No instructions were given	1) Visual only, one of the videos without sound. 2) Auditory only, one of the four sounds against a black background. 3) Audiovisual congruent, video with synchronized original sound. 4) Audiovisual incongruent, auditory /fu/ combined with visual /bi/, auditory /bi/	EEG; ERP	“N1 and P2 amplitude and latency scores were separately entered in a repeated measures MANOVA with the between-subject factor Group, the within-subject factors Stimtype and Modality. An additional analysis involved the spatiotemporal dynamics of the	“To conclude, we found clear evidence that patients with schizophrenia show degraded audiovisual integration at the neural level. From a functional perspective, patients with schizophrenia may have difficulties in predicting sound onset and

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics					Experiment			Brain imaging analysis	Main results and major conclusions
	Participants (number; female/male; age interval; mean (SD))		Clinical and cognitive assessment (Test; mean total score)		Controlled/matched for	Tasks and instructions	Conditions	Brain imaging technique		
	Schizophrenia/ Schizoaffective	Healthy controls	Subjects	Control group						
					which would affect EEG.		with visual /fu/, auditory hand clapping with visual tapping of a spoon, and auditory tapping of a spoon with visual hand clapping		difference of AV interaction between the two groups by conducting point-by-point two-tailed <i>t</i> -tests at each electrode in a 1–300 ms window. In the first analysis the intersensory effect was investigated. We computed the (AVC – [A + V]) difference wave for both groups and performed the running <i>t</i> -tests between both difference waves. In the second analysis we examined the AV congruency effect by comparing the (AVC – AVI) difference wave between groups. As a final step, the neural sources underlying audiovisual interactions were estimated by using a linear distributed inverse solution based on a Local Auto-Regressive Average (LAURA) model of the unknown current density in the brain.”	content from vision.”
Stone et al. (2011)	14; 3/11; -; 42.21 (12.69)	15; 5/10; -; 37.80 (13.81)	WASI FSIQ; 98.29 (19.06), PANSS: 54.29 (15.79). Antipsychotic medication, 14	WASI FSIQ; 113.27 (13.47)	Age, gender, education, ethnicity.	Audiovisual discrimination task Instructions: Participants were asked to indicate whether the stimulus was near to them or far from them by pressing one of two buttons on a response device with either their right index finger or right middle finger, respectively	1) Near visual stimulus in the peripheral view. 2) Far visual stimulus in the central view. 3) Near auditory 80 dB sound. 4) Far auditory 64 dB sound. 5) Near multisensory, visual stimuli and auditory stimuli presented 0-5 ms later (appears simultaneously). 6) Far multisensory visual stimuli and auditory	EEG; evoked potentials	“EEG evoked potentials in response to visual-only stimuli (VEP), auditory-only stimuli (AEP) and audio-visual stimuli (AVEP) were compared in SPs and HNVS. Maps were generated every 15 ms from 50 to 200 ms post-stimulus. SP and HNV maps were then visually compared at each 15 ms time point. The electrodes	“Schizophrenia patients displayed increased physiological multisensory facilitation compared to the control group despite significant unisensory deficits. Our results indicate that basic multisensory processing, at least under some conditions, may represent one

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics					Experiment			Brain imaging analysis	Main results and major conclusions
	Participants (number; female/male; age interval; mean (SD))		Clinical and cognitive assessment (Test; mean total score)		Controlled/ matched for	Tasks and instructions	Conditions	Brain imaging technique		
	Schizophrenia/ Schizoaffective	Healthy controls	Subjects	Control group						
Stone et al. (2014)	46; 7/39; -; 39.2 (13.9)	57; 17/40; -; 39.4 (12.7)	DSM-IV, WASI; SCID NP, WASI; 119.9 (11.3)		Age, gender, no neurological disorders or prior head trauma, no substance abuse. For HC: history of psychiatric disorder in oneself or their first-degree relatives.	Audiovisual discrimination task Instructions: Participants were asked to indicate whether the stimulus was near to them or far from them by pressing one of two buttons on a response device with either their right index finger or right middle finger, respectively	1) Near visual stimulus. 2) Far visual stimulus. 3) Near auditory stimulus. 4) Far auditory stimulus. 5) Near audiovisual synchronous stimulus. 6) Far audiovisual synchronous stimulus. 7) Near audiovisual asynchronous stimulus. 8) Far audiovisual asynchronous stimuli	MEG; oscillatory activity	and time points detected by <i>t</i> -tests delineated spatio-temporal regions where significant group differences in amplitudes occurred. In the present study, the sum of VEP and AEP amplitudes (SUM) was subtracted from the AVEP amplitude at each electrode and time point [AVEP – (AEP + VEP)]. This measure was calculated for each participant separately and group averaged. Between-group comparisons of this measure were assessed with time point-by-time point independent samples <i>t</i> -tests. MANOVAs were used to test specific interactions and directions of any effects observed in the <i>t</i> -tests.”	type of spared functioning.”

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics				Experiment			Brain imaging analysis	Main results and major conclusions	
	Participants (number; female/male; age interval; mean (SD))		Clinical and cognitive assessment (Test; mean total score)		Controlled/ matched for	Tasks and instructions	Conditions			
	Schizophrenia/ Schizoaffective	Healthy controls	Subjects	Control group						Brain imaging technique
								baseline. Spectral power was compared between HC and SP across the 30–50 Hz (gamma-band) frequency range from 0 to 480 ms post-stimulus for each condition. Independent sample <i>t</i> -tests comparing SP to HC were applied at each sensor, time, and frequency point. The results of these <i>t</i> -tests were then used to identify candidate clusters of significant group differences in gamma-band power.”		
Straube et al. (2013)	16; 6/10; 19–53; 38 (-).	16; 0/16; 19–47; 27.9 (-)	DSM-IV; PANSS: 64(19); SAPS: 37(27); SANS: 22(27).	–	Handedness, normal hearing, normal vision, German was their primary language. For HC: no medical, neurological or psychiatric illness, past or present. For SP: No additional neurological and medical impairments as well as any cerebral abnormality, stable doses of atypical antipsychotic medication	Implicit audiovisual encoding task. Instructions: Participants instructed to watch the videos and respond each time they saw a new picture appear by pressing a left index finger (to ensure attention to task).	Combinations of speech and gestures: 1) Co-verbal gestures; concrete sentence content 2) Co-verbal gestures; abstract sentence content. 3) Sentences without gestures. 4) Gestures without sentences	fMRI	Analysis was performed by entering contrast images into a flexible factorial analysis as implemented in SPM8, in which subjects are treated as random variables. Baseline contrasts was the four conditions. The conjunction of bimodal in contrast to both unimodal conditions was calculated. The interaction between group and co-verbal gesture condition was calculated. A PPI analysis was conducted to assess differences in neural connectivity of the area around the left STS (including parts of the MTG and STG) for groups and conditions. On the group level, a random effects analysis (full factorial	“Our data indicate intact neural integration processes for iconic co-verbal gestures in patients with schizophrenia and specific aberrant activation for metaphoric co-verbal gestures. This indicates intact perceptual-matching processes within the pMTG/STS and dysfunctional higher order relational processes in the left IFG.” “Data indicate a reduced connectivity in patients, between brain regions normally involved in the processing of gestures in an abstract sentence context. It seems that processes that require the
Straube et al. (2014)	16; 6/10; 19–53; 38 (-). Subjects and paradigm is the same as in Straube et al. (2013) , just other analysis.	16; 0/16; 19–47; 27.9 (-)	DSM-IV; PANSS: 64(19)	–	Handedness, normal hearing, normal vision, German was their primary language. For HC: no medical, neurological or psychiatric illness, past or present. For SP: No additional neurological and medical impairments as	Implicit audiovisual encoding task. Instructions: Participants instructed to watch the videos and respond each time they saw a new picture appear by pressing a left index	Combinations of speech and gestures: 1) Co-verbal gestures; concrete sentence content 2) Co-verbal gestures; abstract sentence content. 3) Sentences without gestures. 4) Gestures without sentences	fMRI	“Data indicate a reduced connectivity in patients, between brain regions normally involved in the processing of gestures in an abstract sentence context. It seems that processes that require the	

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics				Experiment			Brain imaging analysis	Main results and major conclusions
	Participants (number; female/male; age interval; mean (SD))	Clinical and cognitive assessment (Test; mean total score)		Controlled/matched for	Tasks and instructions	Conditions	Brain imaging technique		
	Schizophrenia/ Healthy Schizoaffective controls	Subjects	Control group						
				well as any cerebral abnormality, stable doses of atypical antipsychotic medication.	finger (to ensure attention to task).			analysis) was conducted. Common connectivity patterns across groups and conditions were analyzed and interaction analyzes of group by condition.	connection of brain regions in close proximity are intact in patients, whereas processes relying on the interplay of more distant regions are disturbed most dramatically. Thus, the disconnectivity between STS and IFG might be the basis of dysfunctional integration of metaphoric verbal gestures and related interpersonal communication problems in patients with schizophrenia."
Szycik et al. (2009)	15; 8/7; -; 38.2 (10.6)	15; 8/7; -; 36.5 (9.4)	DSM-IV; BPRS - 24.1 (8.2).	Age, gender, handedness, native speakers of German. For HC: No previous or current psychiatric or neurological disorders. For SP: under antipsychotic treatment (chlorpromazine equivalent in mg/day was 380 ± 222), mean illness duration was 9.9 ± 9.4 years.	Audiovisual speech perception task Instructions: "To keep participants attending to the stimuli, they were required to identify words belonging to a specific semantic target category (i.e. animals) by pressing a button with the left/right index finger depending on whether a target was present/not present."	1) "Video segments with congruent audiovisual information (lip movements fitting to the spoken word; AV-congruent condition)". 2) "Video sequences with incongruent information in the audio and video channels (lip movements did not match the spoken word; AV-incongruent condition; e.g., video: Insel/ island, audio: hotel/hotel).	fMRI	Statistic model: design matrix including all conditions of interest specified using a hemodynamic response function. Multi-subject random effects (RFX). ANOVA, with within-subject factor AV-Con or AV-Incon and between-subject SP and HC.	"Robust differences were found between schizophrenia patients and matched control participants in the processing of audiovisual speech stimuli. The overall pattern of results suggests an abnormality of the patients in right medial and inferior frontal brain regions, possibly related to impaired processing of audiovisual speech by a mirror neuron system."
Szycik et al. (2013)	15; 8/7; -; 38.2 (10.6). Subjects and paradigm is the same as in Szycik et al. (2009), just other analysis.	15; 8/7; -; 36.5 (9.4)	DSM-IV; BPRS - 24.1 (8.2).	Age, gender, handedness, native speakers of German. For HC: No previous or current psychiatric or neurological disorders. For SP: under antipsychotic treatment (chlorpromazine equivalent in mg/	Audiovisual speech perception task Instructions: "To keep participants attending to the stimuli, they were required to identify words belonging to a specific semantic target category (i.e.	1) "Video segments with congruent audiovisual information (lip movements fitting to the spoken word; AV-congruent condition)". 2) "Video sequences with incongruent information in the audio and video channels	fMRI	Functional connectivity analysis with SPM8. General Linear Model. Three seeds were selected a priori: the left IFG, right IFG and right posterior superior temporal sulcus (RpSTS). One-sample t-test for each group in	"Differences in functional connectivity in particular of Broca's area during the processing of AV speech were found between patients with schizophrenia and HC, in particular for incongruent stimuli.

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics						Brain imaging analysis	Main results and major conclusions		
	Participants (number; female/male; age interval; mean (SD))		Clinical and cognitive assessment (Test; mean total score)		Controlled/ matched for	Experiment				
	Schizophrenia/ Schizoaffective	Healthy controls	Subjects	Control group	Tasks and instructions	Conditions			Brain imaging technique	
					day was 380 ± 222), mean illness duration was 9.9 ± 9.4 years.	animals) by pressing a button with the left/right index finger depending on whether a target was present/not present.”	(lip movements did not match the spoken word; AV-incongruent condition; e.g., video: insel/ island, audio: hotel/hotel).	each condition for the left IFG seed (Broca's area). ANOVA with group and stimulus.	Whereas control participants showed a reduced connectivity of Broca's area to other regions in this condition, SP showed increased connectivity which, we propose, is maladaptive in this case as AV integration is futile.”	
Wroblewski et al., 2020	17; 4/13; -; 33.12 (12.35); education years: 11.82 ± 1.77	17; 4/13; -; 32.65 (10.07); education years: 12.76 ± 1.39	ICD-10; SAPS 15 (6.89); SANS 9 (6.02); TMT A 31.49 (10.73); TMT B 68.56 (37.8); Digit Span forward 7.94 (1.75); Digit Span backward 6.59 (2.58); Verbal IQ 28.47 (3.91); Proverb Interpretation Task 1.38 (0.46); BAG; Word fluency test WTF	TMT A 26.45 (10.12); TMT B 54.18 (19.44); Digit Span forward 8.06 (2.51); Digit Span backward 6.59 (2.58); Verbal IQ 28.88 (3.91); Proverb Interpretation Task 1.13 (0.20); BAG; Word fluency test WTF	Age, gender, education, handedness, normal or corrected to normal vision, no neurological diseases, no substance use or alcohol abuse within the past six months, brain injury, native German speakers without any knowledge of Russian language. For SP: All except one received antipsychotic treatment; six were additionally treated with antidepressive or other psychiatric medication.	Content judgement task Instructions: Participants were instructed to decide (via button press with the index or middle finger of the left hand) whether the presented stimulus was rather object- or person-related. Participants were instructed to respond as soon as they decided on an answer.	1) An intrinsically meaningful gesture in context of a German sentence. 2) An intrinsically meaningful gesture in context of a Russian sentence 3) A German sentence with emblematic or tool-related content without being accompanied by any gesture	fMRI	Defined three experimental contrasts of interest. Verbal effect and conjunction null (between both unimodal conditions). First-level design matrices were specified, containing an “integration regressor”, a “verbal regressor” and a “gestural regressor”. Dynamic Causal Modelling to analyze effective connectivity, one-sample t-tests, Kolmogorov-Smirnov test. Comparison between groups we used one-way ANOVAs and t-tests (unpaired, two-samples, one-tailed, Bonferroni-corrected), correctional analyses, within-subject flexible factorial analysis.	“Across all groups, fMRI analyses revealed similar activation patterns [...] However, patients revealed significantly reduced connectivity in the verbal pathway (from left middle temporal gyrus (MTG) to left STS). Taken together our results demonstrate the relevance of the integration site during gesture-speech and show, that these processes are generally intact in SP. However, we provide evidence that patients might suffer from neural impairments within the verbal pathway during gesture-speech integration, which is related to concretism and SANS subscores.”
Wynn et al. (2014)	33; 10/23; -; 48.4 (9.0)	30; 9/21; -; 48.0 (8.1). Antipsychotic medication, 33	SCID, Axis I. SANS, BPRS; 36.7 (10.0)	SCID-I and SCID-II	Age, gender, IQ over 70, neurological illness or significant head	Audiovisual detection task Instructions: Subjects were instructed to	1) Visual stimulus: a black letter X. 2) Auditory stimulus: 1000	EEG; ERP	“To examine MSI effects on ERPs, we compared the AV condition to the sum of the	“In conclusion, our results show that schizophrenia patients exhibit

(continued on next page)

(continued)

Title, first author (year of publication)	Study characteristics				Brain imaging analysis	Main results and major conclusions		
	Participants (number; female/male; age interval; mean (SD))	Clinical and cognitive assessment (Test; mean total score)		Controlled/ matched for			Experiment	
		Schizophrenia/ Healthy Schizoaffective controls	Subjects					Control group
				injury with loss of consciousness for greater than 60 min. No substance dependence in the last 6 months, substance abuse in the last month. For HC: no history of psychotic disorder, bipolar disorder, recurrent depression, substance dependence, paranoid, schizotypal, schizoid, personality disorder or history of psychotic disorder among first-degree relatives.	push the mouse button as quickly as possible whenever they heard or saw any stimulus.	Hz tone. 3) Audiovisual stimulus: X and tone simultaneously	unisensory conditions (A + V) for the same set of electrodes, within each group. We analyzed each of the four ERP components separately using a 2 (group) × 2 (AV vs. A + V) repeated measures ANOVA. For all repeated measures ANOVAs resulting in more than one degree of freedom we used Greenhouse-Geisser correction factors (ε) to account for the effects of "sphericity".	relatively intact MSI, though with some subtle (non-significant) differences at the neural and behavioral levels. Our findings are inconsistent with most (but not all) of the previous studies of MSI in schizophrenia that have reported MSI deficits. Patients with schizophrenia clearly have problems in unisensory processing."

References

- Aine, C.J., Bockholt, H.J., Bustillo, J.R., Canive, J.M., Caprihan, A., Gasparovic, C., Calhoun, V.D., 2017. Multimodal neuroimaging in schizophrenia: description and dissemination. *Neuroinformatics* 15 (4), 343–364.
- Allen, P., Larø, I.F., McGuire, P.K., Aleman, A., 2008. The hallucinating brain: a review of structural and functional neuroimaging studies of hallucinations. *Neurosci. Biobehav. oral Reviews* 32 (1), 175–191. <https://doi.org/10.1016/j.neubiorev.2007.07.012>.
- Arce, E., Leland, D.S., Miller, D.A., Simmons, A.N., Winternheimer, K.C., Paulus, M.P., 2006. Individuals with schizophrenia present hypo- and hyperactivation during implicit cueing in an inhibitory task. *NeuroImage* 32 (2), 704–713. <https://doi.org/10.1016/j.neuroimage.2006.04.189>.
- Balz, J., Roa Romero, Y., Keil, J., Kriebler, M., Niedeggen, M., Gallinat, J., Senkowski, D., 2016. Beta/gamma oscillations and event-related potentials indicate aberrant multisensory processing in schizophrenia. *Front. Psychol.* 7, 1896. <https://doi.org/10.3389/fpsyg.2016.01896>.
- Barta, P.E., Pearlson, G.D., Powers, R.E., Richards, S.S., Tune, L.E., 1990. Auditory hallucinations and smaller superior temporal gyrus volume in schizophrenia. *Am. J. Psychiatr.* 147 (11), 1457–1462. <https://doi.org/10.1176/ajp.147.11.1457>.
- Baylis, G.C., Rolls, E.T., Leonard, C.M., 1987. Functional subdivisions of the temporal lobe neocortex. *J. Neurosci. Off. J. Soc. Neurosci.* 7 (2), 330–342. <https://doi.org/10.1523/JNEUROSCI.07-02-00330.1987>.
- Bodén, R., 2016. Schizofreni, andra psykoser och katatoni. In: Herlofson, J., Ekselius, L., Lundin, A., Mårtensson, B., Åsberg, M. (Eds.), *Psykatri, 2nd ed.* Studentlitteratur, pp. 269–294.
- Boutros, N.N., Arfken, C., Galderisi, S., Warrick, J., Pratt, G., Iacono, W., 2008. The status of spectral EEG abnormality as a diagnostic test for schizophrenia. *Schizophr. Res.* 99 (1–3), 225–237. <https://doi.org/10.1016/j.schres.2007.11.020>.
- Brang, D., Taich, Z.J., Hillyard, S.A., Grabowecky, M., Ramachandran, V.S., 2013. Parietal connectivity mediates multisensory facilitation. *NeuroImage* 78, 396–401. <https://doi.org/10.1016/j.neuroimage.2013.04.047>.
- Braus, D.F., Weber-Fahr, W., Tost, H., Ruf, M., Henn, F.A., 2002. Sensory information processing in neuroleptic-naïve first-episode schizophrenic patients: a functional magnetic resonance imaging study. *Arch. Gen. Psychiatry* 59 (8), 696–701. <https://doi.org/10.1001/archpsyc.59.8.696>.
- Butler, P.D., Javitt, D.C., 2005. Early-stage visual processing deficits in schizophrenia. *Curr. Opin. Psychiatry* 18 (2), 151–157. <https://doi.org/10.1097/00001504-200503000-00008>.
- Calvert, G.A., 2001. Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cereb. Cortex* 11 (12), 1110–1123. <https://doi.org/10.1093/cercor/11.12.1110>.
- Cappe, C., Rouiller, E.M., Barone, P., 2012a. Cortical and thalamic pathways for multisensory and sensorimotor interplay. In: Murray, M.M., Wallace, M.T. (Eds.), *The Neural Bases of Multisensory Processes.* CRC Press/Taylor & Francis.
- Cappe, C., Thelen, A., Romei, V., Thut, G., Murray, M.M., 2012b. Looming signals reveal synergistic principles of multisensory integration. *J. Neurosci.* 32 (4), 1171–1182. <https://doi.org/10.1523/JNEUROSCI.5517-11.2012>.
- Chen, X., Duan, M., Xie, Q., Lai, Y., Dong, L., Cao, W., Yao, D., Luo, C., 2015. Functional disconnection between the visual cortex and the sensorimotor cortex suggests a potential mechanism for self-disorder in schizophrenia. *Schizophr. Res.* 166 (1–3), 151–157. <https://doi.org/10.1016/j.schres.2015.06.014>.
- Crossley, N.A., Mechelli, A., Ginestet, C., Rubinov, M., Bullmore, E.T., McGuire, P., 2016. Altered hub functioning and compensatory activations in the connectome: a meta-analysis of functional neuroimaging studies in schizophrenia. *Schizophr. Bull.* 42 (2), 434–442. <https://doi.org/10.1093/schbul/sbv146>.
- Curcio-Blake, B., Liemburg, E., Vercammen, A., Swart, M., Knegtering, H., Bruggeman, R., Aleman, A., 2013. When Broca goes uninformed: reduced information flow to Broca's area in schizophrenia patients with auditory hallucinations. *Schizophr. Bull.* 39 (5), 1087–1095. <https://doi.org/10.1093/schbul/sbs107>.
- de Boer-Schellekens, L., Stekelenburg, J.J., Maes, J.P., Van Gool, A.R., Vroomen, J., 2014. Sound improves diminished visual temporal sensitivity in schizophrenia. *Acta Psychol.* 147, 136–142. <https://doi.org/10.1016/j.actpsy.2013.06.013>.
- de Dieuleveult, A.L., Siemonsma, P.C., van Erp, J.B.F., Brouwer, A.M., 2017. Effects of aging in multisensory integration: a systematic review. *Front. Aging Neurosci.* 9 (80), 80–94. <https://doi.org/10.3389/fnagi.2017.00080>.
- de Gelder, B., Vroomen, J., Annen, L., Masthof, E., Hodiament, P., 2003. Audio-visual integration in schizophrenia. *Schizophr. Res.* 59 (2–3), 211–218. [https://doi.org/10.1016/S0920-9964\(01\)00344-9](https://doi.org/10.1016/S0920-9964(01)00344-9).
- de Gelder, B., Vroomen, J., de Jong, S.J., Masthoff, E.D., Trompenaars, F.J., Hodiament, P., 2005. Multisensory integration of emotional faces and voices in schizophrenics. *Schizophr. Res.* 72 (2–3), 195–203. <https://doi.org/10.1016/j.schres.2004.02.013>.
- de Jong, J.J., Hodiament, P.P., Van den Stock, J., de Gelder, B., 2009. Audiovisual emotion recognition in schizophrenia: reduced integration of facial and vocal affect. *Schizophr. Res.* 107, 286–293. <https://doi.org/10.1016/j.schres.2008.10.001>.

- DeLisi, L.E., Sakuma, M., Tew, W., Kushner, M., Hoff, A.L., Grimson, R., 1997. Schizophrenia as a chronic active brain process: a study of progressive brain structural change subsequent to the onset of schizophrenia. *Psychiatry Res.* 74 (3), 129–140. [https://doi.org/10.1016/S0925-4927\(97\)00012-7](https://doi.org/10.1016/S0925-4927(97)00012-7).
- Desimone, R., Gross, C.G., 1979. Visual areas in the temporal cortex of the macaque. *Brain Res.* 178 (2–3), 363–380. [https://doi.org/10.1016/0006-8993\(79\)90699-1](https://doi.org/10.1016/0006-8993(79)90699-1).
- Ebdrup, B.H., Glenthøj, B., Rasmussen, H., Aggeraas, B., Langkilde, A.R., Paulson, O.B., Lublin, H., Skimminge, A., Baaré, W., 2010. Hippocampal and caudate volume reductions in antipsychotic-naïve first-episode schizophrenia. *J. Psychiatry Neurosci.* 35 (2), 95–104. <https://doi.org/10.1503/jpn.090049>.
- Engel, A.K., Senkowski, D., Schneider, T.R., 2012. Multisensory integration through neural coherence. In: Murray, M.M., Wallace, M.T. (Eds.), *The Neural Bases of Multisensory Processes*. CRC Press/Taylor & Francis.
- Ferri, F., Costantini, M., Salone, A., Di Iorio, G., Martinotti, G., Chiarelli, A., Merla, A., Di Giannantonio, M., Gallese, V., 2014. Upcoming tactile events and body ownership in schizophrenia. *Schizophr. Res.* 152 (1), 51–57. <https://doi.org/10.1016/j.schres.2013.06.026>.
- Foucher, J.R., Lacambre, M., Pham, B.T., Giersch, A., Elliott, M.A., 2007. Low time resolution in schizophrenia: lengthened windows of simultaneity for visual, auditory and bimodal stimuli. *Schizophr. Res.* 97 (1–3), 118–127. <https://doi.org/10.1016/j.schres.2007.08.013>.
- Friston, K., Brown, H.R., Siemerkus, J., Stephan, K.E., 2016. The dysconnection hypothesis. *Schizophr. Res.* 176 (2–3), 83–94. <https://doi.org/10.1016/j.schres.2016.07.014>.
- Frith, C., 2005. The neural basis of hallucinations and delusions. *C.R.Biol.* 328 (2), 169–175.
- Frumin, M., Golland, P., Kikinis, R., Hirayasu, Y., Salisbury, D., Hennen, J., Dickey, C., Anderson, M., Jolesz, F., Grimson, W., McCarley, R., Shenton, M., 2002. Shape differences in the corpus callosum in first-episode schizophrenia and first-episode psychotic affective disorder. *Am.J.Psychiatry* 159 (5), 866–868. <https://doi.org/10.1176/appi.ajp.159.5.866>.
- Ghazanfar, A.A., Schroeder, C., 2006. Is neocortex essentially multisensory? *Trends Cogn.Sci.* 10 (6), 278–285. <https://doi.org/10.1016/j.tics.2006.04.008>.
- Goghari, V.M., Sponheim, S.R., MacDonald, A.W., 2010. The functional neuroanatomy of symptom dimensions in schizophrenia: a qualitative and quantitative review of a persistent question. *Neurosci.Biobehav.Rev.* 34, 468–486. <https://doi.org/10.1016/j.neubiorev.2009.09.004>.
- Gur, R.E., Maany, V., Mozley, P.D., Swanson, C., Bilker, W., Gur, R.C., 1998. Subcortical MRI volumes in neuroleptic naïve and treated patients with schizophrenia. *Am.J. Psychiatr.* 155 (12), 1711–1717. <https://doi.org/10.1176/ajp.155.12.1711>.
- Gur, R.E., Loughhead, J., Kohler, C.G., Elliott, M.A., Lesko, K., Ruparel, K., Wolf, D.H., Bilker, W.B., Gur, R.C., 2007. Limbic activation associated with misidentification of fearful faces and flat affect in schizophrenia. *Arch.Gen. Psychiatry* 64 (12), 1356–1366. <https://doi.org/10.1001/archpsyc.64.12.1356>.
- Hampshire, A., Chamberlain, S.R., Monti, M.M., Duncan, J., Owen, A.M., 2010. The role of the right inferior frontal gyrus: inhibition and attentional control. *NeuroImage* 50 (3), 1313–1319. <https://doi.org/10.1016/j.neuroimage.2009.12.109>.
- Hanlon, F.M., Shaff, N.A., Dodd, A.B., Ling, J.M., Bustillo, J.R., Abbott, C.C., Stromberg, S.F., Abrams, S., Lin, D.S., Mayer, A.R., 2016. Hemodynamic response function abnormalities in schizophrenia during a multisensory detection task. *Hum. Brain Mapp.* 37 (2), 745–755. <https://doi.org/10.1002/hbm.23063>.
- Heinrichs, D.W., Buchanan, R.W., 1988. Significance and meaning of neurological signs in schizophrenia. *Am.J. Psychiatry* 145 (1), 11–18. <https://doi.org/10.1176/ajp.145.1.11>.
- Hershenson, M., 1962. Reaction time as a measure of intersensory facilitation. *J.Exp. Psychol.* 63 (3), 289–293. <https://doi.org/10.1037/h0039516>.
- Jardri, R., Pins, D., Bubrowsky, M., Lucas, B., Lethuc, V., Delmaire, C., Vantghem, V., Desprez, P., Thomas, P., 2009. Neural functional organization of hallucinations in schizophrenia: multisensory dissolution of pathological emergence in consciousness. *Cons. Cog.* 18 (2), 449–457.
- Jensen, A., Merz, S., Spence, C., Frings, C., 2019. Overt spatial attention modulates multisensory selection. *J.Exp.Psychol.Hum.Percept.Perform.* 45 (2), 174–188. <https://doi.org/10.1037/xhp0000595>.
- Kambeitz, J., Kambeitz-Ilanovic, L., Leucht, S., Wood, S., Davatzikos, C., Malchow, B., Falkai, P., Koutsouleris, N., 2015. Detecting neuroimaging biomarkers for schizophrenia: a meta-analysis of multivariate pattern recognition studies. *Neuropsychopharmacology* 40 (7), 1742–1751. <https://doi.org/10.1038/npp.2015.22>.
- Kaposvari, P., Cséte, G., Bogner, A., Csibri, P., Toth, E., Szabo, N., Vecsei, L., Sary, G., Kincses, Z.T., 2015. Audio-visual integration through the parallel visual pathways. *Brain Res.* 1624, 71–77. <https://doi.org/10.1016/j.brainres.2015.06.036>.
- Kayser, C., Logothetis, N.K., 2007. Do early sensory cortices integrate cross-modal information? *Brain Struct.Funct.* 212 (2), 121–132. <https://doi.org/10.1007/s00429-007-0154-0>.
- Keil, J., Senkowski, D., 2018. Neural oscillations orchestrate multisensory processing. *Neuroscientist* 24 (6), 609–626. <https://doi.org/10.1177/1073858418755352>.
- Kim, J.J., Crespo-Facorro, B., Andreasen, N.C., O'Leary, D.S., Magnotta, V., Nopoulos, P., 2003. Morphology of the lateral superior temporal gyrus in neuroleptic naïve patients with schizophrenia: relationship to symptoms. *Schizophr. Res.* 60 (2–3), 173–181. [https://doi.org/10.1016/S0920-9964\(02\)00299-2](https://doi.org/10.1016/S0920-9964(02)00299-2).
- Kinchla, R.A., 1974. Detecting target elements in multielement arrays: a confusability model. *Percept.Psychophys.* 15 (1), 149–158. <https://doi.org/10.3758/BF03205843>.
- Kolb, B., Whishaw, I.Q., 2015. *Fundamentals of Human Neuropsychology*, 7th ed. Worth Publishers.
- Kopell, N., Whittington, M., Kramer, M., 2011. Neuronal assembly dynamics in the beta1 frequency range permits short-term memory. *Proc.Natl. Acad.Sci.USA* 108 (9), 3779–3784. <https://doi.org/10.1073/pnas.1019676108>.
- Koychev, I., El-Deredey, W., Deakin, J.F., 2011. New visual information processing abnormality biomarker for the diagnosis of schizophrenia. *Exp.Opin.Med. Diagn.* 5 (4), 357–368. <https://doi.org/10.1517/17530059.2011.586029>.
- Krishnan, G.P., Hetrick, W.P., Brenner, C.A., Shekhar, A., Steffen, A.N., O'Donnell, B.F., 2009. Steady state and induced auditory gamma deficits in schizophrenia. *NeuroImage* 47 (4), 1711–1719. <https://doi.org/10.1016/j.neuroimage.2009.03.085>.
- Kunzelmann, K., Grieder, M., van Swam, C., Homan, P., Winkelbeiner, S., Hubl, D., Dierks, T., 2019. Am I hallucinating or is my fusiform cortex activated? Functional activation differences in schizophrenia patients with and without hallucinations. *Eur.J.Psychiatry* 33 (1), 1–7. <https://doi.org/10.1016/j.ejpsy.2018.06.002>.
- Kwon, J.S., O'Donnell, B.F., Wallenstein, G.V., Greene, R.W., Hirayasu, Y., Nestor, P.G., Hasselmo, M.E., Potts, G.F., Shenton, M.E., McCarley, R.W., 1999. Gamma frequency-range abnormalities to auditory stimulation in schizophrenia. *Arch.Gen. Psychiatry* 56 (11), 1001–1005. <https://doi.org/10.1001/archpsyc.56.11.1001>.
- Lieberman, J.A., Alvir, J.M., Korean, A., Geisler, S., Chakos, M., Sheitman, B., Woerner, M., 1996. Psychobiologic correlates of treatment response in schizophrenia. *Neuropsychopharmacology* 14 (3), 13–21. [https://doi.org/10.1016/0893-133X\(95\)00200-W](https://doi.org/10.1016/0893-133X(95)00200-W).
- Liu, T., Pinheiro, A.P., Zhao, Z., Nestor, P.G., McCarley, R.W., Niznikiewicz, M., 2016. Simultaneous face and voice processing in schizophrenia. *Behav. Brain Res.* 305, 76–86. <https://doi.org/10.1016/j.bbr.2016.01.039>.
- MacDonald, A.W., Carter, C.S., Kerns, J.G., Ursu, S., Barch, D.M., Holmes, A.J., Stenger, V.A., Cohen, J.D., 2005. Specificity of prefrontal dysfunction and context processing deficits to schizophrenia in never-medicated patients with first-episode psychosis. *Am.J. Psychiatry* 162 (3), 475–484. <https://doi.org/10.1176/appi.ajp.162.3.475>.
- Magné, M., Oranje, B., Engeland, H., Kahn, R., Kemner, C., 2009. Cross-sensory gating in schizophrenia and autism spectrum disorder: EEG evidence for impaired brain connectivity? *Neuropsychologia* 47 (7), 1728–1732. <https://doi.org/10.1016/j.neuropsychologia.2009.02.012>.
- Mäntylä, T., Nummenmaa, L., Rikandi, E., Lindgren, M., Kiesepää, T., Hari, R., Suvisaari, J., Raji, T.T., 2018. Aberrant cortical integration in first-episode psychosis during natural audiovisual processing. *Biol. Psychiatry* 84 (9), 655–664. <https://doi.org/10.1016/j.biopsych.2018.04.014>.
- McGurk, H., MacDonald, J., 1976. Hearing lips and seeing voices. *Nature* 264, 746–748. <https://doi.org/10.1038/264746a0>.
- Meredith, M.A., 2002. On the neuronal basis for multisensory convergence: a brief overview. *Cogn. Brain Res.* 14 (1), 31–40. [https://doi.org/10.1016/S0926-6410\(02\)00059-9](https://doi.org/10.1016/S0926-6410(02)00059-9).
- Meredith, M.A., Stein, B.E., 1983. Interactions among converging sensory inputs in the superior colliculus. *Science* 221 (4608), 389–391. <https://doi.org/10.1126/science.6867718>.
- Meredith, M.A., Stein, B.E., 1986. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. *J.Neurophysiol.* 56 (3), 640–662.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* 6 (7), 264–269. <https://doi.org/10.1371/journal.pmed.1000097>.
- Moher, D., Shamseer, L., Clarke, M., et al., 2015. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst.Rev.* 4 (1), 1–9. <https://doi.org/10.1186/2046-4053-4-1>.
- Mohr, P., Cheng, C., Claxton, K., Conley, R., Feldman, J., Hargreaves, W., Lehman, A., Lenert, L., Mahmoud, R., Marder, S., Neumann, P., 2004. The heterogeneity of schizophrenia in disease states. *Schizophr. Res.* 71 (1), 83–95. <https://doi.org/10.1016/j.schres.2003.11.008>.
- Müller, V.I., Kellerman, T.S., Seligman, S.C., Turetsky, B.I., Eickhoff, S.B., 2014. Modulation of affective face processing deficits in schizophrenia by congruent emotional sounds. *Soc.Cogn.Affect.Neurosci.* 9 (4), 436–444. <https://doi.org/10.1093/scan/nss107>.
- Norgren, E., Gröhn, C., 2020. Master's Thesis. Karlstad University, Karlstad, Sweden.
- Nuechterlein, K.H., Green, M.F., Calkins, M.E., Greenwood, T.A., Gur, R.E., Gur, R.C., Lazzeroni, L.C., Light, G.A., Radant, A.D., Seidman, L.J., Siever, L.J., Silverman, J.M., Sprock, J., Stone, W.S., Sugar, C.A., Swerdlow, N.R., Tsuang, D.W., Tsuang, M.T., Turetsky, B.I., Braff, D.L., 2015. Attention/vigilance in schizophrenia: performance results from a large multi-site study of the consortium on the genetics of schizophrenia (COGS). *Schizophr. Res.* 163 (1–3), 38–46. <https://doi.org/10.1016/j.schres.2015.01.017>.
- Ojanen, V., Möttönen, R., Pekola, J., Jääskeläinen, I.P., Joensuu, R., Autti, T., Sams, M., 2005. Processing of audiovisual speech in Broca's area. *NeuroImage* 25 (2), 333–338. <https://doi.org/10.1016/j.neuroimage.2004.12.001>.
- Onitsuka, T., Shenton, M.E., Kasai, K., Nestor, P.G., Toner, S.K., Kikinis, R., Jolesz, F.A., McCarley, R.W., 2003. Fusiform gyrus volume reduction and facial recognition in chronic schizophrenia. *Arch.Gen. Psychiatry* 60 (4), 349–355. <https://doi.org/10.1001/archpsyc.60.4.349>.
- Onitsuka, T., Oribe, N., Nakamura, I., Kanba, S., 2013. Review of neurophysiological findings in patients with schizophrenia. *Psychiatry Clin. Neurosci.* 67 (7), 461–470. <https://doi.org/10.1111/pcn.12090>.
- Pearl, D., Yodanis-Porat, D., Katz, N., Valevski, A., Aizenberg, D., Sigler, M., Weizman, A., Kikinzon, L., 2009. Differences in audiovisual integration, as measured by McGurk phenomenon, among adult and adolescent patients with schizophrenia and age-matched healthy control groups. *Compr. Psychiatry* 50 (2), 186–192. <https://doi.org/10.1016/j.comppsy.2008.06.004>.

- Pehrs, C., Deserno, L., Bakels, J.-H., Schlottermeier, L.H., Kappelhoff, H., Jacobs, A.M., Fritz, T., et al., 2014. How music alters a kiss: superior temporal gyrus controls fusiform-amygdala connectivity. *Soc.Cogn.Affect.Neurosci.* 9 (11), 1770–1778. <https://doi.org/10.1093/scan/nst169>.
- Pekkola, J., Ojanen, C.A., Jääskeläinen, I.P., Möttönen, R., Tarkiainen, A., Sams, M., Ville, A.T., 2005. Primary auditory cortex activation by visual speech: an fMRI study at 3 T. *NeuroReport* 16 (2), 125–128. <https://doi.org/10.1097/00001756-200502080-00010>.
- Pittman-Polletta, B.R., Kocsis, B., Vijayan, S., Whittington, M.A., Kopell, N.J., 2015. Brain rhythms connect impaired inhibition to altered cognition in schizophrenia. *Biol. Psychiatry* 77 (12), 1020–1030. <https://doi.org/10.1016/j.biopsych.2015.02.005>.
- Postmes, L., Sno, H., Goedhart, S., Van Der Stel, J., Heering, H., De Haan, L., 2014. Schizophrenia as a self-disorder due to perceptual incoherence. *Schizophr.Res.* 152 (1), 41–50. <https://doi.org/10.1016/j.schres.2013.07.027>.
- Roa Romero, Y., Keil, J., Balz, J., Gallinat, J., Senkowski, D., 2016a. Reduced frontal theta oscillations indicate altered crossmodal prediction error processing in schizophrenia. *J.Neurophysiol.* 116 (3), 1396–1407. <https://doi.org/10.1152/jn.00096.2016>.
- Roa Romero, Y., Keil, J., Balz, J., Niedeggen, M., Gallinat, J., Senkowski, D., 2016b. Alpha-band oscillations reflect altered multisensory processing of the McGurk illusion in schizophrenia. *Front.Hum.Neurosci.* 10, 41. <https://doi.org/10.3389/fnhum.2016.00041>.
- Ross, L.A., Saint-Amour, D., Leavitt, V.M., Molholm, S., Javitt, D.C., Foxe, J.J., 2007. Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophr.Res.* 97 (1–3), 173–183. <https://doi.org/10.1016/j.schres.2007.08.008>.
- Salvador, R., Canales-Rodríguez, E., Guerrero-Pedraza, A., Sarró, S., Tordesillas-Gutiérrez, D., Maristany, T., Crespo-Pacorro, B., McKenna, P., Pomarol-Clotet, E., 2019. Multimodal integration of brain images for MRI-based diagnosis in schizophrenia. *Front.Neurosci.* 13, 1203.
- Sanfratello, L., Aine, C., Stephen, J., 2018. Neuroimaging investigations of dorsal stream processing and effects of stimulus synchrony in schizophrenia. *Psychiatry Res. Neuroimaging* 278, 56–64. <https://doi.org/10.1016/j.pscychres.2018.05.005>.
- Sass, K., Heim, S., Sachs, O., Straube, B., Schneider, F., Habel, U., Kircher, T., 2013. Neural correlates of semantic associations in patients with schizophrenia. *Eur.Arch. PsychiatryClin.Neurosci.* 264 (2), 143–154. <https://doi.org/10.1007/s00406-013-0425-0>.
- Shenton, M.E., Dickey, C.C., Frumin, M., McCarley, R.W., 2001. A review of MRI findings in schizophrenia. *Schizophr.Res.* 49 (1–2), 1–52. [https://doi.org/10.1016/S0920-9964\(01\)00163-3](https://doi.org/10.1016/S0920-9964(01)00163-3).
- Silbersweig, D.A., Stern, E., Frith, C., Cahill, C., Holmes, A., Grooten, S., Seaward, J., McKenna, P., Chua, S.E., Schnorr, L., Jones, T., Frackowiak, R.S.J., 1995. A functional neuroanatomy of hallucinations in schizophrenia. *Nature* 378 (6553), 176–179. <https://doi.org/10.1038/378176a0>.
- Silverstein, S.M., Keane, B.P., 2011. Perceptual organization impairment in schizophrenia and associated brain mechanisms: review of research from 2005 to 2010. *Schizophr.Bull.* 37 (4), 690–699. <https://doi.org/10.1093/schbul/sbr052>.
- Stein, B.E., Meredith, M.A., 1993. *Cognitive Neuroscience. The Merging of the Senses.* The MIT Press.
- Stein, B.E., Stanford, T.R., 2008. Multisensory integration: current issues from the perspective of the single neuron. *Nat.Rev.Neurosci.* 9 (4), 255–266. <https://doi.org/10.1038/nrn2331>.
- Steinmann, S., Leicht, G., Mulert, C., 2019. The interhemispheric miscommunication theory of auditory verbal hallucinations in schizophrenia. *Int.J.Psychophysiol.* 145, 83–90. <https://doi.org/10.1016/j.ijpsycho.2019.02.002>.
- Stekelenburg, J.J., Maes, J.P., Van Gool, A.R., Sitskoorn, M., Vroomen, J., 2013. Deficient multisensory integration in schizophrenia: an event-related potential study. *Schizophr.Res.* 147 (2–3), 253–261. <https://doi.org/10.1016/j.schres.2013.04.038>.
- Stevenson, R.A., Park, S., Cochran, C., McIntosh, L.G., Noel, J.P., Barense, M.D., Ferber, S., Wallace, M.T., 2017. The associations between multisensory temporal processing and symptoms of schizophrenia. *Schizophr.Res.* 179, 97–103. <https://doi.org/10.1016/j.schres.2016.09.035>.
- Stone, D.B., Urrea, L.J., Aine, C.J., Bustillo, J.R., Clark, V.P., Stephen, J.M., 2011. Unisensory processing and multisensory integration in schizophrenia: a high-density electrical mapping study. *Neuropsychologia* 49 (12), 3178–3187. <https://doi.org/10.1016/j.neuropsychologia.2011.07.017>.
- Stone, D.B., Coffman, B.A., Bustillo, J.R., Aine, C.J., Stephen, J.M., 2014. Multisensory stimuli elicit altered oscillatory brain responses at gamma frequencies in patients with schizophrenia. *Front.Hum.Neurosci.* 8 <https://doi.org/10.3389/fnhum.2014.00788>.
- Straube, B., Green, A., Sass, K., Kirner-Veselinovic, A., Kircher, T., 2013. Neural integration of speech and gesture in schizophrenia: evidence for differential processing of metaphorical gestures. *Hum. Brain Mapp.* 34 (7), 1696–1712. <https://doi.org/10.1002/hbm.22015>.
- Straube, B., Green, A., Sass, K., Kircher, T., 2014. Superior temporal sulcus disconnectivity during processing of metaphorical gestures in schizophrenia. *Schizophr.Bull.* 40 (4), 936–944. <https://doi.org/10.1093/schbul/sbt110>.
- Surguladze, S.A., Calvert, G.A., Brammer, M.J., Campbell, R., Bullmore, E.T., Giampietro, V., David, A.S., 2001. Audio-visual speech perception in schizophrenia: an fMRI study. *Psychiatry Res. Neuroimaging* 106 (1), 1–14. [https://doi.org/10.1016/S0925-4927\(00\)00081-0](https://doi.org/10.1016/S0925-4927(00)00081-0).
- Surguladze, S., Russell, T., Kucharska-Pietura, K., Travis, M.J., Giampietro, V., David, A.S., Phillips, M.L., 2006. A reversal of the normal pattern of parahippocampal response to neutral and fearful faces is associated with reality distortion in schizophrenia. *Biol. Psychiatry* 60 (5), 423–431. <https://doi.org/10.1016/j.biopsych.2005.11.021>.
- Szyck, G.R., Münte, T.F., Dillo, W., Mohammadi, B., Samii, A., Emrich, H.M., Dietrich, D.E., 2009. Audiovisual integration of speech is disturbed in schizophrenia: an fMRI study. *Schizophr.Res.* 110 (1–3), 111–118. <https://doi.org/10.1016/j.schres.2009.03.003>.
- Szyck, G.R., Ye, Z., Mohammadi, B., Dillo, W., te Wildt, B.T., Samii, A., Frieling, H., Bleich, S., Münte, T.F., 2013. Maladaptive connectivity of Broca's area in schizophrenia during audiovisual speech perception: an fMRI study. *Neuroscience* 253, 274–282. <https://doi.org/10.1016/j.neuroscience.2013.08.041>.
- Talsma, D., Senkowski, D., Soto-Faraco, S., Woldorff, G.M., 2010. The multifaceted interplay between attention and multisensory integration. *Trends Cogn.Sci.* 14 (9), 400–410. <https://doi.org/10.1016/j.tics.2010.06.008>.
- Taylor, S.F., Welsh, R.C., Chen, A.C., Veldner, A.J., Liberzon, I., 2007. Medial frontal hyperactivity in reality distortion. *Biol. Psychiatry* 61 (10), 1171–1178. <https://doi.org/10.1016/j.biopsych.2006.11.029>.
- Tseng, H.H., Bossong, M.G., Modinos, G., Chen, K.M., McGuire, P., Allen, P., 2015. A systematic review of multisensory cognitive-affective integration in schizophrenia. *Neurosci.Biobehav.Rev.* 55, 444–452. <https://doi.org/10.1016/j.neubiorev.2015.04.019>.
- Uhlhaas, P.J., Singer, W., 2010. Abnormal neural oscillations and synchrony in schizophrenia. *Nat. Rev.* 11, 100–113. <https://doi.org/10.1038/nrn2774>.
- Uhlhaas, P.J., Singer, W., 2013. High-frequency oscillations and the neurobiology of schizophrenia. *Dialogues Clin.Neurosci.* 15, 301–313. <https://doi.org/10.3389/fpsyg.2013.00621>.
- van den Heuvel, M.P., Fornito, A., 2014. Brain networks in schizophrenia. *Neuropsychol. Rev.* 24 (1), 32–48. <https://doi.org/10.1007/s11065-014-9248-7>.
- Van den Stock, J., de Jong, S.J., Hodiament, P.P., de Gelder, B., 2011. Perceiving emotions from bodily expressions and multisensory integration of emotion cues in schizophrenia. *Soc.Neurosci.* 6 (5–6), 537–547. <https://doi.org/10.1080/17470919.2011.568790>.
- Vlcek, P., Bob, P., Raboch, J., 2014. Sensory disturbances, inhibitory deficits, and the P50 wave in schizophrenia. *Neuropsychiatr.Dis.Treat.* 10, 1309–1315. <https://doi.org/10.2147/NDT.S64219>.
- Wallace, M.T., Stevenson, R.A., 2014. The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia* 64, 105–123. <https://doi.org/10.1016/j.neuropsychologia.2014.08.005>.
- Wallace, M.T., Woynaroski, T.G., Stevenson, R.A., 2020. Multisensory integration as a window into orderly and disrupted cognition and communication. *Annu.Rev. Psychol.* 71, 193–219. <https://doi.org/10.1146/annurev-psych-010419-051112>.
- Wells, G., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., Tugwell, P., 2012. The Newcastle-Ottawa (NOS) for assessing the quality of non-randomised studies in meta-analysis. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
- White, T.P., Joseph, V., O'Regan, E., Head, K.E., Francis, S.T., Liddle, P.F., 2010. Alpha-gamma interactions are disturbed in schizophrenia: a fusion of electroencephalography and functional magnetic resonance imaging. *Clin. Neurophysiol.* 121 (9), 1427–1437. <https://doi.org/10.1016/j.clinph.2010.03.024>.
- Williams, L.M., Das, P., Liddell, B.J., Olivieri, G., Peduto, A.S., David, A.S., Gordon, E., Harris, A.W., 2007. Fronto-limbic and autonomic disjunctions to negative emotion distinguish schizophrenia subtypes. *Psychiatry Res.* 155 (1), 29–44. <https://doi.org/10.1016/j.pscychres.2006.12.018>.
- Williams, L.E., Light, G.A., Braff, D.L., Ramachandran, V.S., 2010. Reduced multisensory integration in patients with schizophrenia on a target detection task. *Neuropsychologia* 48 (10), 3128–3136. <https://doi.org/10.1016/j.neuropsychologia.2010.06.028>.
- Wroblewski, A., He, Y., Straube, 2020. Dynamic Causal Modelling suggests impaired effective connectivity in patients with schizophrenia spectrum disorders during gesture-speech integration. *Schizophr. Res.* 216, 175–183. <https://doi.org/10.1016/j.schres.2019.12.005>.
- Wynn, J.K., Jahshan, C., Green, M.F., 2014. Multisensory integration in schizophrenia: a behavioral and event-related potential study. *Cognit.Neuropsychiatry* 19 (4), 319–336. <https://doi.org/10.1080/13546805.2013.866892>.
- Zhang, L., Li, B., Wang, H., Li, L., Liao, Q., Liu, Y., Bao, X., Liu, W., Yin, H., Lu, H., Tan, Q., 2017. Decreased middle temporal gyrus connectivity in the language network in schizophrenia patients with auditory verbal hallucinations. *Neurosci. Lett.* 653, 177–182. <https://doi.org/10.1016/j.neulet.2017.05.042>.
- Zhou, H.-Y., Cai, X.-L., Weigl, M., Bang, P., Cheung, E.F.C., Chan, R.C.K., 2018. Multisensory temporal binding window in autism spectrum disorders and schizophrenia spectrum disorders: a systematic review and meta-analysis. *Neurosci. Biobehav.Rev.* 86, 66–76. <https://doi.org/10.1016/j.neubiorev.2017.12.013>.