# Emotional, Temporal, and Spatial Multisensory Integration and Crossmodal Recalibration in Individuals with Psychosis Proneness or Psychosis

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem Institut für Psychologie
in der Fakultät für Psychologie und Bewegungswissenschaften
an der Universität Hamburg
vorgelegt

von
Andreas Weiß, M.Sc.

Hamburg, 2024

**Tag der Disputation:** 15.05.24

**Promotionsprüfungsausschuss:**

Vorsitzender: Prof. Dr. Frank Rösler

1. Dissertationsgutachterin: Prof. Dr. Brigitte Röder

2. Dissertationsgutachterin: Prof. Dr. Tania Lincoln

1. Disputationsgutachterin: Prof. Dr. Anja Riesel

2. Disputationsgutachter: Prof. Dr. Philipp Sterzer

## Abstract

Since the first phenomenological descriptions of psychosis, severe alterations in sensory processing and perception have been assumed as one key factor at the root of the disorder. Studies on sensory processing in psychosis have repeatedly observed impairments in information processing within individual sensory systems such as vision or audition. It has been discussed that psychosis might be associated with pronounced deficits in multisensory processing, i.e. the processing of stimuli conveyed via multiple senses. Dysfunctions in multisensory integration (MSI) might lead to fragmented or distorted perception and facilitate the development of psychotic disorders via maladaptive perceptual learning mechanisms. Previous studies on MSI in psychosis have observed deficits such as reduced behavioral benefits during processing of congruent crossmodal stimuli or inappropriate integration. Impairments have repeatedly been observed in temporal, sensorimotor, and linguistic processes. However, findings in emotional MSI have been inconsistent, with reports of deficient, intact, and excessive integration. Only a few earlier studies suggested intact spatial MSI in psychosis and their findings require replication. No published study investigated if crossmodal recalibration, i.e. the short- and long-term adaptation to changes in crossmodal characteristics, is altered and as such might contribute to altered perception in psychosis. Further, little is known about the developmental trajectory of multisensory processes in psychosis and no convincing mechanism has been proposed, which describes how impaired multisensory processing might contribute to psychosis.

To address these gaps in the literature, this dissertation project aimed to investigate different aspects of multisensory processes in psychosis. We examined emotional MSI in psychosis proneness as well as MSI and crossmodal recalibration in the temporal and spatial domains in patients with the diagnosis of a psychotic disorder. First, we addressed the question, if impaired MSI might reflect a global underlying deficit or might be rather specific for some domains such as temporal and social information processing. Second, we examined crossmodal spatial and temporal recalibration in order to investigate if both processes might be altered in psychosis. Third, we aimed to contribute to the question if altered MSI might be a consequence of disorder manifestation or if it might play a role in the development of psychosis. Finally, we addressed the issue of a lack of disorder model-based research and aimed to provide initial evidence for a potential mechanism, via which altered multisensory processing might contribute to psychosis.

To investigate these questions, we conducted three different experimental studies. In the first study, we investigated emotional MSI in psychosis proneness by means of an emotion

categorization task with affective facial expressions and vocal prosody, expecting impaired emotional MSI in subjects with high psychosis proneness. In the second study, we investigated temporal MSI and crossmodal temporal recalibration in psychosis using an audiovisual (AV) simultaneity judgement (SJ) paradigm. We expected to replicate previous findings on impaired temporal MSI and provide first evidence for impaired crossmodal temporal recalibration in psychosis. In the third study, we conducted a paradigm on the spatial ventriloquist effect and aftereffect to investigate spatial MSI and crossmodal spatial recalibration in psychosis. We expected to replicate previous findings on intact spatial MSI in psychosis and provide first results on crossmodal spatial recalibration in psychosis.

Results showed that high and low proneness subjects did not differ in their emotional categorization performance in unimodal, bimodal emotionally congruent and bimodal emotionally incongruent stimuli. This indicates typical emotional MSI in psychosis proneness. Further, patients with the diagnosis of a psychotic disorder and healthy controls showed comparable performance in judging the synchrony of AV stimuli and similarly adjusted their responses to changes in stimulus asynchrony. This indicates typical temporal MSI and crossmodal temporal recalibration in psychosis. Finally, patients and controls showed similar ventriloquist effects and aftereffects, suggesting intact spatial MSI and crossmodal spatial recalibration in psychosis. In all studies, no correlation between measures of multisensory processes and psychotic symptoms could be observed.

While our findings on emotional MSI in psychosis proneness and temporal MSI in psychosis are in contrast with previous findings in patient samples, we successfully replicated earlier findings on intact spatial MSI. The lack of group differences in all three of our studies neither supports the hypothesis of a global underlying impairment nor of a domain-specificity of multisensory processing dysfunctions in psychosis. Further, our results indicate intact crossmodal recalibration in the spatial and temporal domain, suggesting that patients with the diagnosis of a psychotic disorder are able to adapt to changes in crossmodal spatial and temporal characteristics. Moreover, our findings do not offer support for the assumption that deficient emotional MSI plays a role in the development of psychosis. Taken together, our findings could suggest that altered perception in psychosis might not be generally driven by deficits in multisensory processes.

## Contents

## Abbreviations

| | |
|---|---|
| ANOVA | Analysis of Variance |
| ASD | Autism Spectrum Disorder |
| AV | Audiovisual |
| BCSS | Brief Core Schema Scales |
| BF | Bayes Factor |
| CAPE | Community Assessment of Psychic Experiences |
| CAPS | Cardiff Anomalous Perception Scale |
| CPT-IP | Continuous Performance Test Identical Pairs |
| DSM-V | Diagnostic and Statistical Manual of Mental Disorders, 5th Edition |
| EEG | Electroencephalography |
| EHI | Edinburgh Handedness Inventory |
| EOS | Early Onset Schizophrenia |
| ERP | Event-Related Potential |
| FDR | False Discovery Rate |
| fMRI | Functional Magnetic Resonance Imaging |
| IE | Inverse Efficiency |
| ITI | Inter-Trial-Interval |
| LSHS-E | Launey-Slade Hallucinations Scale Extended |
| MEG | Magnetencephalography |
| MSI | Multisensory Integration |
| PANSS | Positive and Negative Syndrome Scale |
| PCL | Paranoia Checklist |
| PSS | Point of Subjective Simultaneity |
| PSYRATS | Psychotic Symptom Rating Scales |
| RT | Reaction Time |
| SCID-V | Structural Clinical Interview for DMS-V Diagnoses |
| SJ | Simultaneity Judgement |
| SOA | Stimulus Onset Asynchrony |
| TBW | Temporal Binding Window |
| TMT | Trail Making Test |
| TOJ | Temporal Order Judgement |
| VAE | Ventriloquist Aftereffect |
| VE | Ventriloquist Effect |
| UA | Unimodal Auditory |

# Chapter I: General Introduction

It has been discussed that psychosis, a severe mental syndrome characterized by distressing symptoms such as hallucinations and delusions (American Psychiatric Association, 2022), might partly be rooted in marked alterations in sensory processing and perception (e.g. Mettler, 1955; Uhlhaas & Mishara, 2006). Starting in the second half of the 20th century up until today, research on perceptual impairments in psychosis has mainly focused on changes within one sensory modality, revealing impaired sensory processing of e.g. visual or auditory information (e.g. Javitt, 2009a, 2009b; Javitt & Freedman, 2015). However, in everyday life information usually is perceived via multiple sensory modalities and these so-called crossmodal stimuli interact with each other (Koelewijn et al., 2010). Research interest in sensory processing of crossmodal information in psychosis has increased over the last 20 years. It has been discussed that dysfunctional processing and integration of crossmodal stimuli might reflect a key impairment driving perceptual deficits in psychosis, facilitating distorted or confusing perception of the self and the world as well as potentially contributing to the development of psychotic symptoms (de Jong et al., 2009; Postmes et al., 2014; Tseng et al., 2015).

In this dissertation project, we investigated multisensory processing in psychosis. This projects aimed to pursue the questions, which aspects of multisensory processing might be impaired in psychosis and how these impairments might be associated with the development of psychotic symptoms. In the following, I will first briefly discuss previous literature on sensory processing in psychosis and then introduce multisensory processes and their basic principles and function. After the short introduction of these two research fields, I will transfer to the topic of this dissertation, discuss the literature on multisensory processes in psychosis and reveal open research questions, which were experimentally investigated within the scope of this dissertation project.

## 1 Psychosis and Sensory Processing

Psychosis is a severe mental syndrome affecting about 0.4% (12-month prevalence) of the population (Moreno-Küstner et al., 2018). Psychotic disorders such as schizophrenia have been rated among the worldwide top 15 disability causes (Vos et al., 2017). Patients with the diagnosis of a psychotic disorder suffer from changes in e.g. cognitive, affective, and social functioning severely affecting their everyday life and mental well-being (American Psychiatric Association, 2022). Psychotic symptoms such as hallucinations and delusions cause significant distress for patients and often occur in psychotic disorders such as

schizophrenia or schizoaffective disorder, but can also develop in affective disorders (American Psychiatric Association, 2022; Gaebel & Zielasek, 2015).

It has been observed that psychotic symptoms do not only occur in patients with the diagnosis of a psychotic disorder. Previous studies reported that they also occur in the general population in reduced intensity and/or frequency as well as in stages prior to disorder onset, i.e. in psychosis proneness (for a review, see Verdoux & van Os, 2002). Based on these observations, it has been discussed that psychosis as a mental syndrome might not reflect a dichotomous category, i.e. being either present or not, but might rather constitute a dimensional phenomenon. This so-called continuum hypothesis of psychosis proposed that psychotic symptoms occur on a symptomatic dimension from no psychosis to full-blown psychosis (Verdoux & van Os, 2002).

Already about 70 years ago, Mettler (1955) proposed that distorted perception might underlie the severe dysfunctions in cognition and emotional processes characterizing psychosis. Subjective reports of patients with the diagnosis of a psychotic disorder often described how they perceived the self and the environment to be altered and how this led to confusion and distress (e.g. Javitt & Freedman, 2015; McGhie & Chapman, 1961). Intuitively, one might think of psychosis as a disorder of perception since characteristic symptoms of psychosis such as hallucinations reflect erroneous perceptual experience and hence suggest dysfunctions in sensory processes (Wallace & Stevenson, 2014). Thus, psychosis might be partly described as a disorder of altered perception and cognition, in which impaired processing of sensory information drastically impacts the experience and interpretation of the surrounding world and the self in it.

Previous research on sensory processes in psychosis mainly focused on processing of visual or auditory information. Earlier studies reported numerous impairments in sensory processing and perception such as sensory gating dysfunctions, impaired sound detection and discrimination, deficient facial emotion recognition or reduced sensitivity to moving or ambiguous visual information (for reviews, see e.g. Bob et al., 2014; Javitt & Freedman, 2015). It has been discussed that deficient sensory processing might play a pivotal role in influencing cognitive functioning in psychosis in a bottom-up driven way. For example, a deficit in tone discrimination observed in psychosis might lead to impaired speech processing and emotional prosody recognition, thus exacerbating interpersonal difficulties (for a review, see Javitt, 2009a). This approach suggests a link between dysfunctional perception and higher-order cognitive impairments characterizing psychosis.

Beyond reports on impaired information processing in terms of dysfunctional perception in psychosis (Javitt & Freedman, 2015), it has been discussed that psychotic disorders might also been described as a disorder of cognitive disintegration (Bob et al., 2014). It has been reported that patients with the diagnosis of a psychotic disorder showed deficits in integrating information processed in distributed brain areas, suggesting a disconnectivity in neural functions essential for holistic information processing and potentially facilitating fragmentation of consciousness and cognition (for a review, see e.g. Bob et al., 2016).

While most of previous research on sensory processing in psychosis focused on vision or audition individually (Javitt & Freedman, 2015), the perceptual system does not merely rely on individual senses for sensory information processing. Rather, the brain needs to simultaneously process and integrate stimuli conveyed via multiple senses in order to perceive a coherent world and initiate appropriate behavior (e.g. de Jong et al., 2009). This phenomenon is called multisensory processing. Multisensory processes at least partly rely on unisensory processes, thus it is likely that deficient processing of unimodal information might give rise to pronounced deficits in multisensory processes (Foxe & Molholm, 2009). Since psychosis is associated both with impaired unisensory processing and perception as well as disintegration of information processing (Bob et al., 2016; Javitt & Freedman, 2015), it can be assumed that multisensory processes might be impaired in psychosis. This might potentially entail detrimental consequences for holistic perception and for subsequent cognitive processes as well as adaptive behavior.

# 2 Multisensory Processing & Perception

## 2.1 Multisensory Integration

MSI, i.e. the integration of stimuli conveyed via different sensory modalities such as vision and audition, describes the process, in which crossmodal stimuli are combined to form a new response differing from the individual cues (Stein et al., 2010). It has been discussed that MSI might reflect a non-linear gain mechanism, in which the resulting response goes beyond mere addition of the contributing sensory signals (e.g. Meredith & Stein, 1986; Schroeder & Foxe, 2005). MSI has been observed in a multitude of modality combinations such as AV, audio-tactile, visuo-tactile or sensorimotor information, including stimuli of various complexity such as simple flashes and sinusoidal sounds, stimuli with emotional content or speech stimuli (Murray et al., 2016; Wallace et al., 2020).

Several features of crossmodal stimuli influence the likelihood of integration. The integration of crossmodal stimuli is more likely when they are temporally and spatially aligned or congruent in supramodal features like semantic or affective meaning (Collignon et al., 2008; Doehrmann & Naumer, 2008; Meredith et al., 1987; Meredith & Stein, 1986; Stein et al., 1989). The integration of crossmodal stimuli has been found to hold crossmodal gain effects, i.e. neural and behavioural benefits, compared to processing of unimodal cues alone. This phenomenon is called congruency facilitation and can be measured by higher detection rates and faster response times when stimuli are congruent (Miller, 1982; Schroger & Widmann, 1998). Moreover, incongruency of crossmodal information is associated with e.g. lower detection rates, slower responses and higher error rates, called incongruency interference (Welch & Warren, 1980). Furthermore, it has been observed that crossmodal gain is greater for less effective, e.g. noisy or ambiguous, contributing unimodal signals. This principle is called inverse effectiveness (Meredith & Stein, 1986).

Thus, it is beneficial to appropriately integrate crossmodal information in order to optimize perceptual processes and subsequent cognition, e.g. through improving performance under noisy sensory conditions (Meredith & Stein, 1986; Miller, 1982; Schroger & Widmann, 1998). However, given the constant overflow of incoming information, the perceptual system is faced with the problem to decide when to integrate crossmodal stimuli and when not. This means that the brain has to constantly infer which crossmodal stimuli belong to the same and which to different events. It has been discussed that the brain solves this so-called causal inference problem by means of Bayesian inferencing. Within this process, prior knowledge about the association of crossmodal stimuli is combined with the sensory input stemming from the respective signals to infer if it is more likely that crossmodal stimuli stem from the same or different sources (Körding et al., 2007; Sato et al., 2007; Shams & Beierholm, 2010).

MSI can be experimentally investigated by means of various experimental paradigms. For example, the ventriloquist effect (VE) describes the crossmodal capture of a sound stimulus by a concurrently presented visual stimulus, which shifts the perceived location of the sound in direction of the visual stimulus position (Alais & Burr, 2004; Radeau & Bertelson, 1974). The McGurk effect, which is discussed to reflect linguistic MSI, describes the fusion of incongruent AV syllables into the perception of a third syllable, which differs from the visual and auditory cues (McGurk & MacDonald, 1976). In the temporal domain, the double-flash illusion describes the illusionary percept of two visual stimuli when one visual stimulus is presented temporally close to two sound stimuli (Shams et al., 2002). These

examples of multisensory effects are well suited to experimentally investigate MSI in the laboratory. Further, they illustrate how crossmodal stimuli influence each other during MSI, resulting in a response differing from the individual unimodal responses.

MSI has been discussed to be linked to the development of higher order cognitive skills, e.g. selective attention, association learning, language acquisition, rule learning and affect discrimination (Wallace et al., 2020). Thus, the intact development of multisensory processes might be crucial for the maturation of several cognitive processes and skills (Murray et al., 2016; Wallace et al., 2020). In sum, MSI constitutes a central process for the perceptual system to order the multitude of incoming sensory information and its functioning is essential for higher-order cognition, adaptive behavior, and for the creation of holistic percepts of the environment and the self (Murray et al., 2016; Wallace et al., 2020).

### 2.2 Crossmodal Recalibration

Beyond the on-line integration of crossmodal information, the perceptual system is able to adapt to changes in crossmodal characteristics. This so-called crossmodal recalibration has been observed both in the temporal and the spatial domain, reflecting adaptations to shifts in crossmodal features such as spatial disparity or temporal asynchrony (for a review, see Chen & Vroomen, 2013). For example, repeated exposure to an AV temporal asynchrony shifts the perceived simultaneity of subsequent AV stimuli in the direction of the previous temporal AV conflict (Fujisaki et al., 2004; Vroomen et al., 2004). In the spatial domain, exposure to an AV spatial disparity influences the perceived stimulus position of a subsequently presented sound. Even after crossmodal exposure, the shift of perceived sound position towards the previously presented visual stimulus persists when the subject is asked to localize the sound without any visual stimulus present, an effect called ventriloquist aftereffect (VAE; Radeau & Bertelson, 1974). These crossmodal temporal and spatial recalibration processes have been discussed to reflect crossmodal learning mechanisms. It has been discussed that the perceptual system adapts to a constantly changing world in order reduce discrepancy between the senses and thereby maintains coherent and stable representations of the environment and the self (Chen & Vroomen, 2013).

### 2.3 Potential Consequences of Impaired Multisensory Processes

In sum, MSI and crossmodal recalibration have been discussed to reflect processes essential for ordering an environment overflowing with sensory input and for the creation as well as maintenance of holistic and stable percepts of the world and the self (Chen & Vroomen, 2013; Murray et al., 2016). This seems to be important for the development of

higher order cognitive skills (Wallace et al., 2020). Thus, it can be argued that dysfunctions in multisensory processes might hold detrimental consequences for perception, cognition and behaviour. For example, it could be hypothesized that dysfunctional MSI leads to impaired integration of crossmodal stimuli and thus to reduced crossmodal gain or fragmentation of sensory information stemming from the same event. On the other hand, excessive MSI potentially leads to inaccurate binding of information, which did not originate from the same event. This might result in confusing percepts or the learning of erroneous associations (Odegaard & Shams, 2017). It has been discussed that such dysfunctions in MSI might be at the root of aberrant perceptual experiences in psychosis, leading to fragmented and unstructured perception of the environment and the self and potentially even facilitating the development of psychotic symptoms such as hallucinations or self-disorders (Postmes et al., 2014; Uhlhaas & Mishara, 2006).

# 3 Multisensory Processing in Psychosis

## 3.1 Previous Evidence on MSI in Psychosis

Comparing previous results on MSI in psychosis remains difficult due to a multitude of different experimental methods such as varying targeted sensory modalities, experimental paradigms or stimulus complexity and content (for reviews, see Tseng et al., 2015; Wallace et al., 2020). To date, it is not sufficiently clear, which exact multisensory processes might be impaired, if impairments might manifest as deficient or excessive integration or how impaired MSI might express across various sensory modalities. Thus, the generalizability of findings on MSI in psychosis so far is still limited, reducing the possibility to infer how globally MSI might be impaired in patients with the diagnosis of a psychotic disorder. In order to get an overview over the extent of dysfunctions in multisensory processes in psychosis, in the following prior evidence will be briefly reviewed separately for each investigated domain, i.e. temporal, spatial, linguistic, emotional or sensorimotor processing.

One domain of MSI, in which deficient MSI has been quite consistently observed in psychosis, is the temporal domain. Several studies reported dysfunctional temporal MSI in psychosis, with patients with the diagnosis of a psychotic disorder showing an enlarged tolerance for crossmodal asynchronies during SJ or an impaired ability to judge the temporal order of crossmodal stimuli (for a review and meta-analysis, see Zhou et al., 2018). It has been discussed that these impairments might facilitate erroneous binding and deficient temporal structuring of crossmodal information, resulting in confusing percepts (Amadeo et al., 2022; Zhou et al., 2018). Further, such dysfunctions in temporal MSI have also been

reported in psychosis proneness (e.g. Di Cosmo et al., 2021; Marsicano et al., 2022) and it has been hypothesized that deficient temporal MSI might facilitate the development of psychotic symptoms (Amadeo et al., 2022).

In contrast, the spatial domain has received significantly less scientific attention. Only few studies on spatial MSI in psychosis have been published and their results suggested that spatial MSI might be unimpaired in psychosis (Tseng et al., 2015). However, comparability of their findings might be limited due to variations in experimental designs as well as targeted sensory modalities. Moreover, one of the most extensively researched operationalizations of spatial MSI as implemented by a VE paradigm using simple AV stimuli (Bruns, 2019) has been investigated in psychosis in only one of those earlier studies (de Gelder et al., 2003).

Studying basic multisensory processes in psychosis, i.e. in simple stimuli beyond the temporal or spatial domain, might be helpful in evaluating the generalizability of dysfunctional MSI in psychosis. Findings on basic multisensory processes in psychosis so far have been rather inconsistent. Two previous AV target detection studies reporting conflicting findings, with Williams et al. (2010) observing reduced crossmodal gain in patients with the diagnosis of a psychotic disorder compared to controls, i.e. less benefit in AV compared to unimodal trials, whereas Wynn et al. (2014) reported comparable AV target detection performance in patients and controls. However, a recent visuo-tactile target detection study observed typical crossmodal target detection in patients with the diagnosis of a psychotic disorder (Moran et al., 2021), corroborating the findings of Wynn et al. (2014).

In domains with more complex stimuli such as stimuli with socially relevant content, the picture so far is not sufficiently clear. On the one hand, patients with the diagnosis of a psychotic disorder consistently showed deficient MSI in the linguistic domain compared to healthy controls in previous studies. For example, results of previous studies indicated a significantly reduced susceptibility to the McGurk effect or reduced crossmodal gain in AV lip-reading paradigms during crossmodal compared to unimodal trials in patients compared to controls (for a review, see Tseng et al., 2015). On the other hand, previous findings on MSI of emotional information so far have been inconsistent, with reports on deficient, excessive as well as intact integration of crossmodal affective stimuli (for reviews, see Lin et al., 2020; Tseng et al., 2015). Nevertheless, Lin et al. (2020) and Tseng et al. (2015) proposed that the inconsistency of findings on emotional MSI in psychosis might be largely attributable to methodological differences in previous studies. Further, unisensory emotion processing deficits are well documented in psychosis (D. Martin et al., 2020; Seo et al., 2020; Tripoli et

al., 2022; van Donkersgoed et al., 2015), so it could be argued that emotional multisensory processes might also be impaired in psychosis. Thus, Lin et al. (2020) and Tseng et al. (2015) assumed that emotional MSI – analogue to linguistic MSI – might show pronounced deficits compared to multisensory processes involving simple stimuli (Tseng et al., 2015). This might suggest selective impairments during processing of socially relevant crossmodal information in psychosis, potentially contributing to interpersonal difficulties in individuals with the diagnosis of a psychotic disorder (Lin et al., 2020; Tseng et al., 2015).

In the sensorimotor domain, i.e. stimuli including proprioception, multisensory processes have been quite consistently found to be impaired in psychosis. For example, patients with the diagnosis of a psychotic disorder showed an increased susceptibility to the rubber hand illusion (Noel, Cascio, et al., 2017). This sensorimotor illusion describes the effect that simultaneous tactile stimulation of a visible rubber hand and of the occluded own hand leads to a proprioceptive shift, resulting in a sense of ownership over the rubber hand (Botvinick & Cohen, 1998). Another example for dysfunctional sensorimotor processing in psychosis can be found in impaired corollary discharge. In psychosis, a failure in the appropriate integration of motor commands and the sensory consequences of one's own behaviour might lead to anomalous percepts of own actions. This might cause confusion about the origin of own actions and facilitate psychotic symptoms linked to own behaviour (Bansal et al., 2018; Thakkar et al., 2017). It has been discussed that previous findings on sensorimotor processes in psychosis might indicate a weaker differentiation of the self and external cues (Noel, Cascio, et al., 2017) and an impaired sense of agency over the own body (Klaver & Dijkerman, 2016). Further, previous findings suggested that these dysfunctions might be linked to psychotic symptoms such as delusions of control and hallucinations (Bansal et al., 2018; Noel, Cascio, et al., 2017; Tschacher et al., 2017).

### 3.2 Crossmodal Recalibration in Psychosis

While there has been considerable work on MSI in psychosis, too little scientific attention has been paid to crossmodal recalibration processes in psychosis. Published studies on crossmodal spatial or temporal recalibration involving external, i.e. sensory-sensory stimuli in psychosis are essentially non-existent. Thus, the extent to which crossmodal learning processes might be associated with psychosis or if crossmodal recalibration might even play a role in disorder development is poorly understood. It could be argued that impaired crossmodal recalibration might lead to unstable and unreliable representations of the environment and as such contribute to confusing and distorted perception previously reported

in psychosis (Javitt & Freedman, 2015; Uhlhaas & Mishara, 2006). In the sensorimotor domain, previous evidence indicated that sensorimotor recalibration processes such as visuomotor recalibration in tasks such as prism adaptation or saccadic adaptation might be impaired in psychosis, suggesting deficits in perceptual motor learning processes (e.g. Bartolomeo et al., 2020; Rösler et al., 2015). To date, the question however remains if crossmodal learning, i.e. sensory-sensory recalibration might also be impaired in psychosis and how this potentially contributes to altered perceptual experience in psychosis.

### 3.3 Interim Summary of Evidence on Multisensory Processes in Psychosis

In sum, findings on MSI in psychosis are rather heterogeneous and likely depend on several factors such as stimulus complexity and targeted domain. Deficient multisensory processes have been revealed in previous studies especially in the temporal, sensorimotor, linguistic domain (Noel, Cascio, et al., 2017; Tseng et al., 2015; Zhou et al., 2018). Further, alterations have also been observed in the emotional domain, albeit less consistently (Lin et al., 2020; Tseng et al., 2015). A few previous studies indicated that MSI might be intact e.g. in the spatial domain (Tseng et al., 2015), suggesting that multisensory processing might be selectively impaired in more complex information such as social cues (Tseng et al., 2015) and stimuli involving temporal and proprioceptive/sensorimotor signals (Noel, Cascio, et al., 2017; Zhou et al., 2018). Due to inconsistencies in findings and/or limited number of published studies, so far it however remains largely unanswered if deficits in MSI might reflect a global deficit or are rather domain-specific. Further, to date it is unclear if deficits in multisensory processing can primarily be observed in on-line integration processes or if they extend to crossmodal recalibration. Previous findings on sensorimotor recalibration in patients with the diagnosis of a psychotic disorder suggested deficits in perceptual motor learning (e.g. Bartolomeo et al., 2020; Rösler et al., 2015). Studies are however needed, which investigate if sensory-sensory recalibration deficits can be observed in psychosis.

Although it has been discussed that deficits in multisensory processes might be linked to the *development* of psychosis (Tseng et al., 2015; Wallace & Stevenson, 2014), most of the previous studies have administered cross-sectional designs in samples already diagnosed with a psychotic disorder. However, to evaluate if dysfunctions in multisensory processing might precede and potentially even play a role in disorder development or rather manifest as consequence of disorder progress, it is crucial to investigate subjects with psychosis proneness. So far, few studies on MSI in psychosis proneness have been published, and they do not sufficiently cover the whole range of the above-mentioned domains of MSI.

Nonetheless, previous work suggested deficits in temporal MSI (Dalal et al., 2021; Di Cosmo et al., 2021; Ferri et al., 2017, 2018; Marsicano et al., 2022) and sensorimotor processing (Germine et al., 2013; Kállai et al., 2015; Torregrossa & Park, 2022) in psychosis proneness. Findings on linguistic MSI in psychosis proneness have been inconclusive (Muller et al., 2020, 2021). To our knowledge, no published study investigated spatial or emotional MSI in psychosis proneness. This might especially surprise in the case of emotional MSI. Emotional MSI can be assumed to be at least altered in psychosis and it has been discussed that deficient emotional MSI might facilitate experiences of interpersonal difficulties (Lin et al., 2020; Tseng et al., 2015). Further, impairments in unisensory emotional processing have repeatedly been observed in psychosis proneness and a role in disorder development has been discussed (D. Martin et al., 2020; Seo et al., 2020; Tripoli et al., 2022; van Donkersgoed et al., 2015). Thus, studies on emotional MSI in psychosis proneness are necessary to gather support for above-mentioned assumptions of its role in disorder development.

Taken together, while psychosis might at least partly be described as a disorder of sensory processing and disintegration of information processing (Bob et al., 2014, 2016; Javitt & Freedman, 2015), the picture on multisensory processing in psychosis is not yet clear enough to pinpoint which exact processes are disturbed as opposed to those that remain intact. Further, too little scientific attention has been paid to factors, which might influence multisensory processing in psychosis. Moreover, it remains unclear if dysfunctions in multisensory processes in psychosis might reflect a consequence of disorder manifestation or if (potential) deficits in MSI and crossmodal recalibration might play a role in the development of psychosis.

# 4 Linking Dysfunctional Multisensory Processing and the Development of Psychosis

## 4.1 A Lack of Theoretical Models on Multisensory Processes in Psychosis

Although previous studies assumed that dysfunctions in multisensory processes facilitate the development of psychosis (e.g. Tseng et al., 2015), little attention has been paid to revealing potential mechanisms, via which deficits in multisensory processes might contribute to psychosis. It appears that previous research so far has been mainly driven by research questions directly derived from experimental findings of preceding studies and less based on models, which connect experimental findings and phenomenology of psychosis.

In their review, Postmes et al. (2014) proposed how self-disorders in psychosis, i.e. deviations in the experience of the self, might be rooted in a deficit in the adequate processing of proprioceptive and external cues. Dysfunctions in (multi)sensory processing, especially

sensorimotor and proprioceptive processes, give rise to uncertain and irritating perception. In turn, the attentional focus shifts to internal processes such as thoughts and emotions, both losing focus on sensory contextual and other meaningful information and increasing the focus on potentially irrelevant details of internal processes. The authors hypothesized that this process of impaired sensory input, attentional shift and subsequent deficient integration of internal and external cues into a holistic percept occurs in a persistent cycle, leading to an estranged experience of the self in the surrounding world. Further, they proposed that psychotic symptoms can develop as a maladaptive mechanism to resolve or explain these experiences by means of mental functions such as imagination or memory. Thus, sensory conflicts and resulting confusing perception are resolved at the cost of reality monitoring (Postmes et al., 2014).

Although the proposed account of Postmes et al. (2014) generated interesting hypotheses on the potential role of (multi-)sensory processing in psychosis, their proposed mechanism heavily focused on proprioceptive processes. Only very few published studies on dysfunctions in sensory-sensory, i.e. non-proprioceptive processes in psychosis have been considered in the account of Postmes et al. (2014). In this account, impairments in sensory-sensory processing have been largely neglected, subordinated into general sensory impairments and as such were not sufficiently addressed as a potential key deficit driving perceptual dysfunctions in psychosis, as it has been previously proposed (e.g. Tseng et al., 2015). Moreover, Postmes et al. (2014) claimed that any sensory process, including the processing of crossmodal information, is inevitably linked with interoceptive processes, with the self serving as a reference for all sensory experiences (Postmes et al., 2014). It could be argued that this claim might not strictly describe a sensory process per se but rather involve cognitive, i.e. higher-order processes. This raises the question if the proposed mechanism by (Postmes et al., 2014) coherently refers to sensory disintegration in terms of MSI or rather disintegration between sensory processing and other mental processes. Further, while (Postmes et al., 2014) discussed previous literature and a potential mechanism by which altered (multi-)sensory processes might be linked to psychosis, they distinctly focus on self disorders in psychosis. Hallucinations and delusions are only briefly mentioned in the mechanism, if at all, limiting the explanatory value to an individual symptom of psychosis rather than comprehensively addressing the syndrome of psychosis. Finally, despite the proposition by Postmes et al. (2014) to explain the role of dysfunctions in multisensory processing for psychosis, to our knowledge they did not elaborate a disorder model with subsequent empirical testing of derived hypotheses.

To conclude, model accounts focusing on multisensory processing in psychosis and its potential link to psychosis as a syndrome rather than isolated symptoms are lacking. Given at least partly consistent reports of dysfunctional multisensory processes beyond sensorimotor processes, i.e. in sensory-sensory processing (Lin et al., 2020; Tseng et al., 2015; Wallace et al., 2020; Zhou et al., 2018), it is necessary to bring impaired sensory-sensory processing into the focus of discussions on potential mechanisms of dysfunctional perception in the development of psychosis.

### 4.2 Anomalous Perceptual Experiences as a Potential Link between Multisensory Processing and Psychosis

To reconcile experimental findings on MSI in psychosis and phenomenology, it might be fruitful to link experimental findings with previously published and empirically tested disorder models of psychosis, which contain model components on perceptive processes. A candidate for such a model is the cognitive disorder model by Garety et al. (2001; see paragraphs below). Embedding evidence on associations between dysfunctional multisensory processing and psychotic symptoms in a published disorder model will benefit the explanatory value of the model and contribute to the understanding how altered (multi-) sensory perception might impact the development of psychosis.

An extensive review of the published literature on multisensory processes in psychosis revealed that previous studies directly examined correlations between experimental findings and measures for psychotic symptoms such as the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987) as opposed to investigating a potentially mediating construct. The findings of these studies were largely inconsistent, with both reports on correlations (Castagna et al., 2013; Ferri et al., 2014, 2017; Germine et al., 2013; He et al., 2021; Muller et al., 2021; Müller et al., 2014; Portnova et al., 2021; Prikken et al., 2019; Sanfratello et al., 2018; Stekelenburg et al., 2013; Stephen et al., 2013; Stevenson, Park, et al., 2017; Thakkar et al., 2011; Torregrossa & Park, 2022; White et al., 2014; Williams et al., 2010; Wynn et al., 2014; Zhou, Cui, et al., 2022) and no correlations (de Boer-Schellekens et al., 2014; Giannitelli et al., 2015; Hanlon et al., 2016; Haß et al., 2017; Mangelinckx et al., 2017; B. Martin et al., 2013; Pearl et al., 2009; Roa Romero, Keil, Balz, Gallinat, et al., 2016; Roa Romero, Keil, Balz, Niedeggen, et al., 2016; Senkowski & Moran, 2022; Seubert, 2010; Stone et al., 2011, 2014; Surguladze et al., 2001; Vanes et al., 2016; Zhou et al., 2020, 2021; Zhou, Lai, et al., 2022; Zvyagintsev et al., 2017) between MSI and psychotic symptoms. Further, about one third of previous studies did not report any correlation analysis between their experimental

findings and symptom measures (de Gelder et al., 2003, 2005; de Jong et al., 2009, 2010; Fiszdon & Bell, 2009; Liu et al., 2016; Moran et al., 2021; Noel et al., 2018; Ross et al., 2007; Sestito et al., 2013; Simpson et al., 2013; Szycik et al., 2009, 2013; Thaler et al., 2013; Van den Stock et al., 2011; Wroblewski et al., 2020; Zvyagintsev et al., 2013). It might be difficult to directly relate experimental operationalizations of multisensory processes and phenomenology of psychosis, with both reflecting distinct levels of psychological observation. Thus, we aimed to explore a conceptual construct able to relate to both components and thus to obtain first data for hypothesis generation on mechanisms linking MSI and psychotic symptoms.

A phenomenological concept, which potentially links experimental findings and psychotic symptoms, might lie in so-called anomalous perceptual experiences. Anomalous perceptual experiences describe unusual sensory phenomena in all senses, such as perception of unusually increased stimulus intensity or sensory experiences originating from an inexplicable source (Bell et al., 2006; Wright et al., 2018). The concept of anomalous perceptual experiences as operationalized by Bell et al. (2006) does not exclusively rely on a clinical perspective. This is in contrast with earlier operationalizations of aberrant perception, which often are biased towards clinical symptoms such as hallucinations rather than reflecting a distinct phenomenon (Bell et al., 2006). Adopting the approach by Bell et al. (2006) provides the opportunity to investigate perceptual anomalies both in clinical populations and individuals from the general population with various degrees of psychotic experiences such as persons with high psychosis proneness. Thus, this conceptualization of anomalous perceptual experiences supports the continuum hypothesis of psychosis (Verdoux & van Os, 2002) and enables the examination of altered perception separated of psychotic symptoms and as such of altered perception in stages before psychotic symptoms manifest (Bell et al., 2006).

Aberrant perception in terms of anomalous perceptual experiences has been described in the above-mentioned cognitive disorder model of psychosis by Garety et al. (2001). This model incorporates aberrant perception as one key factor facilitating the development of psychotic symptoms (see Figure 1 for a depiction of the model). Garety et al. (2001) proposed that a stressful event can trigger a disruption of cognitive processes in individuals with a predisposition for psychosis, leading to anomalous perceptual experiences such as hearing one's own thoughts aloud or distorted sensory input. These experiences are attention-capturing, irritating as well as arousing and demand the search for a cognitive explanation. Confounded by cognitive biases such as the jumping-to-conclusion bias (Huq et al., 1988) and

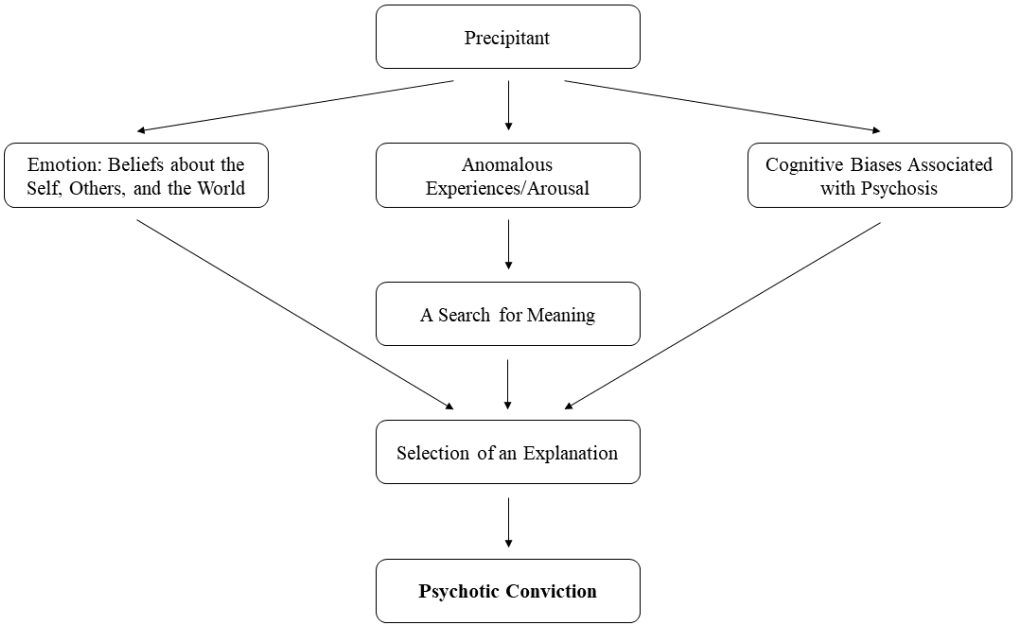
negative beliefs about the self and the world, the interpretation of the anomalous experience is biased towards negative, potentially threatening explanations and as such eventually facilitates the construction of psychotic convictions (Garety et al., 2001).

While cognitive disorder model by Garety et al. (2001) introduced a mechanism how aberrant perception might be linked to psychotic symptoms, the description what the component *anomalous perceptual experience* entails might be rather vague and did not address anomalous multisensory perception. In this dissertation project, we propose that dysfunctions in multisensory processing, which lead to e.g. fragmentation of information originating from the same event or erroneous binding of stimuli from separate sources, might be associated with the emergence of conscious anomalous perceptual experiences. In this perspective, dysfunctional multisensory perception might facilitate the manifestation of psychotic symptoms in accordance to the mechanism described by Garety et al. (2001). Following this approach might aid in addressing the above discussed question if impairments in multisensory processes might play a role in the development of psychosis rather than reflecting a consequence of disorder manifestation.

Thus, in order to generate hypotheses regarding anomalous perceptual experiences as a potential conceptual link between experimental findings on deficient multisensory processing in psychosis and psychotic symptoms, an additional aim of this dissertation project was to investigate associations between measures of multisensory processes, anomalous perceptual experiences as operationalized in the questionnaire Cardiff Anomalous Perception Scale (CAPS) by Bell et al. (2006) and psychotic symptoms such as hallucinations and delusions.

***Figure 1***

*Depiction of the Cognitive Disorder Model for Psychosis by Garety et al. (2001)*




*Note*. The depicted path model was adapted from Freeman et al. (2002).

## 5 Summary of Open Questions

Within the scope of this dissertation project, four overarching open questions could be derived from the literature on multisensory processing in psychosis and were subsequently addresses in three experimental studies.

The first question followed previous discussions if dysfunctions in MSI in psychosis could be observed on a global scope, suggesting a general underlying dysfunction, or might be rather specific for certain domains of MSI (Tseng et al., 2015). As outlined above, previous evidence rather tends to support the latter, but a lack of studies in spatial MSI and inconsistencies of findings in emotional MSI in psychosis so far prevented clear conclusions. Thus, we followed up on the first question "*1: Are Multisensory Integration Deficits in Psychosis Global or Domain-Specific?"* by investigating both emotional and spatial MSI in subjects with psychotic symptoms. We aimed to contribute to previous evidence on emotional MSI in psychosis and to address the issue if emotional MSI might be deficient or excessive. Further, we aimed to gather evidence on spatial MSI in psychosis, targeting the relative scarcity of data in this field.

The second question addressed the lack of evidence on crossmodal learning in psychosis. As discussed above, studies investigated sensory-sensory recalibration in psychosis

are lacking, whereas previous studies indicated deficits in sensorimotor recalibration processes in psychosis (e.g. Bartolomeo et al., 2020; Rösler et al., 2015). We aimed to provide first evidence on sensory-sensory recalibration in psychosis and by this address the question "*2: Is Crossmodal Learning Deficient in Psychosis?*". We experimentally examined crossmodal temporal recalibration in patients with the diagnosis of a psychotic disorder and by this investigated if impairments temporal MSI previously reported in psychosis (Zhou et al., 2018) also extend to crossmodal temporal recalibration. Further, we investigated crossmodal spatial recalibration in patients with the diagnosis of a psychotic disorder, addressing the scarcity of findings on spatial multisensory processes in psychosis and providing first evidence on crossmodal spatial recalibration.

The third question addressed the relative lack of studies on MSI in psychosis proneness. This question followed up upon the above outlined discussion if impaired MSI might contribute to the development of psychosis rather than reflecting a consequence of it (Tseng et al., 2015). Most of previous research has been conducted in patients with the diagnosis of a psychotic disorder and published studies in psychosis proneness did not sufficiently target all domains of MSI. While studies have been conducted in the temporal (Dalal et al., 2021; Di Cosmo et al., 2021; Ferri et al., 2017, 2018; Marsicano et al., 2022), sensorimotor (Germine et al., 2013; Kállai et al., 2015; Torregrossa & Park, 2022), and linguistic domains (Muller et al., 2020, 2021), none have been published in the spatial and emotional domain. Given previous discussions of a potential role of dysfunctional emotional MSI for the development of psychosis (Lin et al., 2020; Tseng et al., 2015) and consistent reports of impaired unisensory emotion processing in psychosis proneness (D. Martin et al., 2020; Seo et al., 2020; Tripoli et al., 2022; van Donkersgoed et al., 2015), we aimed to contribute to the question "3: *Are Deficits in Multisensory Processing in Psychosis a Precedent or a Consequence of Disorder Development?*" by investigating emotional MSI in subjects with psychosis proneness.

The fourth and last question addressed in this dissertation project aimed to target the lack of attempts to reveal mechanisms, which link dysfunctional multisensory processing and the phenomenology of psychosis. As outlined above, we proposed the concept of anomalous perceptual experiences described in the cognitive disorder model of psychosis by Garety et al. (2001) as a potential mechanism linking experimental measures of multisensory processing in psychosis and psychotic symptoms. This approach further contributed to the question if dysfunctional multisensory processing might play a role in the development of psychosis by

providing initalevidence for a specific mechanism, via which multisensory processing might contribute to psychosis. Thus in order to address the question "4: *Can Anomalous Perceptual Experiences Conceptually Link Multisensory Processing and Phenomenology of Psychosis?*", we aimed to reveal potential associations between experimental findings on emotional, temporal, and spatial multisensory processes in psychosis, anomalous perceptual experiences as operationalized by the CAPS (Bell et al., 2006) and self-reports on psychotic symptoms. By this, we aimed to gather evidence for future hypothesis generation regarding mechanisms linking multisensory processing and psychotic symptoms.

## 6 Outlook on the Experimental Studies in this Dissertation Project

In order to investigate the questions outlined above, we conducted three different studies covering various multisensory processes in individuals with different levels of psychotic symptoms.

In our first study, which is addressed in **Chapter II "Emotional Multisensory Integration in Psychosis Proneness"**, emotional MSI was investigated in psychosis proneness administering a paradigm on AV emotion processing by Föcker et al. (2011). The aim of this study was to investigate if findings on dysfunctional emotional MSI in individuals with psychotic symptoms can be replicated and if emotional MSI might be already altered in psychosis proneness. This study addressed questions 1 and 3 of the above outlined open questions.

**Chapter III "Temporal Multisensory Integration and Crossmodal Temporal Recalibration in Psychosis"** focuses on our second study on temporal multisensory processing in psychosis. Temporal MSI and crossmodal temporal recalibration were investigated in patients with the diagnosis of a psychotic disorder and healthy controls by means of a paradigm adapted from Van der Burg et al. (2015). This study's objective was to both replicate earlier findings on impaired temporal MSI in psychosis and to provide first evidence if dysfunctions in crossmodal temporal recalibration can be observed in adults with the diagnosis of a psychotic disorder. This study aimed to contribute to questions 1 and 2 of the outlined open questions.

**Chapter IV "Spatial Multisensory Integration and Crossmodal Spatial Recalibration in Psychosis"** covers our third study on spatial multisensory processing in psychosis. Both spatial MSI and crossmodal spatial recalibration were investigated in patients with the diagnosis of a psychotic disorder and healthy controls with a paradigm on the VE and

VAE adopted from Bruns & Röder (2015). This study aimed to replicate earlier findings on spatial MSI in psychosis and provide first experimental evidence on crossmodal spatial recalibration in psychosis. The third study addressed questions 1 and 2 of the open questions.

To evaluate if dysfunctional MSI might be linked to the phenomenology of psychosis via the mechanism of anomalous perceptual experiences, we investigated associations between multisensory processing, anomalous perceptual experiences as operationalized by the CAPS (Bell et al., 2006) and psychotic symptoms in all three studies. Based on the cognitive disorder model of psychosis by Garety et al. (2001), we aimed to provide first evidence for future hypothesis generation on a specific mechanism, via which impaired multisensory processes might contribute to the development of psychosis. This approach addressed question 4 of the outlined open questions.

# Chapter II: Emotional Multisensory Integration in Psychosis Proneness (Study 1)

# 1 Introduction

## 1.1 Multisensory Processing of Emotional Information

In our daily life, we constantly process social information to take part in society. Social information can be verbal, such as spoken or written language, but also non-verbal, such as emotional face expressions or vocal prosody. Most of the time, we perceive social cues via multiple sensory modalities, which need to be adequately processed. For example, it is crucial to adequately identify another person's facial expression or vocal utterances in order to induce appropriate reactions during social interactions.

Social information is often presented in noisy surroundings, e.g. in crowded spaces, and comes with a high inter-individual ambivalence in terms of meaning and intent. In noisy social situations, it is advantageous to process cues conveyed via multiple modalities for recognition of social information. In line with the principle of inverse effectiveness (Meredith & Stein, 1986), this crossmodal performance benefit increases when unimodal social stimuli are highly noisy and ambivalent (e.g. Izen & Ciaramitaro, 2020; Tatz et al., 2021). This highlights the importance of appropriate integration of crossmodal social cues for daily social functioning.

Prior studies of emotional MSI in healthy subjects showed that emotion recognition accuracy increased and response time decreased when emotional AV information was presented emotionally congruent compared to unimodal stimuli. Presenting emotionally incongruent AV stimuli lead to reduced recognition accuracy and higher reaction times (RT) compared to unimodal stimuli (e.g. Collignon et al., 2008; Föcker et al., 2011). These congruency facilitation and incongruency interference effects are thought to reflect emotional MSI and seemed to be automatic and relatively independent of selective attention towards either sensory modality (Collignon et al., 2008; Föcker et al., 2011; Föcker & Röder, 2019). However, a recent study by Gao et al. (2021) found evidence for increased crossmodal integration effects when participants were instructed to attend the auditory stimulus compared to when attending the visual stimulus, which was discussed as suggesting a visual dominance in AV emotion recognition. Crucially, modality dominance in multisensory processing is not rigid and flexibly adapts to stimulus reliability. A general principle in multisensory processing is that the more reliable modality dominates MSI. When the signal-to-noise ratio in the usually dominant modality significantly decreased, the crossmodal impact of the usually non-

dominant modality increased, eventually leading to a shift in modality dominance (e.g. Alais & Burr, 2004; Collignon et al., 2008; Piwek et al., 2015).

### 1.2 Psychosis and Emotion Processing

In patients with the diagnosis of a psychotic disorder, deficits in social cognition are well-documented (for reviews, see Healey et al., 2016; Savla et al., 2013). A large body of studies on social cognition in psychosis investigated facial emotion recognition. These studies indicated that the recognition of unimodally presented affective facial expressions was deficient in patient samples (for reviews, see Barkl et al., 2014; Kohler et al., 2010), with patients showing significantly lower recognition rates or mislabeling emotional content of facial expressions compared to healthy controls (e.g. Darke et al., 2021; Tripoli et al., 2022). Further, first-episode patients have been found to show neural hypoactivation towards emotional face stimuli in neuroimaging studies, suggesting impaired neural processing of affective information in psychosis (for a review, see Lukow et al., 2021). While not being as extensively studied as facial emotion processing, several prior studies showed that processing of unimodally presented emotional vocal prosody also was disturbed in patients with the diagnosis of a psychotic disorder, with patients showing difficulties in identifying and discriminating target emotions in vocal prosody (for reviews, see Gong et al., 2021; Lin et al., 2018).

The magnitude of emotion processing deficits in psychosis might differ depending on the specific target emotion (Barkl et al., 2014). Recent studies indicated that aberrances of emotion processing in patients with the diagnosis of a psychotic disorder were more pronounced for negative emotions compared to positive emotions, especially for fear and anger (Tripoli et al., 2022; Won et al., 2019). Emotion processing aberrances were also present with regard to affectively neutral information, with patients erroneously rating neutral face expressions as having emotional content (Mitrovic et al., 2020). Compared to healthy controls, affectively neutral information was rated as more intense in patient samples (Dondaine et al., 2014) as well as mislabeled as negative more often in psychosis proneness (Seo et al., 2020). Further, patients with the diagnosis of a psychotic disorder have shown interference effects during emotion recognition, i.e. a tendency to confuse non-target emotions with the target emotion (Dondaine et al., 2014), as well as a negativity bias, i.e. a tendency to erroneously rate the target emotion as negatively valenced (e.g. Pinkham et al., 2011). In sum, psychosis seems to be associated with especially negative-biased emotion

processing deficits, which might play a crucial role in experiencing psychotic symptoms. Supporting this assumption, a recent study by Krkovic et al. (2020) revealed a predictive value of negative emotions, especially fear, for persecutory ideation in both patients and subjects with psychosis proneness. Further, experiencing negative emotions in turn also followed persecutory beliefs, which suggests a vicious cycle between emotion processing and specific psychotic symptoms (Krkovic et al., 2020). This highlights the potential detrimental effects of emotion processing deficits, especially concerning the bias towards negative emotions, for individuals suffering from psychosis.

Deficient processing of the emotional content of face or voice stimuli has been shown to be associated with various psychotic symptoms like hallucinations and delusions, (Tseng et al., 2013), as well as to be related to problems in social functioning (Brittain et al., 2010; Pinkham et al., 2003). Further, Chapellier et al. (2022) recently showed that in stable out-patients with the diagnosis of a psychotic disorder deficient processing of non-verbal social cues was associated with impaired functional outcomes, such as engagement in everyday activities and skills related to work or personal care, putting emphasis on the importance of functional processing of social information for everyday life.

Impaired processing of affective facial expressions or vocal prosody have also been identified in persons with high psychosis proneness (Seo et al., 2020; van Donkersgoed et al., 2015) and first-degree relatives of patients with the diagnosis of a psychotic disorders (D. Martin et al., 2020). In several studies, persons with high psychosis proneness showed dysfunctional recognition and discrimination of affective facial expressions and vocal prosody (for a review, see van Donkersgoed et al., 2015). Further, Seo et al. (2020) recently found a negativity bias in facial expression recognition, i.e. a bias to rate neutral face expressions as negative, as well as associations of impaired emotion recognition with higher scores in schizotypy and paranoia in subjects with high psychosis proneness. Corcoran et al. (2015) found that emotion recognition deficits have predictive value for the transition from high proneness status to full-blown psychosis, suggesting a potential role of dysfunctional emotion processing in the development of psychosis. Supporting this assumption, it has been discussed that dysfunctional social information processing, such as facial or vocal emotion processing deficits, might constitute a risk factor for the development of psychotic symptoms (D. Martin et al., 2020; Seo et al., 2020; Tripoli et al., 2022; Tseng et al., 2016).

## 1.3 Psychosis and Emotional Multisensory Processing

While social cognition in terms of unisensory processing has been extensively studied in psychosis, fewer studies have examined multisensory processing of social stimuli. Several of these studies investigated processing of verbal social cues, i.e. crossmodal speech stimuli, and consistently revealed impairments in MSI of verbal social cues in patients with the diagnosis of a psychotic disorder (for a review, see Tseng et al., 2015). For example, studies using the McGurk illusion (McGurk & MacDonald, 1976) indicated that patients show a significantly reduced proportion of fused percepts compared to controls (de Gelder et al., 2003; Pearl et al., 2009), suggesting impaired MSI of AV linguistic information in psychosis.

While linguistic MSI in psychosis has been consistently shown to be deficient in prior studies, processing of arbitrary AV stimuli seemed to be rather intact. These findings might reflect a pronounced deficit in processing crossmodal stimuli with a communicative value, i.e. social stimuli, compared to non-social stimuli (Tseng et al., 2015). This suggests specific difficulties regarding multisensory processing in social situations, with possible detrimental effects on evaluating social encounters and potentially aggravating interpersonal problems in psychosis (Tseng et al., 2015).

In contrast to consistent evidence of impaired MSI of verbal social cues in psychosis (Tseng et al., 2015), empirical evidence regarding MSI of non-verbal social cues, i.e. emotional stimuli, in psychosis has been rather inconsistent in previous studies (for reviews, see Lin et al., 2020; Tseng et al., 2015). Compared to healthy controls, patient samples showed either impaired MSI of emotional information (Castagna et al., 2013; de Gelder et al., 2005, Exp.1; de Jong et al., 2009, 2010; Fiszdon & Bell, 2009; Jeong et al., 2021; Mangelinckx et al., 2017; Portnova et al., 2021; Seubert, 2010), no group differences in emotional MSI (Müller et al., 2014; Sestito et al., 2013; Simpson et al., 2013; Zvyagintsev et al., 2013) or even an increased crossmodal influence during emotional MSI (de Gelder et al., 2005, Exp.2; Van den Stock et al., 2011).

Multiple factors might have contributed to these inconsistent findings: varying levels of symptom load and cognitive dysfunctions due to heterogeneous patient samples, potential underlying unisensory processing deficits, aberrant attentional effects, differences in experimental methods complicating comparability, or differences in target emotions or modalities, (Lin et al., 2020; Tseng et al., 2015).

In line with the latter, de Gelder et al. (2005) found opposite group differences between patients with the diagnosis of a psychotic disorder and healthy controls depending on the task-relevant sensory modality: compared to healthy controls, patients showed a reduced influence of vocal prosody on categorizing emotional facial expressions, while an excessive influence of facial expressions on categorizing vocal prosody was found. de Gelder et al. (2005) discuss their findings as suggesting an increased visual dominance during AV emotion processing in patients with the diagnosis of a psychotic disorder. Supporting these findings, Sestito et al. (2013) found that patients relied more on the visual than the auditory cue when facial expressions and vocal prosody were emotionally incongruent. In contrast, other authors presented findings of reduced crossmodal influence of emotional face expressions on judging vocal prosody (de Jong et al., 2009) or an increased crossmodal influence of vocal prosody on facial expression categorization (Van den Stock et al., 2011), suggesting an auditory dominance effect in psychosis. In line with a potential auditory predominance, Thaler et al. (2013) showed that patients with the diagnosis of a psychotic disorder showed a similar positivity mislabeling bias in AV emotion recognition as in auditory only trials, while visual trials were mislabeled more negatively-biased. In sum, while prior studies suggested visual dominance during AV emotion recognition in healthy participants (Collignon et al., 2008; Gao et al., 2021), findings on modality dominance in psychosis are inconsistent (Lin et al., 2020; Tseng et al., 2015), supporting either visual dominance (de Gelder et al., 2005; Sestito et al., 2013) or auditory dominance (de Jong et al., 2009; Thaler et al., 2013; Van den Stock et al., 2011). Possibly, modality dominance during AV emotion processing might be more variable in patients with the diagnosis of a psychotic disorder than the general population, potentially contributing to the inconsistency of findings on aberrant emotional MSI in previous studies (Lin et al., 2020; Tseng et al., 2015).

To sum up, prior evidence suggested altered MSI of emotional-laden information in psychosis, although the results did not support clear conclusions regarding the direction of impairments. Possibly, disturbances in emotional MSI in psychosis might not be generalizable as being generally reduced or excessive. Further, they might as well largely depend on the situation, i.e. the given task (Tseng et al., 2015).

The inconsistency of evidence regarding emotional MSI in psychosis does not lend support for the assumption that psychosis is associated with deficient MSI in terms of generally insufficient integration, assumed to be even more pronounced in social compared to

non-social stimuli (Tseng et al., 2015). However, the majority of prior studies suggested at least dysfunctional aberrant emotional MSI in psychosis. Apart above mentioned factors, which potentially influenced the inconsistent results, it might be speculated that impairments of emotional MSI in psychosis might not manifest in a one-dimensional, generalized deficit but rather a multi-faceted, dysfunctional alteration in processing crossmodal emotional stimuli. In that perspective, dysfunctional emotional MSI in psychosis might express as reduced integration under certain circumstances and excessive in other circumstances, with potential maladaptive cognitive or behavioral consequences in either case.

### 1.4 A Potential Role of Dysfunctional Emotional MSI for Psychosis Development

While theoretical approaches regarding the influence of dysfunctional emotional MSI on psychosis speculated that it might play a role in disorder development (Tseng et al., 2015), previous studies of emotional MSI in psychosis mostly investigated samples of patients who had already been diagnosed with a psychotic disorder. However, to our knowledge no study of emotional MSI in a sample with psychosis proneness has been published.

Giannitelli et al. (2015) investigated emotional MSI in children with the diagnosis of early-onset schizophrenia (EOS). Their findings of poorer uni-and crossmodal performance compared to healthy controls mirror the deficit pattern in adult patients (Giannitelli et al., 2015). This suggested that emotional MSI impairments might already be present in the early developmental trajectory of psychosis[1]. However, caution in generalizing the findings of Giannitelli et al. (2015) to the development trajectory of adult-onset psychosis might be advised: EOS is a particularly rare form of schizophrenia (Driver et al., 2013) has been found to show significant higher premorbid impairments and symptom load (e.g. Vyas et al., 2011). This raises the question if the development trajectories concerning emotion processing deficits are sufficiently similar between EOS and adult-onset psychosis.

Further, several studies on unimodal emotion processing indicated deficient recognition of facial and vocal affect in psychosis proneness (van Donkersgoed et al., 2015). Since MSI reflects a non-linear gain mechanism beyond simple addition of unisensory processes (Stein et al., 2010), it can be assumed that deficits in MSI might reflect specific dysfunctions beyond summation of unisensory processing deficits.

[1] Throughout this chapter, the term „psychosis“ refers to adult-onset psychosis. EOS will be referred to as such to avoid confusion between both types of disorder manifestation.

In sum, previous studies of unisensory emotion processing in psychosis proneness as well as prior findings of deficient emotional MSI in EOS offer hints towards potential impairments of emotional MSI in psychosis proneness. However, to contribute to the assumptions of deficient MSI of social cues constituting a risk factor in the development of psychotic symptoms, studies with proneness samples or first-degree relatives are needed.

### 1.5 Research Question

Therefore, in this study we aimed to investigate MSI of emotional AV information in a healthy sample with high psychosis proneness compared to a sample with low proneness. Subjects categorized the emotional content of facial expressions as well as vocal prosody and rated the respective perceived emotional intensity while one modality was attended and the other ignored (paradigm adopted from Föcker et al., 2011).

In line with findings of altered processing of affective laden AV information in patient samples as well as deficient unisensory facial or vocal emotion processing deficits in at-risk samples, we expected the high proneness group to show aberrant emotion categorization performance benefits and interference by crossmodal emotional in-/congruency compared to a low proneness group, reflecting atypical emotional MSI. Further, we expected in-/congruency effects during emotion categorization to be associated with self-reported psychotic experiences such as everyday anomalous percepts, hallucinatory experiences and paranoid ideation. Since the interpretation of perceptual anomalies is influenced by negative schemata about the self and others (Garety et al., 2001), we also expected an association between in-/congruency effects during categorization performance and negative schemata.

# 2 Methods

## 2.1 Participants

We recruited healthy participants from the local community and amongst university students. After initial contact, participants received a link leading to an online screening (see **2.3 General Procedure**), which served to allocate subjects to either a high proneness group or a low proneness group by means of a cut-off value (≥9) of the positive subscale score of the Community Assessment of Psychic Experiences (CAPE; Stefanis et al., 2002).

Subjects with a present or past diagnosis of a mental disorder, especially of a psychotic disorder, bipolar disorder or autism, as well as acute suicidality, neurological disorders, regular and/or recent consumption of psychedelic substances, uncorrected sight or hearing impairment or inability to follow task instructions were excluded. Since empirical evidence indicated higher occurrence of substance abuse (Addington et al., 2014), anxiety (McAusland et al., 2017) and depression (Verdoux et al., 1999) in samples with proneness to psychosis, we did not exclude subjects with mild forms of substance abuse, anxiety disorders and depressive disorders.

87 subjects fulfilled our criteria and were invited to the experimental session. We had to exclude $n = 3$ in the high and $n = 3$ in the low proneness group due to technical difficulties during data collection, $n = 1$ in the high and $n = 5$ in the low proneness group due to inability to follow task instructions (see **2.4 Data Analysis**), and n = 2 in the high proneness and $n = 1$ in the low proneness group since behavioural performance deviated more than two *SD* from the respective group mean.

The final sample consisted of $N = 72$ participants, with $n = 36$ in the high proneness group and $n = 36$ in the low proneness group. There were no associations between group and both gender, $X^2(2) = 2.16$, $p = .339$, $\varphi = 0.17$, and education level, $X^2(2) = 3.01$, $p = .222$, $\varphi = 0.21$, and groups did not differ in age, $t(70) = 0.87$, $p = .385$, Cohen's $d = 0.21$ (see Table 1 for demographic data). Participants gave written informed consent prior to participation and received 10€/hour or course credits as compensation. This study was approved by the Local Ethics Committee of the University of Hamburg.

***Table 1***

*Demographic and Diagnostic Data*

| | High Proneness | Low Proneness |
|---|---|---|
| *n* | 36 | 36 |
| Age, *M* (range) | 24.8 (19-39) | 25.9 (18-42) |
| Gender, *n* (%) | | |
| female | 29 (80.6) | 26 (72.2) |
| male | 6 (16.7) | 10 (27.8) |
| diverse | 1 (2.8) | 0 |
| Education, *n* | | |
| 10 years | 1 | 2 |
| 12-13 years [a] | 27 | 20 |
| university degree [b] | 8 | 14 |
| CAPE positive, *M* (*SD*) | 14.0 (6.2) | 4.3 (2.1) |
| CAPS, *M* (*SD*) | 11.3 (6.4) | 5.7 (4.3) |
| LSHS-E, *M* (*SD*) | 15.3 (9.7) | 8.6 (7.1) |
| PCL, *M* (*SD*) [c] | 47.3 (27.1) | 34.8 (26.3) |
| BCSS n.S., *M* (*SD*) | 4.2 (3.1) | 3.3 (2.7) |
| BCSS n.O., *M* (*SD*) | 4.8 (3.3) | 3.8 (3.9) |

*Notes.* $N = 72$. CAPE positive = Community Assessment of Psychic Experiences, positive subscale; CAPS = Cardiff Anomalous Perception Scale, total score; LSHS-E = Launey-Slade Hallucination Scale-Extended, total score; PCL= Paranoia Checklist, total score; BCSS n.S. = Brief Core Schema Scales, negative self subscale; BCSS n.O. = Brief Core Schema Scales, negative other subscale.

[a] due to changes in German education law in 2007, the years of education for acquire the Abitur (A-level/high school equivalent) vary between 12-13 years depending on school and Bundesland; [b] university degree: Bachelor level or higher. [c] one subject in the low and one in the high proneness group had ≥50% missings in the PCL. Their PCL scores were corrected by the respective group mean.

## 2.2 Stimuli and Experimental Paradigm

Experimental procedure was adopted from Föcker et. (2011), using the Software Presentation® (Version 21.1; Neurobehavioural Systems Inc., Berkeley, California, USA). In total, 216 emotionally laden AV stimuli were presented centrally on a computer screen (61cm diameter, screen resolution: 1920 x 1200) at 65cm viewing distance (4° width and 9° height of visual angle) and via computer screen loudspeakers (60 to 65 dB, measured at participants' head position). Stimuli consisted of visual and acoustic recordings of 4 actors (2 female, 2 male) uttering bisyllabic German pseudowords ("lolo", "tete", "gigi"). Visual stimuli consisted of video recordings of the actor's inner facial features during pseudoword utterance, acoustic stimuli of the auditory stream of the pseudoword utterance (for stimulus creation and selection, see Föcker et al., 2011). Face expressions and prosody contained one of 4 emotions: happy, angry, sad or neutral. Stimuli were either unimodal visual (mean duration 761ms, range 500-1400), unimodal acoustic (mean duration 761ms, range 500-1400), bimodal emotionally congruent (mean duration 946ms, range 734-1468) or bimodal emotionally incongruent (mean duration 904ms, range 633-1501), with 48 stimuli per stimulus condition. For both bimodal conditions, the auditory stream and the video originated from different recordings. Additionally, visual and acoustic deviant stimuli (6 per stimulus condition) were presented to check whether participants paid attention to the task over the course of the experiment. Visual deviants included a black dot (0,6°/ 6mm diameter, duration 100ms) in one of 4 possible locations (forehead, nose, left or right cheek) during last 130-330ms of the video, acoustic deviants one of four tones (600/900/1200 or 1500Hz, duration 100ms) during the last 130-330ms of the auditory stream.

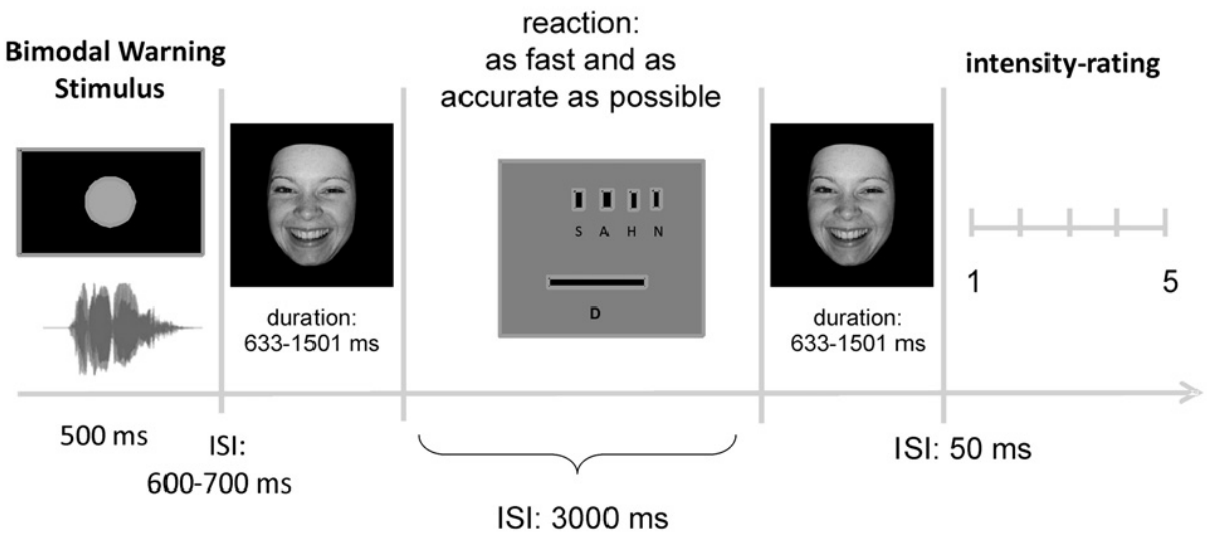
After a bimodal warning stimulus (grey circle with 2° of visual angle combined with multispeaker noise, duration = 500ms), each stimulus was presented twice in a row after a random Inter-Stimulus-Interval (600-700ms, uniform distribution), Following the first onset, subjects were asked to categorize the perceived emotion as fast and accurately as possible by button press on a custom German keyboard. Each of the four buttons was designated to one of the four emotions as well as to one of four fingers of the right hand (index to little finger). The assignment order of emotion category to the buttons was balanced across participants. After an Inter-Stimulus-Interval of 3000ms, the same stimulus was presented a second time. After offset, a rating scale appeared on the screen and participants had to rate the intensity of the perceived emotion from 1 (very low) to 5 (very high) by pressing the respective key (1-5) on

the keyboard. Immediately after intensity rating, the next trial started. In case of deviants, participants had to press the space bar following the first presentation to indicate that they recognized the deviant. After pressing the space bar, the next trial started immediately (see Figure 2 for a depiction of the trial structure).

Stimuli were presented randomly in 12 blocks of 27 stimuli each. For half of the blocks, participants were asked to attend to the face and ignore the voice and vice versa. Unimodal stimuli were only used in blocks in which the corresponding modality had to be attended. The order of blocks was randomized across participants. Half of the participants started with attend face, the other half with attend voice. After each block, participants could take a short break of 2-3 minutes.

***Figure 2***

*Experimental Design*




*Note.* ISI = Inter-Stimulus-Interval. Capital letters represent buttons for emotion categorization: S = sad, A = angry, H = Happy, N = Neutral, D = deviant recognition via space bar. Adopted from "Preattentive processing of audio-visual emotional signals" by J. Föcker, M. Gondan, & B. Röder, 2011, *Acta Psychologica, 137*(1), 36-47 (https://doi.org/ 10.1016/j.actpsy.2011.02.004).

## 2.3 General Procedure

Prior to the experimental session, subjects completed an online screening. In the screening, demographic data was collected and in- and exclusion criteria were checked. The main aim of the screening was administering the 28-item version of the CAPE (Stefanis et al., 2002), a self-report questionnaire measuring life-time occurrence of psychotic symptoms, as a measure for proneness to psychosis. We applied a cut-off value of (≥9) of the positive

subscale to assign subjects to one of two subgroups (high proneness ≥9 & low proneness <9). This value has been used previously to distinguish subjects with low and high proneness to psychosis (Krkovic et al., 2020).

The experimental session took place in a sound attenuated and dimly lit room at the University of Hamburg. Prior to the experimental run, baseline diagnostic measures were conducted. These included personal and familiar history of psychotic disorders, regular medication intake, recent substance consumption, sight or hearing impairment and respective corrections, handedness, the Trail Making Test (TMT) Version A and Version B (Reitan & Wolfson, 1985) for measuring attention and processing speed and the Continuous Performance Test – Identical Pairs (CPT-IP; Nuechterlein et al., 2008) for measuring vigilance.

Subsequently, participants completed the experimental paradigm (adopted from Föcker et al., 2011) as described above. 10 practice trials were conducted prior to the actual experimental run to familiarize participants with the task. The experimental run took approximately 50 min to complete.

After the experimental run, participants completed the following self-report instruments: The Launey-Slade Hallucination Scale-Extended (LSHS-E; Siddi et al., 2019; Lincoln et al., 2009) for measuring hallucinatory experiences, the Paranoia Check-List (PCL; Freeman et al., 2005) for measuring paranoid thoughts and convictions, the Cardiff Anomalous Perception Scale (CAPS; Bell et al., 2006) for measuring anomalous perceptual experiences and the Brief Core Schema Scales (BCSS; Fowler et al., 2006) for measuring positive and negative schemata about the self and others (a glossary of the instruments and questionnaires used in this dissertation project can be found in Table A.1 in Appendix A).

## 2.4 Data Analysis

To examine group differences in age or in distributions of gender and education level, we calculated an unpaired t-test for age and $X^2$-tests for gender and education level.

For each subject, experimental data was preprocessed using RStudio Version 1.0.143 (RStudio Team, 2016) by calculating mean RT (in ms) and accuracy (proportion of correct emotion categorizations, in %) for categorization data as well as mean perceived emotional intensity, separately per stimulus condition (unimodal, bimodal congruent and bimodal incongruent) and attention condition (attend face vs. attend voice). Incorrectly categorized

stimuli or responses latencies below 200ms and above 3950ms were excluded from analysis. Deviant trials were analyzed separately (see paragraph below). To account for a possible speed-accuracy trade-off, we calculated inverse efficiency (IE) scores for categorization data by dividing mean RT by accuracy (Townsend & Ashby, 1978). Statistical analyses were performed using IBM SPSS Statistics for Windows, Version 27.0 (IBM Corp., 2020).

To examine whether participants followed task instructions, we calculated the percentage of total recognized deviants for each participant. If less than 60 percent were recognized, the participant was excluded from analysis. Further, we calculated an unpaired *t*-test to check for group differences in deviant recognition.

IE scores were analyzed for group differences by means of mixed ANOVAs with the between-factor *Group* (high vs. low proneness) and the within-factor *Stimulus Condition* (unimodal vs. bimodal emotionally congruent vs. bimodal emotionally incongruent). Since we were not primarily interested in effects of attention condition for our hypotheses, we ran one mixed ANOVA each for both attention conditions. We applied the Huynh-Feldt correction (Huynh & Feldt, 1976) in case of sphericity violations. Post-hoc pairwise comparisons were calculated by means of Bonferroni corrected *t*-test. The same analyses were run for mean perceived intensity and, following the analysis strategy of Föcker et al. (2011) additionally for mean RT and accuracy (see Appendix B).

Further, we explored if both groups differed in emotion recognition performance regarding the target emotion. Therefore, we additionally calculated mean RT, accuracy and mean perceived intensity separately for each of the four target emotions per participant. IE scores could not be calculated due to low accuracy scores in some target emotion * stimulus condition * attention condition combinations. We observed that calculating IE scores would result in very high extreme values and violations of analysis requirements. Mean RT, accuracy and mean perceived intensity were exploratory analyzed by means of mixed ANOVAs with the factors *Group* and *Stimulus Condition*, separately for each target emotion and attention condition, followed by Bonferroni-corrected post-hoc t-tests (for an overview over inferential statistics of all emotion-specific mixed ANOVAs, *F*-statistic, *p*-values and $\eta^2_p$ for each mixed ANOVA are additionally reported in Table B.2 in Appendix B). Post-hoc comparisons for the within-factor *Stimulus Condition* are reported between congruent and unimodal trials, regarding a possible congruency facilitation effect (lower RT, higher

accuracy, and higher perceived emotional intensity in congruent compared to unimodal trials), and between incongruent and unimodal trials, regarding a possible incongruency interference effect (higher RT, lower accuracy, and lower perceived emotional intensity in incongruent compared to unimodal trials). Comparisons between congruent and incongruent trials are not informative for the research question and are hence not reported for readability reasons.

Potential group differences in attention and processing speed were analyzed by means of unpaired t-Tests entering TMT-A and -B completion times. To check for group differences in vigilance, CPT-IP *d'* scores, a sensitivity score reflecting attentional capacity (Cornblatt et al., 1988), were analyzed with a mixed ANOVA with the between factor *Group* (high vs. low proneness) and the within-factor *Digit Load* (2 vs. 3 vs. 4 digits), followed by post-hoc t-tests. In case of significant group differences in the TMT or CPT-IP, we assessed the influence of the control variable on the main dependent variable (IE) by means of Pearson's correlation coefficients *r*.

CAPS, LSHS-E, PCL, BCSS negative self and BCSS negative other scores were compared between high and low proneness groups by means of unpaired *t*-tests, separately for each questionnaire.

To examine associations between in-/congruency effects and psychotic experiences as well as negative schemata, we first calculated IE-difference scores between the IE scores of each stimulus condition pairing (congruent - unimodal, unimodal - incongruent, congruent - incongruent), separately for both attention conditions. We then calculated Bonferroni-corrected Pearson correlation coefficients (*r*) between IE-difference scores and CAPS, LSHS-E, PCL as well as BCSS negative subscale scores.

Additionally, we calculated Bonferroni-corrected Pearson's correlation coefficients (*r*) between IE scores and CAPS, LSHS-E, PCL, as well as BCSS negative subscale scores to explore associations between emotional categorization performance and psychotic experiences as well as negative schemata. For all correlations, Pearson's correlation coefficients $r > .3$ were highlighted as potentially meaningful for further discussion.

# 3 Results

## 3.1 Experimental Data

### *3.1.1 Deviant Recognition*

Groups did not differ in percentage of recognized deviants (high proneness: $M$ = 88.65, $SD$ = 10.23; low proneness: $M$ = 84.61, $SD$ = 10.33), $t(70) = -1.67$, $p = .099$, Cohen's $d$ = 0.39.

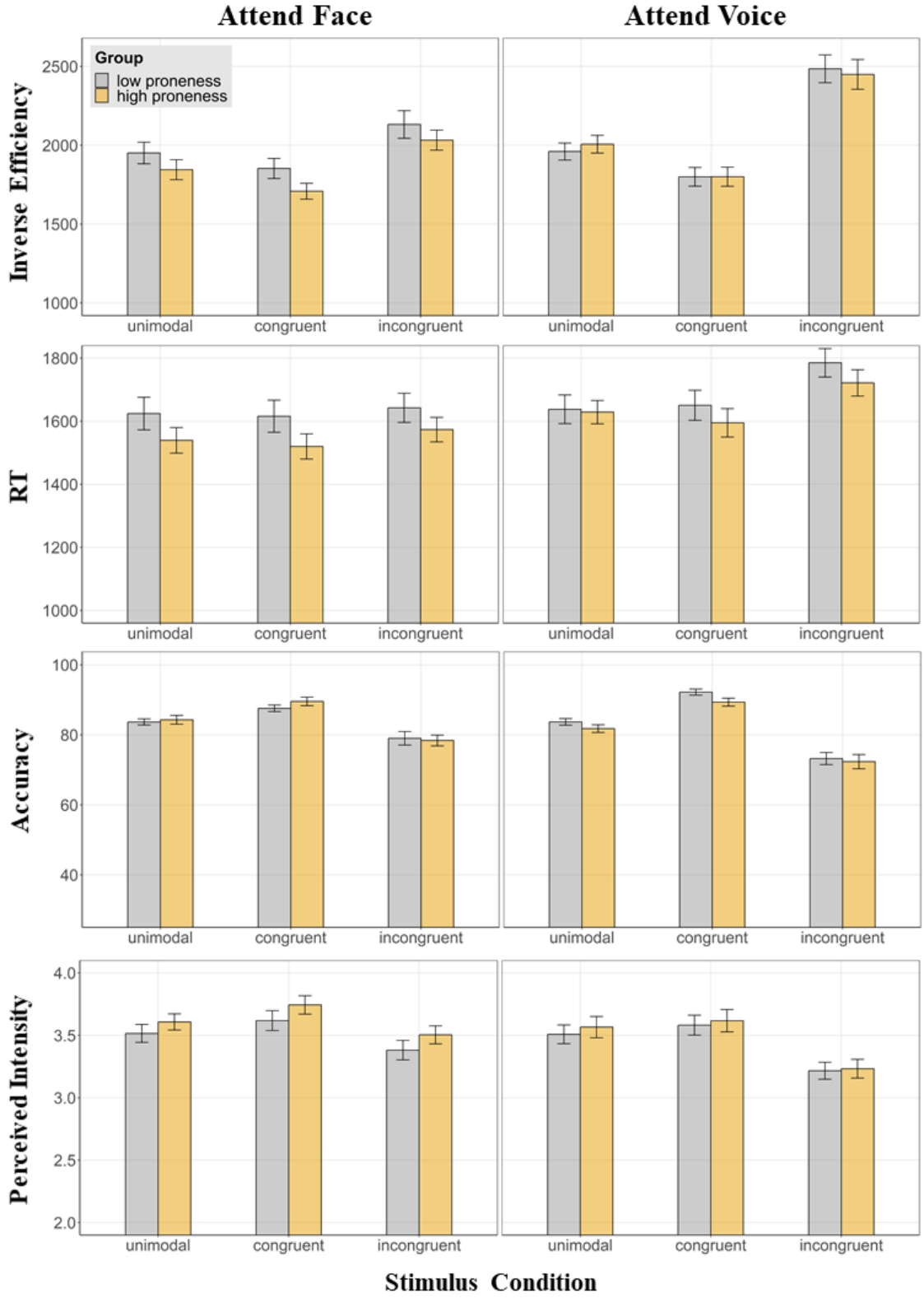
### *3.1.2 Inverse Efficiency Analysis*

**Attend Face.** A 2 (*Group*) x 3 (*Stimulus Condition*) mixed ANOVA for IE scores in attend face blocks revealed a significant main effect of *Stimulus Condition*, $F(1.6, 112.07) = 35.94$, $p < .001$, $\eta^2_p = 0.34$, indicating lower IE-scores in the congruent condition ($M$ = 1780.67, SD = 348.64) compared to both the unimodal ($M$ = 1898.02, SD = 395.64), $t(71) = -4.70$, $p < .001$, and the incongruent condition ($M$ = 2082.07, SD = 454.11), $t(71) = -7.13$, $p < .001$, as well as lower IE scores in the unimodal condition compared to the incongruent condition, $t(71) = -4.92$, $p < .001$. There was no main effect of *Group*, $F(1, 70) = 1.88$, $p = .175$, $\eta^2_p = 0.03$, or *Group* * *Stimulus Condition* interaction, $F(1.6, 112.07) = 0.23$, $p = .748$, $\eta^2_p = 0.003$ (see Figure 3 top row for IE scores per group and stimulus condition).

**Attend Voice.** A 2 (*Group*) x 3 (*Stimulus Condition*) mixed ANOVA for IE scores in attend voice blocks revealed a significant main effect of *Stimulus Condition*, $F(1.26, 88.05) = 120.59$, $p < .001$, $\eta^2_p = 0.63$, indicating lower IE-scores in the congruent condition ($M$ = 1800.01, $SD$ = 357.85) compared to both the unimodal ($M$ = 1983.38, $SD$ = 327.15), $t(71) = -8.64$, $p < .001$, and the incongruent condition ($M$ = 2466.80, $SD$ = 544.06), $t(71) = -12.19$, $p < .001$, as well as lower IE scores in the unimodal condition compared to the incongruent condition, $t(71) = -9.84$, $p < .001$. There was no main effect of *Group*, $F(1, 70) = 0.002$, $p = .964$, $\eta^2_p < 0.001$, and no *Group* * *Stimulus Condition* interaction, $F(1.26, 88.05) = 0.43$, $p = .559$, $\eta^2_p = 0.006$ (see Figure 3 top row for IE scores per group and stimulus condition).

***Figure 3***

*Inverse Efficiency Scores, mean RT, Accuracy and mean Perceived Emotional Intensity per Group and Stimulus Condition, Separated by Attention Condition*



*Notes.* $n_{low}$ = 36, $n_{high}$ = 36. Error bars denote ±1 standard error of the mean. congruent = bimodal emotionally congruent; incongruent = bimodal emotionally incongruent.

**Exploratory Analysis with Attention as Within-Factor.** To check for potential attention-specific group differences, we calculated an exploratory mixed ANOVA with the between-factor *Group* (high vs. low proneness) and the within-factors *Stimulus Condition* (unimodal vs. bimodal emotionally congruent vs. bimodal emotionally incongruent) and *Attention* (attend face vs. attend voice). There were significant main effects of the within-factors *Stimulus Condition*, $F(1.34, 93.6) = 121.2$, $p < .001$, $\eta^2_p = 0.63$, and *Attention*, $F(1, 70) = 38.5$, $p < .001$, $\eta^2_p = 0.36$, indicating overall lower IE-scores in attend face blocks ($M$ = 1920.26.7, $SD$ = 363.11) compared to attend voice blocks ($M$ = 2083.40, $SD$ = 360.80). *Stimulus Condition* and *Attention* significantly interacted with each other, $F(1.45, 101.33) = 30.92$, $p < .001$, $\eta^2_p = 0.31$. Post-hoc $t$-tests for the *Stimulus Condition* * *Attention* interaction revealed lower IE scores in the attend face compared to the attend voice condition in incongruent trials (face: $M$ = 2082.07, $SD$ = 457.12; voice: $M$ = 2466.80, $SD$ = 544.06) $t(71) = -6.91$, $p < .001$, and in unimodal trials (face: $M$ = 1898.02, $SD$ = 395.64; voice: $M$ = 1983.39, $SD$ = 327.15), $t(71) = -3.18$, $p = .002$, but not in congruent trials (face: $M$ = 1780.68, $SD$ = 348.64; voice: $M$ = 1800.01, $SD$ = 357.85), $t(71) = -0.68$, $p = .499$. Further, there was a significant interaction between *Group* * *Attention*, $F(1, 70) = 5.26$, $p = .025$, $\eta^2_p = 0.07$. Post-hoc $t$-tests for the *Group* * *Attention* interaction revealed that the high proneness group showed significantly lower IE scores in attend face trials ($M$ = 1861.91, $SD$ = 323.51) compared to attend voice trials ($M$ = 2085.36, $SD$ = 375.34), $t(35) = -6.55$, $p < .001$, Cohen's $d = -1.09$. For the low proneness group, the difference in IE scores between attend face ($M$ = 1978.60, $SD$ = 394.73) and attend voice trials ($M$ = 2081.44, $SD$ = 350.97) was also significant, $t(35) = -2.57$, $p = .015$, Cohen's $d = -.043$, but smaller compared to the high proneness group. There was no main effect of *Group*, $F(1, 70) = 0.48$, $p = .49$, $\eta^2_p = 0.007$, no *Group* * *Stimulus Condition* interaction, $F(1.34, 93.6) = 0.26$, $p = .677$, $\eta^2_p = 0.004$, and no *Group* * *Stimulus Condition* * *Attention* interaction, $F(1.45, 101.33) = 0.49$, $p = .554$, $\eta^2_p = 0.007$.

**Correlation Between IE Difference Scores.** Pearson correlation coefficients calculated over all subjects revealed highly significant positive correlations between the IE difference score congruent – incongruent and both the IE difference scores congruent – unimodal and unimodal – incongruent, respectively, for both the attend face and attend voice trials, all $r > .4$, all Bonferroni-corrected $p < .001$. There was no other significant correlation

between IE difference scores, all $p > .05$ (see Table 2 for Pearson's r for correlations between IE difference scores).

**Correlations Between IE Scores.** Pearson correlation coefficients over all subjects showed that IE scores for all stimulus conditions and both attention conditions revealed highly significant, strong correlations, all $r \geq .48$ and Bonferroni-corrected $p < .001$ (see Table B.1 in Appendix B for Pearson's *r* for correlations between IE scores).

***Table 2***

*Pearson Correlation Coefficients r Between IE Difference Scores and Questionnaire Data*

| | | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. | 9. | 10. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Attend Face | 1. IE: C-U | | | | | | | | | | |
| | 2. IE: U-I | -.13 | | | | | | | | | |
| | 3. IE: C-I | **.48***** | **.81***** | | | | | | | | |
| Attend Voice | 4. IE: C-U | .21 | .11 | .22 | | | | | | | |
| | 5. IE: U-I | .19 | .17 | .27 | .06 | | | | | | |
| | 6. IE: C-I | .25 | .20 | **.32** | **.44***** | **.92***** | | | | | |
| | 7. CAPS | -.24 | -.03 | -.17 | -.15 | -.03 | -.08 | | | | |
| | 8. LSHS-E | -.19 | .01 | -.1 | -.18 | -.4 | -.1 | **.7***** | | | |
| | 9. PCL [a] | -.21 | .08 | -.05 | .01 | -.2 | -.18 | **.42***** | **.52***** | | |
| | 10.BCSS nS | -.12 | .15 | .06 | .04 | .02 | .03 | .14 | .27 | **.63***** | |
| | 11.BCSS nO | .02 | -.03 | -.02 | -.18 | -.18 | -.23 | .28 | **.40*** | **.52***** | .29 |

*Notes.* $N = 72$. Pearson correlation coefficients $r > .3$ are marked in bold. *p*-values are Bonferroni-corrected. IE = Inverse Efficiency; C-U = Difference in IE scores between congruent and unimodal conditions; U-I = Difference in IE scores between unimodal and incongruent conditions; C-I = Difference in IE scores between congruent and incongruent conditions; LSHS-E = Launey-Slade Hallucination Scale-Extended; PCL total = Paranoia Checklist, total score; CAPS total = Cardiff Anomalous Perception Scale, total score; BCSS nS = Brief Core Schema Scales, negative self subscale; BCSS nO = Brief Core Schema Scales, negative other subscale.

[a] One subject in the low and one in the high proneness group had ≥50% missings in the PCL. Their PCL scores were corrected by the respective group mean.

* $p < .05$, ** $p < .01$, *** $p < .001$

### *3.1.3 Intensity Analysis*

**Attend Face.** A 2 (*Group*) x 3 (*Stimulus Condition*) mixed ANOVA for mean perceived emotional intensity in attend face blocks revealed a significant main effect of *Stimulus Condition*, $F(1.86, 130.44) = 28.42$, $p < .001$, $\eta^2_p = 0.29$, indicating higher intensity ratings in the congruent condition ($M = 3.68$, $SD = 0.46$) compared to both the unimodal ($M = 3.56$, $SD = 0.41$), $t(71) = 4.56$, $p < .001$, and the incongruent condition ($M = 3.44$, $SD = 0.45$), $t(71) = 6.71$, $p < .001$, as well as higher intensity ratings in the unimodal condition compared to the incongruent condition, $t(71) = 3.73$, $p = .001$. There was no main effect of *Group*, $F(1, 70) = 1.36$, $p = .248$, $\eta^2_p = 0.02$ , and no *Group* * *Stimulus Condition* interaction, $F(1.86, 130.44) = 0.17$, $p = .827$, $\eta^2_p = 0.002$ (see Figure 3 bottom row for mean perceived emotional intensity per group and stimulus condition).

**Attend Voice.** A 2 (*Group*) x 3 (*Stimulus Condition*) mixed ANOVA for mean perceived emotional intensity in attend voice blocks revealed a significant main effect of *Stimulus Condition*, $F(1.88, 131.39) = 86.22$, $p < .001$, $\eta^2_p = 0.55$, indicating higher intensity ratings in the congruent condition ($M = 3.60$, $SD = 0.51$) compared to both the unimodal ($M = 3.54$, $SD = 0.48$), $t(71) = 2.41$, $p = .019$, and the incongruent condition ($M = 3.22$, $SD = 0.43$), $t(71) = 10.87$, $p < .001$, as well as higher intensity ratings in the unimodal compared to the incongruent condition, $t(71) = 10.28$, $p < .001$. There was no main effect of *Group*, $F(1, 70) = 0.12$, $p = .731$, $\eta^2_p = 0.002$ , and no *Group* * *Stimulus Condition* interaction, $F(1.88, 131.39) = 0.23$, $p = .783$, $\eta^2_p = 0.003$ (see Figure 3 bottom row for mean perceived emotional intensity per group and stimulus condition).

**Exploratory Analysis with Attention as Within-Factor.** Analogue to the IE analysis, we calculated an exploratory 2 (*Group*) * 3 (*Stimulus Condition*) * 2 (*Attention*) mixed ANOVA for mean perceived emotional intensity. There were significant main effects of the within-factors *Stimulus Condition*, $F(1.68, 117.48) = 90.18$, $p < .001$, $\eta^2_p = 0.56$, and *Attention*, $F(1, 70) = 20.47$, $p < .001$ , $\eta^2_p = 0.23$, revealing overall higher intensity ratings in attend face blocks ($M = 3.56$, $SD = 0.41$) compared to attend voice blocks ($M = 3.45$, $SD = 0.45$). The factors *Stimulus Condition* and *Attention* interacted significantly with each other, $F(2, 140) = 11.88$, $p < .001$ , $\eta^2_p = 0.15$. Post-hoc *t*-tests for the *Stimulus Condition* * *Attention* interaction revealed higher intensity ratings in the attend face compared to the attend voice condition in congruent (face: $M = 3.68$, $SD = 0.46$; voice: $M = 3.60$, $SD = 0.51$), $t(71) = 2.6$, $p = .011$, and incongruent trials (face: $M = 3.44$, $SD = 0.45$; voice: $M = 3.22$, $SD = 0.43$), $t(71) =$

6.16, $p < .001$, but not in unimodal trials (face: $M = 3.56$, $SD = 0.41$; voice: $M = 3.53$, $SD = 0.48$), $t(71) = 0.72$, $p = .475$. There was no main effect of *Group*, $F(1, 70) = 0.57$, $p = .451$, $\eta^2_p = 0.008$, no *Group* * *Stimulus Condition*, $F(1.68, 117.48) = 0.03$, $p = .953$, $\eta^2_p = < 0.001$, no *Group* * *Attention*, $F(1, 70) = 2.57$, $p = .113$, $\eta^2_p = 0.04$, and no *Group* * *Stimulus Condition* * *Attention* interaction, $F(2, 140) = 0.42$, $p = .657$, $\eta^2_p = 0.006$.

### *3.1.4 Reaction Time and Accuracy*

The supplementary analyses of mean RT and accuracy revealed no main effects and interactions of the between-factor *Group* on neither mean RT nor accuracy. Further, within-effects of the factor *Stimulus Condition* comparable to the effects published by Föcker et al. (2011) could be observed on both mean RT and accuracy (see Appendix B for details on the supplementary analyses of mean RT and accuracy).

### *3.1.5 Exploratory Emotion-Specific Analysis*[2]

**Target Emotion Happy.** Separate *Group* * *Stimulus Condition* mixed ANOVAs in attend face trials with target emotion happy revealed neither a main effect of *Group*, $F(1, 70) = 0.93$, $p = .338$, $\eta^2_p = 0.01$, nor a main effect of *Stimulus Condition*, $F(2, 140) = 1.79$, $p = .171$, $\eta^2_p = 0.03$, on mean RT. There was a non-significant trend for a *Group* * *Stimulus Condition* interaction on mean RT in attend face trials, $F(2, 140) = 2.97$, $p = .055$, $\eta^2_p = 0.04$. Post-hoc comparisons revealed a non-significant trend for lower mean RT in the high proneness compared to the low proneness group in congruent happy face trials, $t(70) = -1.83$, $p = .072$, Cohen's $d = -0.43$, but not in unimodal, $t(70) = -0.67$, $p = .505$, Cohen's $d = -0.16$, or incongruent trials, $t(70) = -0.29$, $p = .775$, Cohen's $d = -0.07$. The high proneness group showed a congruency facilitation effect, $t(35) = -2.22$, $p = .033$, Cohen's $d = -0.37$, but no incongruency interference effect $t(35) = 1.11$, $p = .275$, Cohen's $d = 0.19$, on mean RT in happy face stimuli, whereas the low proneness group neither showed a congruency facilitation, $t(35) = 0.41$, $p = .688$, Cohen's $d = 0.07$, nor an incongruency interference effect, $t(35) = 0.12$, $p = .907$, Cohen's $d = 0.02$. There was neither a main effect of *Group*, $F(1, 70) = 2.06$, $p = .156$, $\eta^2_p = 0.03$, nor a *Group* * *Stimulus Condition* interaction, $F(1.74, 121.74) = 0.99$, $p = .366$, $\eta^2_p = 0.01$, on accuracy in attend face trials. The factor *Stimulus Condition* showed a non-significant trend for a main effect on accuracy in attend face trials, $F(1.74, 121.74) = 2.75$, $p = .075$, $\eta^2_p = 0.04$, revealing a congruency facilitation effect on accuracy in happy face stimuli, $t(71) = 2.72$, $p =$

[2] For an overview over inferential statistics of all emotion-specific mixed ANOVAs, *F*-statistic, *p*-value and $\eta^2_p$ for each mixed ANOVA are additionally reported in Table B.2 in Appendix B.

.022, Cohen's $d$ = 0.32 . Further, there was no main effect of *Group*, $F(1, 70) = 1.04$, $p = .312$, $\eta^2_p = 0.02$, and no *Group* * *Stimulus Condition* interaction on mean perceived intensity in attend face trials, $F(2, 140) = 0.6$, $p = .549$, $\eta^2_p = 0.009$. There was a significant main effect of *Stimulus Condition* on mean perceived intensity in attend face trials, $F(2, 140) = 5.46$, $p = .005$, $\eta^2_p = 0.07$, indicating a non-significant trend for a congruency facilitation effect when rating happy face stimuli, $t(71) = 2.30$, $p = .073$, Cohen's $d = 0.27$.

Separate *Group* * *Stimulus Condition* mixed ANOVAs in attend voice trials with target emotion happy revealed no main effect of *Group* on mean RT, $F(1, 68) = 0.51$, $p = .479$, $\eta^2_p = 0.007$. There was a highly significant main effect of *Stimulus Condition*, $F(2, 136) = 42.31$, $p < .001$, $\eta^2_p = 0.38$, and a non-significant trend for a *Group* * *Stimulus Condition* interaction, $F(2, 136) = 3$, $p = .053$, $\eta^2_p = 0.04$, on mean RT in attend voice trials. Post-hoc comparisons revealed similar mean RT for both groups in unimodal, $t(70) = -0.1$, $p = .919$, Cohen's $d = -0.02$, congruent, $t(70) = -0.75$, $p = .454$, Cohen's $d = -0.18$, and incongruent happy voice trials, $t(68) = .1{,}61$, $p = .113$, Cohen's $d = -0.38$. The high proneness group showed both a congruency facilitation, $t(35) = -2.29$, $p = .029$, Cohen's $d = -0.38$, and an incongruency interference effect, $t(34) = 2.63$, $p = .013$, Cohen's $d = 0.44$, on mean RT in happy voice trials, whereas the low proneness group showed an incongruency interference, $t(34) = 6.37$, $p < .001$, Cohen's $d = 1.08$, but no congruency facilitation effect, $t(35) = -0.92$, $p = .363$, Cohen's $d = -0.15$. There was no main effect of *Group* on accuracy in attend voice trials, $F(1, 70) = 0.21$, $p = .645$, $\eta^2_p = 0.003$. There was a highly significant main effect of *Stimulus Condition*, $F(1.63, 114.25) = 153.04$, $p < .001$, $\eta^2_p = 0.69$, and a non-significant trend for a *Group* * *Stimulus Condition* interaction, $F(1.63, 114.25) = 2.54$, $p = .094$, $\eta^2_p = 0.04$, on accuracy in attend voice trials. Post-hoc comparisons revealed similar accuracy for both groups in unimodal, $t(70) = -1.0$, $p = .321$, Cohen's $d = -0.24$, congruent, $t(70) = 0.74$, $p = .465$, Cohen's $d = 0.17$, and incongruent happy voice trials, $t(70) = 1.13$, $p = .264$, Cohen's $d = 0.27$. The high proneness group showed both a congruency facilitation, $t(35) = 7.61$, $p < .001$, Cohen's $d = 1.27$, and incongruency interference effect, $t(35) = -5.49$, $p < .001$, Cohen's $d = 0.92$, on accuracy in happy voice trials. Both a congruency facilitation, $t(35) = 4.85$, $p < .001$, Cohen's $d = 0.81$, and an incongruency interference effect, $t(35) = -9.57$, $p < .001$, Cohen's $d = -1.59$, on accuracy in happy voice trials could also be observed in the low proneness group. There was no main effect of *Group*, $F(1, 70) = 0.15$, $p = .703$, $\eta^2_p = 0.002$, and no *Group* * *Stimulus Condition* interaction, $F(2, 140) = 0.2$, $p = .819$, $\eta^2_p = 0.003$, on

mean perceived intensity in attend voice trials. The main effect of *Stimulus Condition* on mean perceived intensity in attend voice trials was highly significant, $F(2, 140) = 54.95$, $p < .001$, $\eta^2_p = 0.44$, revealing both a congruency facilitation effect, $t(71) = 5.54$, $p < .001$, Cohen's $d = 0.65$, and an incongruency interference effect, $t(71) = -5.62$, $p < .001$, Cohen's $d = -.066$, when rating happy face stimuli.

In sum, both groups overall showed similar mean RT, accuracy and mean intensity ratings in both attend face and voice trials with target emotion happy. Overall, congruency facilitation and/or incongruency interference effects could be found for all measures in trials with target emotion happy in both attention conditions except for mean RT in attend face trials. In attend face trials, the high proneness group showed both a congruency facilitation and an incongruency interference effect on mean RT, whereas the low proneness group showed neither effect. Apart from this, both groups showed a comparable pattern of in-/congruency effects for all measures in happy face and voice trials (see Figure 4 top row for mean RT, accuracy and mean perceived intensity for target emotion happy).

**Target Emotion Angry.** Separate *Group* * *Stimulus Condition* mixed ANOVAs in attend face trials with target emotion angry revealed neither a main effect of *Group*, $F(1, 70) = 1.7$, $p = .196$, $\eta^2_p = 0.02$, nor of *Stimulus Condition*, $F(2, 140) = 2.30$, $p = .104$, $\eta^2_p = 0.03$, and no *Group* * *Stimulus Condition* interaction, $F(2, 140) = 0.77$, $p = .464$, $\eta^2_p = 0.01$, on mean RT. There was no main effect of *Group* on accuracy in attend face trials, $F(1, 70) = 1.26$, $p = .265$, $\eta^2_p = 0.02$, and no *Group* * *Stimulus Condition* interaction, $F(1.58, 110.73) = 0.95$, $p = .372$, $\eta^2_p = 0.01$ (note: there was no homogeneity of covariances in accuracy in angry face trials, Box-test $p = .001$, prohibiting an interpretation of the interaction). The main effect of *Stimulus Condition* on accuracy in attend face trials was highly significant, $F(1.58, 110.73) = 14.52$, $p < .001$, $\eta^2_p = 0.17$, indicating an incongruency interference effect on accuracy in angry face stimuli, $t(71) = -4$, $p < .001$, Cohen's $d = -0.47$. There was no main effect of *Group*, $F(1, 70) = 0.64$, $p = .425$, $\eta^2_p = 0.009$, and no *Group* * *Stimulus Condition* interaction, $F(1.53, 107.28) = 0.002$, $p = .994$, $\eta^2_p < 0.001$, on mean perceived intensity in attend face trials. The main effect of *Stimulus Condition* on mean perceived intensity in attend face trials was highly significant, $F(1.53, 107.28) = 14.47$, $p < .001$, $\eta^2_p = 0.17$, indicating an incongruency interference effect on intensity ratings in angry face stimuli, $t(71) = -3.30$, $p = .005$, Cohen's $d = -0.39$.

Separate *Group* * *Stimulus Condition* mixed ANOVAs in attend voice trials with target emotion angry revealed no main effect of *Group* on mean RT, $F(1, 70) = 0.09$, $p = .769$, $\eta^2_p = 0.001$. There was a highly significant main effect of *Stimulus Condition*, $F(2, 140) = 20.76$, $p < .001$, $\eta^2_p = 0.23$, and a significant *Group* * *Stimulus Condition* interaction, $F(2, 140) = 3.24$, $p = .042$, $\eta^2_p = 0.04$, on mean RT in attend voice trials. Post-hoc comparisons revealed similar mean RT for both groups in unimodal, $t(70) = 0.70$, $p = .487$, Cohen's $d = 0.17$, congruent, $t(70) = -0.6$, $p = .549$, Cohen's $d = -0.14$, and incongruent angry voice trials, $t(70) = 0.78$, $p = .438$, Cohen's $d = 0.18$. The high proneness group showed an incongruency interference, $t(35) = 4.21$, $p < .001$, Cohen's $d = 0.7$, but no congruency facilitation effect, $t(35) = -1.02$, $p = .314$, Cohen's $d = -0.17$, on mean RT in angry voice trials, whereas the low proneness group showed an incongruency interference effect, $t(35) = 4.47$, $p < .001$, Cohen's $d = 0.74$, and a non-significant trend for higher mean RT in congruent compared to unimodal angry voice trials, $t(35) = 2.03$, $p = .051$, Cohen's $d = 0.34$. A mixed ANOVA on accuracy in attend voice trials could not be interpreted due to violations of homogeneity of error variances (Levene-test $p < .05$). Following a recommendation by Hsu (1996), post-hoc comparisons are reported. Post-hoc comparisons revealed lower accuracy in the high proneness ($M = 78.88$, $SD = 10.72$) compared to the low proneness group ($M = 84.51$, $SD = 6.93$) during categorization of angry voice stimuli, $t(59.89) = -2.64$, $p = .010$, Cohen's $d = -0.62$. A congruency facilitation effect on accuracy could be observed in angry voice stimuli, $t(71) = 17.02$, $p < .001$, Cohen's $d = 2.01$. Further, there was no main effect of *Group*, $F(1, 70) = 0.86$, $p = .356$, $\eta^2_p = 0.01$, and no *Group* * *Stimulus Condition* interaction, $F(2, 140) = 0.39$, $p = .68$, $\eta^2_p = 0.005$, on mean perceived intensity in attend voice trials. The main effect of *Stimulus Condition* on mean perceived intensity in attend voice trials was highly significant, $F(2, 140) = 39.32$, $p < .001$, $\eta^2_p = 0.36$, indicating both a congruency facilitation effect, $t(71) = 3.42$, $p = .003$, Cohen's $d = 0.4$, and an incongruency interference effect, $t(71) = -5.44$, $p < .001$, Cohen's $d = -0.64$, on intensity ratings in angry voice stimuli.

In sum, both groups overall showed similar mean RT, accuracy and mean intensity ratings in both attend face and voice trials with target emotion angry, with the exception of accuracy scores in attend voice trials, with the high proneness group being less accurate compared to the low proneness group. Congruency facilitation and/or incongruency interference effects could be observed for all measures in trials with target emotion angry in both attention conditions, except for mean RT in attend face trials. Overall, both groups

showed a comparable pattern of in-/congruency effects for all measures in angry face and voice trials. (see Figure 4 second row for mean RT, accuracy and mean perceived intensity for target emotion angry).

**Target Emotion Sad.** Separate *Group* * *Stimulus Condition* mixed ANOVAs in attend face trials with target emotion sad revealed neither a main effect of *Group*, $F(1, 70) = 1.70$, $p = .197$, $\eta^2_p = 0.02$, nor of *Stimulus Condition,* $F(2, 140) = 1.77$, $p = .175$, $\eta^2_p = 0.03$, and no *Group* * *Stimulus Condition* interaction, $F(2, 140) = 1.12$, $p = .329$, $\eta^2_p = 0.02$, on mean RT. A mixed ANOVA on accuracy in attend face trials was not interpretable due to violations of homogeneity of error variances (Levene-test $p < .05$). Post-hoc comparisons revealed a non-significant trend for lower accuracy in the high proneness ($M = 77.91$, $SD = 10.85$) compared to the low proneness group ($M = 82.25$, $SD = 9.79$) during categorization of sad face stimuli, $t(70) = -1.78$, $p = .079$, Cohen's $d = -0.42$. Both a congruency facilitation, $t(71) = 4.4$, $p < .001$, Cohen's $d = 0.52$, and a incongruency interference effect, $t(71) = -6.5$, $p < .001$, Cohen's $d = -0.77$, could be observed on accuracy in sad face stimuli. Further, there was no main effect of *Group*, $F(1, 70) = 0.01$, $p = .93$, $\eta^2_p = 0.001$, and no Group * *Stimulus Condition* interaction, $F(1.78, 124.37) = 1.61$, $p = .206$, $\eta^2_p = 0.02$, on mean perceived intensity in attend face trials. The main effect of *Stimulus Condition* on mean perceived intensity in attend face trials was highly significant, $F(1.78, 124.37) = 21.03$, $p < .001$, $\eta^2_p = 0.23$, indicating both a congruency facilitation effect, $t(71) = 3.51$, $p = .003$, Cohen's $d = 0.41$, and an incongruency interference effect, $t(71) = -3.6$, $p = .002$, Cohen's $d = -0.43$, on intensity ratings in sad face stimuli.

Separate *Group* * *Stimulus Condition* mixed ANOVAs in attend voice trials with target emotion sad revealed no main effect of *Group* on mean RT, $F(1, 70) = 0.21$, $p = .652$, $\eta^2_p = 0.003$, and no *Group* * *Stimulus Condition* interaction, $F(2, 140) = 0.42$, $p = .959$, $\eta^2_p = 0.001$ (note: there was no homogeneity of covariances in mean RT in sad voice trials, Box-test $p = .001$, not allowing an interpretation of the interaction). There was a highly significant main effect of *Stimulus Condition* on mean RT in attend voice trials, $F(2, 140) = 17.16$, $p < .001$, $\eta^2_p = 0.20$, indicating an incongruency interference effect on mean RT in sad voice stimuli, $t(71) = 5.83$, $p < .001$, Cohen's $d = 0.69$. A mixed ANOVA on accuracy in attend voice trials was not interpretable due to violations of homogeneity of error variances (Levene-test $p < .05$). Post-hoc comparisons revealed a non-significant trend for lower accuracy in the high proneness ($M = 91.60$, $SD = 7.31$) compared to the low proneness group ($M = 94.40$, *SD*

= 5.66) during categorization of sad voice stimuli, $t(70) = -1.82$, $p = .073$, Cohen's $d = -.043$. An incongruency interference effect on accuracy could be observed in sad voice stimuli, $t(71) = -6.21$, $p < .001$, Cohen's $d = -0.73$. Further, there was no main effect of *Group,* $F(1, 70) = 0.14$, $p = .705$, $\eta^2_p = 0.002$, and no *Group* * *Stimulus Condition* interaction, $F(1.83, 128.3) = 0.15$, $p = .845$, $\eta^2_p = 0.002$, on mean perceived intensity in attend voice trials. The main effect of *Stimulus Condition* on mean perceived intensity in attend voice trials was highly significant, $F(1.83, 128.3) = 40.37$, $p < .001$, $\eta^2_p = 0.37$, indicating an incongruency interference effect on intensity ratings in sad voice stimuli, $t(71) = -8.71$, $p < .001$, Cohen's $d = -1.03$, as well as lower intensity ratings in congruent compared to unimodal sad voice trials, $t(71) = -3.95$, $p < .001$, Cohen's $d = -0.47$.

In sum, both groups showed similar mean RT and mean intensity ratings in trials with target emotion sad in both attention conditions. However, the high proneness group showed a non-significant trend for lower accuracy compared to the low proneness group in trials with target emotion sad in both the attend face and attend voice condition. Congruency facilitation and/or incongruency interference effects could be observed for all measures in trials with target emotion sad in both attention conditions except for mean RT in attend face trials. Both groups showed a similar pattern of in-/congruency effects for all measures in sad face and voice trials (see Figure 4 third row for mean RT, accuracy and mean perceived intensity for target emotion sad).

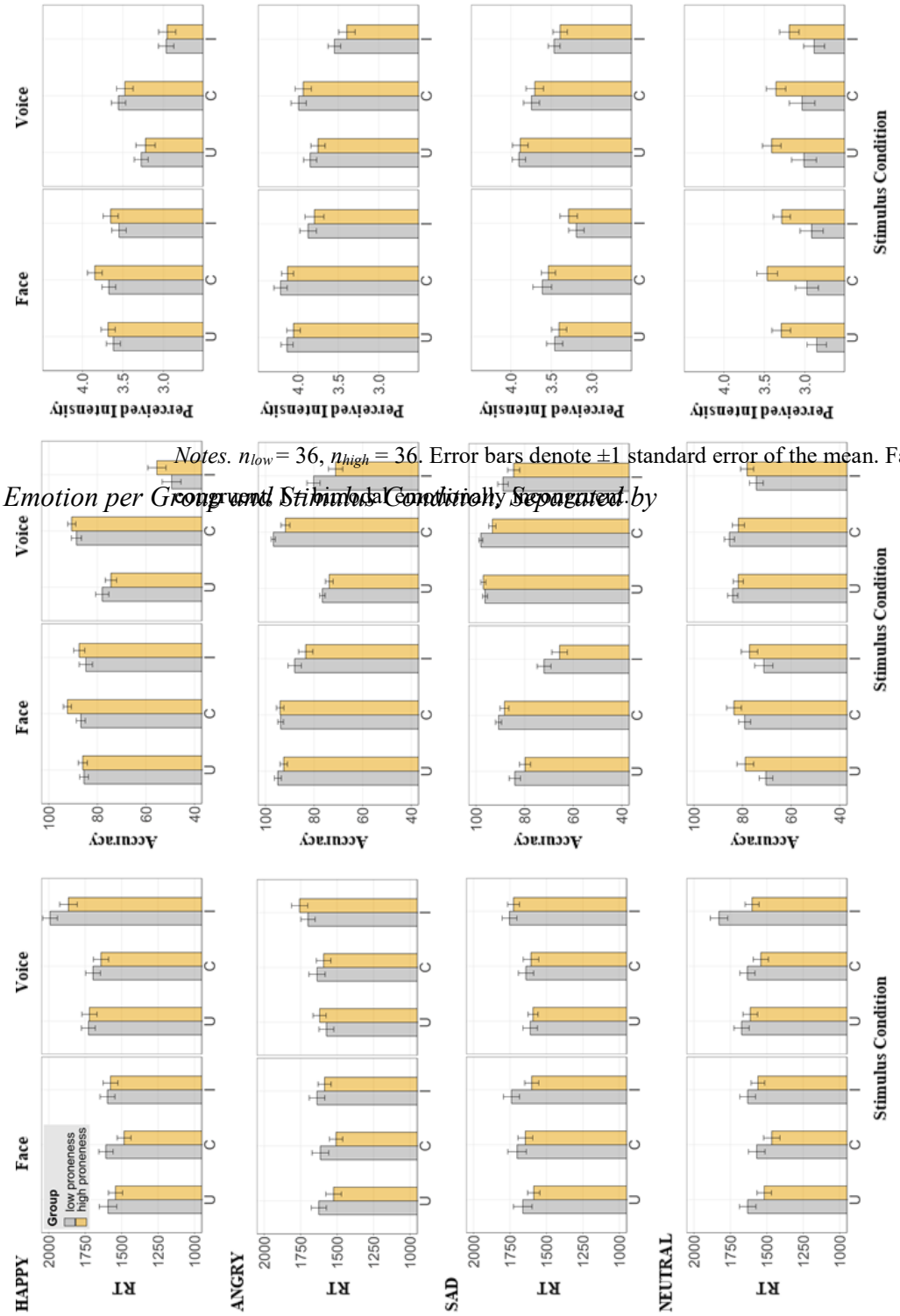
**Target Emotion Neutral.** Separate *Group* * *Stimulus Condition* mixed ANOVAs in attend face trials with target emotion neutral revealed a non-significant trend for a main effect of *Group* on mean RT, $F(1, 70) = 3.15$, $p = .081$, $\eta^2_p = 0.04$, indicating a non-significant trend for lower mean RT in the high proneness ($M = 1494.04$, $SD = 240.02$) compared to the low proneness group ($M = 1613.13$, $SD = 316.63$) when categorizing emotionally neutral face stimuli. There was a highly significant main effect of *Stimulus Condition* on mean RT in attend face trials, $F(2, 140) = 8.97$, $p < .001$, $\eta^2_p = 0.12$, indicating a congruency facilitation effect on mean RT in emotionally neutral face stimuli, $t(71) = -2.49$, $p = .024$, Cohen's $d = -0.29$. There was no *Group* * *Stimulus Condition* interaction on mean RT in attend face trials, $F(2, 140) = 1.79$, $p = .17$, $\eta^2_p = 0.03$. Further, there was no main effect of *Group*, $F(1, 70) = 2.59$, $p = .112$, $\eta^2_p = 0.04$, and no *Group* * *Stimulus Condition* interaction, $F(1.79, 125.2) = 0.59$, $p = .539$, $\eta^2_p = 0.008$, on accuracy in attend voice face trials. The main effect of *Stimulus Condition* on accuracy in attend face trials was highly significant, $F(1.79, 125.2) = 8.64$, $p <$

.001, $\eta^2_p$ = 0.11, indicating a congruency facilitation effect on accuracy in emotionally neutral face stimuli, $t(71) = 4.06$, $p < .001$, Cohen's $d = 0.48$. There was a significant main effect of *Group* on mean perceived intensity in attend face trials, $F(1, 70) = 6.65$, $p = .012$, $\eta^2_p = 0.09$, indicating significantly higher intensity ratings of emotionally neutral faces in the high proneness ($M = 3.35$, $SD = 0.64$) compared to the low proneness group ($M = 2.91$, $SD = 0.77$). The main effect of *Stimulus Condition* on mean perceived intensity was significant, $F(1.84, 128.78) = 4.48$, $p = .016$, $\eta^2_p = 0.06$, indicating a congruency facilitation effect on intensity ratings in emotionally neutral face stimuli, $t(71) = 3.42$, $p = .003$, Cohen's $d = 0.4$. There was no *Group* * *Stimulus Condition* interaction on mean perceived intensity in attend face trials, $F(1.84, 128.78) = 0.65$, $p = .512$, $\eta^2_p = 0.009$.

Separate *Group* * *Stimulus Condition* mixed ANOVAs in attend voice trials with target emotion neutral revealed a non-significant trend for a main effect of *Group* on mean RT, $F(1, 70) = 3.50$, $p = .066$, $\eta^2_p = 0.05$, indicating a non-significant trend for lower mean RT in the high proneness ($M = 1585.00$, $SD = 281.4$) compared to the low proneness group ($M = 1711.25$, $SD = 291.15$) during categorization of emotionally neutral voice stimuli. There was a highly significant main effect of *Stimulus Condition* on mean RT in attend voice trials, $F(2, 140) = 11.84$, $p < .001$, $\eta^2_p = 0.15$. *Group* and *Stimulus Condition* interacted significantly with each other on mean RT, $F(2, 140) = 5.70$, $p = .004$, $\eta^2_p = 0.08$. Post-hoc comparisons revealed lower mean RT in the high proneness compared to the low proneness group in incongruent, $t(70) = -2.98$, $p = .004$, Cohen's $d = -0.7$, but not in unimodal, $t(70) = -0.87$, $p = .414$, Cohen's $d = -0.19$, or congruent neutral voice trials, $t(70) = -1.26$, $p = .212$, Cohen's $d = 0.30$. The high proneness group showed a congruency facilitation, $t(35) = -2.28$, $p = .029$, Cohen's $d = -0.38$, but no incongruency interference effect, $t(35) = -0.41$, $p = .681$, Cohen's $d = -0.07$, on mean RT in neutral voice trials, whereas the low proneness group showed an incongruency interference, $t(35) = 4.2$, $p < .001$, Cohen's $d = 0.7$, but no congruency facilitation effect, $t(35) = -0.9$, $p = .374$, Cohen's $d = -0.15$. There was no main effect of *Group* on accuracy in attend voice trials, $F(1, 70) = 0.07$, $p = .797$, $\eta^2_p = 0.001$. There was a highly significant main effect of *Stimulus Condition*, $F(2, 140) = 11.86$, $p < .001$, $\eta^2_p = 0.15$, and a non-significant trend for a *Group* * *Stimulus Condition* interaction, $F(2, 140) = 2.83$, $p = .062$, $\eta^2_p = 0.04$, on accuracy in attend voice trials. Post-hoc comparisons revealed similar accuracy for both groups in unimodal, $t(70) = -0.77$, $p = .442$, Cohen's $d = -0.18$, congruent, $t(70) = -1.13$, $p = .262$, Cohen's $d = -0.27$, and incongruent neutral voice trials, $t(70) = 0.97$, $p$

= .335, Cohen's $d$ = 0.23. The high proneness group showed neither a congruency facilitation, $t(35)$ = -0.02, $p$ = .982, Cohen's $d$ = -.004, nor an incongruency interference effect, $t(35)$ = -1.36, $p$ = .183, Cohen's $d$ = -0.23, on accuracy in neutral voice trails, whereas the low proneness group showed an incongruency interference, $t(35)$ = -4.21, $p$ < .001, Cohen's $d$ = -0.7, but no congruency facilitation effect, $t(35)$ = 0.61, $p$ = .549, Cohen's $d$ = 0.1. Due to violations of homogeneity of error variances (Levene-test p < .05), a mixed ANOVA on mean perceived intensity was not interpretable. Post-hoc comparisons revealed a non-significant trend for higher intensity ratings of emotionally neutral voice stimuli in the high proneness ($M$ = 3.32, $SD$ = 0.67) compared to the low-proneness group ($M$ = 2.98, $SD$ = 0.85), $t(70)$ = 1.9, $p$ = .061, Cohen's $d$ = 0.45. Further, an incongruency interference effect on mean perceived intensity could be observed in emotionally neutral voice stimuli, $t(71)$ = -3.14, $p$ = .008, Cohen's $d$ = -0.37.

In sum, the high proneness group showed a non-significant trend for lower mean RT in trials with target emotion neutral in both the attend face and attend voice condition as well as higher intensity ratings in attend face and a non-significant trend for higher intensity ratings in attend voice trials with target emotion neutral. Overall, accuracy scores were similar in both groups in trials with target emotion neutral in both attention conditions. Congruency facilitation and/or incongruency interference effects could be observed for all measures in trials with target emotion neutral in both attention conditions. The low proneness group showed a marked incongruency interference effect on mean RT in neutral voice trials, whereas the high proneness group did not show this effect. Further, the high proneness group did not show any in-/congruency effects on accuracy in neutral voice trials, in contrast to the low proneness group (see Figure 4 bottom row for mean RT, accuracy and mean perceived intensity for target emotion neutral).




*Notes.* $n_{low}$ = 36, $n_{high}$ = 36. Error bars denote ±1 standard error of the mean. Face = attend face trial

*l Intensity for each Target Emotion per Group and Stimulus Condition, Separated by*

*roup and Stimulus Condition, Separated by*

## 3.2 Control Variables and Questionnaire data

### *3.2.1 Control Tests for Attention, Processing Speed, and Vigilance*

**TMT.** Mean completion time (in seconds) in the TMT-A version was similar in the high proneness ($M$ = 30.1, $SD$ = 10.9) and the low proneness group ($M$ = 27.0, $SD$ = 8.0), $t(70) = -1.36$, $p = .178$, Cohen's $d = 0.32$. Further, the high proneness ($M$ = 54.1, $SD$ = 13.7) and low proneness group ($M$ = 54.9, $SD$ = 16.5) showed similar mean completion time in the TMT-B version, $t(70) = 0.22$, $p = .824$, Cohen's $d = -0.05$.

**CPT-IP.** The 2 (*Group*) x 3 (*Digit Load*) mixed ANOVA for CPT-d'-scores revealed a significant main effect of *Digit Load*, $F(2, 140) = 269.53$, $p < .001$, $\eta^2_p = 0.79$, indicating higher d'-values in the 2-digit load condition ($M$ = 3.80, $SD$ = 0.47) compared to both the 3-digit load condition ($M$ = 2.84, $SD$ = 0.84), $t(71) = 10.26$, $p < .001$, and the 4-digit load condition ($M$ = 1.63, $SD$ = 0.73), $t(71) = 23.45$, $p < .0001$, as well as higher d'-values in the 3-digit load condition compared to the 4-digit load condition, $t(71) = 12.86$, $p < .001$. There was no main effect of *Group*, $F(1, 70) = 0.36$, $p = .55$, $\eta^2_p = 0.005$, and no *Group* * *Digit Load* interaction, $F(2, 140) = 0.48$, $p = .618$, $\eta^2_p = 0.007$.

### *3.2.2 Questionnaire Data*

**CAPS.** Groups significantly differed in CAPS total scores, with the high proneness group showing higher scores ($M$ = 11.3, $SD$ = 6.4) compared to the low proneness group ($M$ = 5.7, $SD$ = 4.3), $t(61.52) = 4.31$, $p < .001$, Cohen's $d = 1.02$.

**LSHS.** Groups significantly differed in LSHS-E total scores, with the high proneness group showing higher scores ($M$ = 15.3, $SD$ = 9.7) compared to the low proneness group ($M$ = 8.6, $SD$ = 7.1), $t(70) = 3.39$, $p = .001$, Cohen's $d = 0.80$.

**PCL**[3]. Groups did not significantly differ in PCL total scores, with the high proneness group showing a non-significant trend towards higher scores (M = 47.3, SD = 27.1) compared to the low proneness group (M = 34.8, SD = 26.3), t(70) = 1.98, p = .052, Cohen's d = 0.47.

**BCSS.** Groups did neither differ in BCSS negative self scores (high proneness: $M$ = 4.2, $SD$ = 3.1; low proneness: $M$ = 3.3, $SD$ = 2.7), $t(70) = 1.34$, $p = .185$, Cohen's $d = 0.32$,

[3] One subject in the low and one in the high proneness group had ≥50% missings in the PCL. Their PCL scores were corrected by the respective group mean.

nor in BCSS negative other scores (high proneness: $M = 4.8$, $SD = 3.3$; low proneness: $M = 3.8$, $SD = 3.9$), $t(68.5) = 1.17$, $p = .246$, Cohen's $d = 0.28$.

**Correlations Between Questionnaire Scores.** Bonferroni-corrected Pearson correlation coefficients over all subjects revealed highly significant moderate to strong correlations between CAPS and LSHS, between CAPS and PCL, between LSHS and PCL, and between PCL and both BCSS negative subscales, all $r > .4$ and $p < .001$. Further, there was a significant correlation between LSHS and BCSS negative other subscale, $r = .40$, $p < .05$. No other correlation between questionnaires scores reached statistical significance, all $p > .05$ (see Table 2 for Pearson's $r$ for correlations of questionnaire data).

## 3.3 Correlations Between Experimental Data and Questionnaire Data

### *3.3.1 Correlations Between IE Difference Scores and Questionnaire Data*

Pearson correlation coefficients over all subject indicated small negative correlations between the IE difference score congruent-unimodal in attend face trials and the CAPS ($r = -.24$, uncorrected $p = .046$) as well as between the IE difference score congruent-incongruent in attend voice trials and the BCSS negative other scale ($r = -.23$, uncorrected $p = .049$). Neither did survive Bonferroni-correction. No other correlation between IE difference scores and questionnaire data reached statistical significance, all $r < .3$ and $p > .05$ (see Table 2 for Pearson's $r$ for correlations between IE difference scores and questionnaire data).

### *3.3.2 Correlations Between IE Scores and Questionnaire Data*

Pearson correlation coefficients over all subjects indicated negative correlations between IE scores in congruent trials during attend face blocks and both the CAPS ($r = -.33$, uncorrected $p = .005$) and LSHS ($r = -.24$, uncorrected $p = .043$), respectively. Neither did survive Bonferroni-correction, Bonferroni-corrected $p = .15$ and $p = 1$. No other correlation between IE scored and questionnaire data reached statistical significance, all $r < .3$ and $p > .05$ (see Table B.1 in Appendix B for Pearson's $r$ for correlations between IE scores and questionnaire data).

# 4 Discussion

In this study, we investigated potential impairments in emotional MSI in stages prior to psychosis onset. Following up on findings of altered emotional MSI in patients with the diagnosis of a psychotic disorder (Lin et al., 2020; Tseng et al., 2015), we examined emotional MSI in psychosis proneness in order to generate hypotheses regarding a potential

role of impaired emotional MSI for disorder development. Two groups of a healthy community sample, one with low and the other with high psychosis proneness, categorized the perceived emotion and rated the perceived emotional intensity of unimodal, bimodal emotionally congruent and bimodal emotionally incongruent dynamic facial expressions and vocal prosody. Over the course of 12 experimental blocks, the attended, i.e. task-relevant, modality alternated, while the respective other modality had to be ignored (Föcker et al., 2011). Further, participants completed neuropsychological tests for processing speed, attention, and vigilance as well as self-report questionnaires for psychotic experiences and beliefs about the self and others.

In both attention conditions, both groups showed comparable patterns of emotion categorization performance and intensity ratings depending on stimulus congruency. This indicated similar emotional MSI in our high and low psychosis proneness samples. These results will be discussed in more detail in the following with regard to our hypotheses and previous findings on emotional processing in psychosis and psychosis proneness.

## 4.1 Hypothesis 1: Crossmodal Modulation in Emotional MSI and Psychosis Proneness

### *4.1.1 Discussion and Comparison to Previous Evidence*

We expected our high proneness sample to show aberrant crossmodal congruency facilitation and incongruency interference effects during emotion categorization, which would indicate atypical emotional MSI. In contrast to our expectation, our behavioral data did not reveal an interaction between group and stimulus condition on categorization performance: In both attention conditions, both groups showed lower IE scores (i.e. better performance) in emotionally congruent stimuli and higher IE scores (i.e. worse performance) in emotionally incongruent stimuli compared to unimodal stimuli, respectively. This indicates similar congruency facilitation and incongruency interference effects in categorization performance for both groups. Additionally, our data did not reveal a main effect of group on categorization performance, indicating overall functional categorization of emotionally laden face and voice stimuli in both groups. In sum, our findings suggest typically developed emotional MSI in our psychosis proneness sample.

The analyses for mean perceived intensity ratings mirrored the findings of the IE-analyses: we observed similar congruency facilitation and incongruency interference on intensity ratings in both groups. Additionally, high and low psychosis proneness subjects

showed similar overall levels of mean perceived emotional intensity in both attention conditions. The supplementary analyses for mean RT and accuracy provided analogue results: overall, congruency and incongruency effects could be observed in terms of lower RT and better accuracy in congruent trials as well as higher RT and lower accuracy in incongruent trials compared to unimodal trials. In sum, we observed crossmodal modulation on intensity ratings, RT, and accuracy for both the low and the high proneness group during processing of emotionally laden AV stimuli, analogue to the results regarding IE.

Taken together, we were able to observe previously published crossmodal modulation effects in emotional MSI (e.g. Collignon et al., 2008; Föcker et al., 2011; Föcker & Röder, 2019), successfully replicating the findings in subjects from the general population originally published by Föcker et al., (2011) and in a further adoption of the original paradigm (Fengler et al., 2017). We observed congruency facilitation and incongruency interference effects both in emotion categorization and emotional intensity ratings in both attentions conditions, despite the explicit instruction to ignore the task-irrelevant modality. This indicated that the task-irrelevant, non-attended information, i.e. the facial or vocal cue, modulated the processing of emotionally laden face expressions and vocal prosody, respectively. Thus, our findings are consistent with previous work on crossmodal modulation in emotional MSI (e.g. Collignon et al., 2008) and suggest that AV emotional MSI might constitute a bi-directional, automatic process relatively independent of attentional focus (Fengler et al., 2017; Föcker et al., 2011; Föcker & Röder, 2019).

Our results suggest typically developed emotional MSI in our psychosis proneness sample, which differs from the majority of findings in patients with the diagnosis of a psychotic disorder. Thus, our findings indicate a different pattern of emotional MSI in psychosis proneness compared to studies in patients with the diagnosis of a psychotic disorder. Previous studies had reported reduced crossmodal modulation by non-target modality cues (de Gelder et al., 2005, Exp.1; de Jong et al., 2009; Fiszdon & Bell, 2009; Seubert, 2010), reduced AV emotion matching performance (Castagna et al., 2013; Jeong et al., 2021), or lower emotion categorization performance in AV trials (Giannitelli et al., 2015; Mangelinckx et al., 2017; Thaler et al., 2013) in patients compared to healthy controls. Further, Müller et al. (2014) and Portnova et al., (2021) reported altered ERP responses during processing of AV emotional stimuli in patients compared to healthy controls,

especially when visual and auditory cues were incongruent, possibly reflecting deficient neural responses to conflicting AV emotional information in psychosis.

Further, our findings in psychosis proneness also show a different pattern compared to patient studies reporting excessive emotional MSI: both de Gelder et al. (2005, Exp.2) and Van den Stock et al. (2011) reported an increased crossmodal modulation by the non-target modality on AV emotion categorization in patients compared to controls. However, the subjects in the study by Van den Stock et al. (2011) had to categorize the perceived emotion in body expressions rather than facial expressions. It is possible that the specific stimulus characteristics in facial vs. body expression recognition and potential discrepancies in the respective task demand differently impact AV emotion recognition in psychosis. Strauss et al. (2015) had indeed reported deficits in recognizing emotional body expression in patients with the diagnosis of a psychotic disorder compared to healthy controls. To our knowledge, however, no published study directly compared deficits in emotional face and body expressions in psychosis.

On the other hand, our findings in psychosis proneness are in line with findings of studies reporting similar AV emotion categorization performance (Müller et al., 2014; Simpson et al., 2013; Zvyagintsev et al., 2013) and similar AV emotion valence ratings (Sestito et al., 2013) in patients with the diagnosis of a psychotic disorder compared to healthy controls, reflecting unimpaired behavioral performance during multisensory processing of emotional face and voice stimuli.

Taken together, our findings might suggest that emotional MSI in psychosis proneness remains mostly intact prior to disorder onset, with aberrances potentially manifesting later in full-blown psychosis. Unlike multisensory emotion processing reported here, previous studies have reported impairments in unisensory emotion processing in high psychosis proneness (Seo et al., 2020; van Donkersgoed et al., 2015). It is possible that multisensory processes ameliorate these unisensory emotion processing deficits, resulting in similar AV emotion categorization in subjects with high and low proneness. Crossmodal compensatory effects during processing of congruent crossmodal compared to unimodal emotional information, i.e. at least partially intact crossmodal gains, have been reported in patients with the diagnosis of a psychotic disorder (Mangelinckx et al., 2017; Müller et al., 2014; Simpson et al., 2013). It could be speculated that psychosis proneness might be associated with relatively intact

crossmodal gain effects, which reduce with increased symptom load. Future research could aim to disentangle whether a) psychosis proneness is associated with preserved crossmodal gains ameliorating unisensory emotion processing deficits, b) if crossmodal gains change over the course of disorder progression and c) which factors influence these changes. It is important to note, however, that the lack of unisensory emotion processing differences between our high and low proneness groups is inconsistent with reports of reduced facial and vocal emotion recognition and discrimination in high psychosis proneness (Seo et al., 2020; van Donkersgoed et al., 2015) and thus does not offer support for our speculation of preserved crossmodal gains ameliorating unisensory emotion processing deficits in psychosis proneness.

### *4.1.2 A Potential Role of Disorder Progress Stage*

Our sample mainly consisted of university students, who likely showed high levels of emotional and social functioning and thus potentially preserved emotion processing capabilities. Previous studies on unisensory emotion processing focused on individuals with ultra-high risk for psychosis (Seo et al., 2020; van Donkersgoed et al., 2015). Thus, it is possible that our sample did not exhibit similar levels of symptom load as well as perceptual-cognitive impairments as the subjects in previous studies. The lack of group differences in processing speed, attention and vigilance in our data, which have been previously reported to be impaired in psychosis proneness (Mesholam-Gately et al., 2009; Nuechterlein et al., 2004), might support this assumption. However, our sample was allocated to the high proneness group by means of a cut-off score in the CAPE (Stefanis et al., 2002), which has been previously shown to distinguish between low and high proneness subjects (Krkovic et al., 2020). Further, our high proneness group reported significantly higher levels of anomalous perceptual experiences and psychotic symptoms compared to the low proneness group. It is possible that both unisensory and multisensory emotion processing impairments manifest rather in ultra-high risk states and full-blown psychosis than in psychosis proneness. This is supported by studies reporting that age, illness duration and symptom severity influence emotion processing deficits in psychosis (Edwards et al., 2002; Kohler et al., 2010) and by studies reporting associations between impaired face and voice processing and higher levels of perceptual-cognitive dysfunctions and psychotic symptoms (Darke et al., 2021; Lin et al., 2018). On the other hand, unisensory emotion processing deficits in terms of impaired facial affect recognition were already observed in subjects with increased genetic risk for psychosis such as first-degree relatives of patients with the diagnosis of a psychotic disorder (D. Martin

et al., 2020), suggesting early impairments and a potential vulnerability marker for psychosis (Seo et al., 2020). Further research is necessary to disentangle the development of uni- and multisensory emotion processing deficits in psychosis proneness, ultra-high risk groups and psychosis.

### *4.1.3 Attention Effects*

In our exploratory analyses with attention condition as an additional factor, we observed an interaction between group and attention condition, reflecting that the difference in emotion categorization performance between attend face and attend voice trials was larger in high proneness compared to low proneness subjects. This could suggest larger sensitivity to emotional face expressions compared to vocal prosody in our high proneness subjects. While we observed a group * attention condition interaction on IE scores, there was no such interaction on mean perceived emotional intensity ratings. This suggests that a potentially higher sensitivity to emotional face expressions compared to vocal prosody in high proneness subjects might influence crossmodal emotion categorization but not the perception of emotional intensity. The impact of visual vs. auditory modality on emotion recognition in psychosis has not been sufficiently studied and experimental designs and results of previous studies appeared to be too inconsistent for clear conclusions (Dondaine et al., 2014). Further, to our knowledge no published study compared visual and auditory emotion recognition in psychosis proneness. A study by Dondaine et al. (2014) aimed to compare visual and auditory emotion recognition impairments in patients with the diagnosis of a psychotic disorder, but did not observe an impact of modality on emotion recognition. However, Thaler et al. (2013) observed different patterns of emotion misclassification in psychosis for affectively neutral stimuli depending on presentation modality, i.e. visual, auditory or AV. More research comparing visual and auditory emotion recognition deficits in psychosis and psychosis proneness is needed to further investigate a potential role of task-relevant modality.

Further, we observed a significant interaction between attention and stimulus condition in our exploratory 3-way analysis, indicating better categorization performance in attend face compared to attend voice trials in the unimodal and the emotionally incongruent condition. This suggests better face expression than vocal prosody recognition and less incongruency interference of auditory cues when attending face compared to voice stimuli. The lack of performance differences between attend face and attend voice trials in the emotionally congruent condition could reflect greater crossmodal gain of face stimuli when

attending vocal prosody, resulting in similar performance as when attending face stimuli with emotionally congruent vocal prosody. These findings might reflect visual dominance in emotion recognition of dynamic face and voice stimuli and is consistent with previous studies reporting a higher influence of emotional face stimuli compared to vocal stimuli when categorizing emotions (e.g. Collignon et al., 2008).

### *4.1.4 Effects of Emotion Category and Valence*

Our exploratory analysis of RT, accuracy and mean perceived intensity separately for each of the four emotion categories revealed hints regarding a more fine-grained pattern of AV emotion processing in our proneness sample. Our high proneness sample was less accurate in categorizing angry voice stimuli compared to the low proneness group and showed non-significant trends for less accurate categorizations of both sad face and voice stimuli. In contrast, both groups showed similar categorization performance in happy face and voice trials. Our findings are consistent with previous studies on emotion recognition deficits in psychosis reporting pronounced deficits in categorizing negative emotions compared to positive emotions (Bonfils et al., 2020; Tripoli et al., 2022; Won et al., 2019). Negative emotion recognition deficits have been observed in early disorder stages (Won et al., 2019) and it has been discussed that specific impairments in the recognition of negative emotions such as anger might be associated with a higher vulnerability for psychosis (Tripoli et al., 2022). A meta-analysis on emotion perception and functional outcomes in psychosis observed an association between deficient emotion recognition and poor social skills such as social problem solving (Irani et al., 2012). This could suggest that poor emotion recognition might facilitate stressful social encounters and in turn exacerbate psychotic symptoms. However, the meta-analysis did not distinguish between emotions of positive vs. negative valence. It could be speculated that a pronounced deficit in recognizing negative emotions might be particular impactful, since it can be assumed that social situations in which others express negative emotions might be more stressful. Dysfunctional emotion recognition and a deficit in social skills might lead to a particular stressful social encounter for the individual suffering from psychosis, with detrimental consequences for their social behaviour and experience of psychotic symptoms. Initial support for our speculation of a pronounced impact of deficient negative emotion recognition can be found in a study by Modinos et al. (2020) in subjects with high risk for psychosis, who observed an association between deficient anger and fear recognition and poorer functional outcomes in terms of larger impairments in daily life. More

studies on a potentially specific impact of impaired negative emotion recognition on social stress and psychotic experience are needed to corroborate our speculation.

Further, our high proneness group showed trends for faster categorization responses in affectively neutral face and voice stimuli as well as higher intensity ratings in affectively neutral face and a trend for higher intensity ratings in affectively neutral voice stimuli compared to the low proneness group. This might suggest a heightened sensitivity or greater attentional capture by affectively neutral stimuli in our high proneness sample, resulting in faster responses. Further, affectively neutral information might have elicited overly high levels of arousal in our high proneness subjects. This is consistent with previous studies reporting aberrant processing of affectively neutral information in patients with the diagnosis of a psychotic disorder (e.g. Dondaine et al., 2014; Garcia-Leon et al., 2021; Mitrovic et al., 2020; Pinkham et al., 2011) and high psychosis proneness (Seo et al., 2020), suggesting that psychosis might be associated with a greater sensitivity to affectively neutral stimuli. This might also be reflected by aberrant neurophysiological responses towards neutral stimuli, such as neural hyperactivity in frontal regions during emotion processing tasks, in psychosis (Lakis & Mendrek, 2013; Modinos et al., 2015; Underwood et al., 2015) and psychosis proneness (for a review, see Kozhuharova et al., 2020). It has been discussed that subjects with psychotic symptoms misattribute affective meaning to neutral cues (e.g. Kohler et al., 2003; Mitrovic et al., 2020), which is consistent with an erroneous attribution of salience to motivationally irrelevant information in psychosis due to a dopaminergic dysfunction (Howes & Kapur, 2009). This so-called dopamine hypothesis states that an increased salience attribution and insufficient filtering of irrelevant stimuli caused by dysregulated dopamine release and the following biased cognitive interpretation of the resulting perceptual anomalies facilitate the emergence of psychotic symptoms (Howes & Kapur, 2009; Howes & Murray, 2014). Further, it has been discussed that mis-labeling of neutral faces as negative might at least partly be the consequence of reduced visual scanning patterns of salient facial features for emotion recognition observed in psychosis (for a review, see Toh et al., 2011) and psychosis proneness (Hillmann et al., 2015, 2017). The vigilance-avoidance hypothesis assumes that psychosis, especially paranoia, is associated with an attentional bias avoiding threat-related features in facial cues to reduce arousal (Green et al., 2000; Green & Phillips, 2004). The interpretation of the resulting ambiguous experience might then be influenced by a jumping-to-conclusion bias (Freeman, 2016), increased threat expectation (Reininghaus et al.,

2016) and the own predominant emotional state of anxiety (Freeman, 2007). Hillman et al., (2018) argue that these mechanisms might provide a theoretical link between psychotic symptoms and affective misattribution to affectively neutral faces as well as negative mis-labeling of neutral face stimuli observed in psychosis (e.g. Mitrovic et al., 2020; Pinkham et al., 2011).

We observed interactions/non-significant trends for interactions between group and stimulus condition on RT and accuracy in happy, angry and neutral trials, especially in the attend voice condition. While the overall pattern of in-/congruency effects was observable in both groups, the effects of simultaneously presented emotionally in-/congruent stimuli of the unattended modality differed between groups in some cases. Numerically, it seemed that the high proneness group showed smaller in-/congruency effects in affectively neutral attend voice trials compared to the low proneness group. Further, it appeared as the low proneness group showed similar RT in happy face trials regardless of stimulus condition, i.e. no in-/congruency effects, whereas the high proneness group showed a congruency facilitation effect. This could suggest a ceiling effect in the low proneness group, reflecting that it was sufficiently easy to recognize happy faces for low proneness subjects that additional vocal cues had not further impact on RT.

It is important to note that we observed (trends for) group * stimulus condition interactions mainly in attend voice trials. It is possible that emotion-specific MSI deficits in psychosis proneness depend on the target modality, potentially reflecting modality dominance effects. This is consistent with previous studies suggesting that the influence of the unattended modality differed depending on the target modality in psychosis. Earlier studies indicated that the impact of facial expressions on vocal prosody recognition was increased in patients compared to healthy controls (de Gelder et al., 2005; de Jong et al., 2010) whereas the influence of emotional vocal prosody on facial emotion recognition was reduced (de Gelder et al., 2005). In contrast, de Jong et al. (2009) reported a reduced impact of facial expression on emotional vocal prosody recognition. In sum, empirical evidence regarding potential modality dominance aberrances on AV emotional MSI in psychosis is inconsistent and does not support clear conclusions if modality dominance aberrances are specific for psychosis. Future research should specifically investigate the influence of the target modality regarding the crossmodal impact of the unattended modality on AV emotion recognition in psychosis.

Our emotion-specific results and especially the observed (trends for) interactions should be interpreted with caution, since the original paradigm adopted from Föcker et al. (2011) was not designed for such fine-grained approaches. The resulting trial numbers per group * emotion category * attention condition * stimulus condition are small, hampering robust statistical testing. Therefore, these results should be considered rather as hints for future research regarding emotion-specific MSI deficits in psychosis proneness. To our knowledge, our study is the first to follow such a fine-grained approach of emotional MSI in subjects with psychotic symptoms. In previous studies, emotion-specific deficits in unisensory emotion processing have been reported in psychosis (Barkl et al., 2014; Bonfils et al., 2019, 2020; Mitrovic et al., 2020; Tripoli et al., 2022; Won et al., 2019). It is possible that aberrant patterns of emotional MSI in both patients and possibly psychosis proneness samples manifest particularly for those emotions, which have been found to show pronounced unisensory processing deficits, e.g. for negative and affectively neutral stimuli (Bonfils et al., 2020; Dondaine et al., 2014; Mitrovic et al., 2020; Seo et al., 2020; Tripoli et al., 2022; Won et al., 2019). Future research should specifically address potential emotion-specific patterns of emotional MSI in psychosis proneness and patient samples. Initial evidence for emotion-specific differences in multisensory compared to unisensory emotion processing can be found in a study by Thaler et al. (2013), who observed a positive misattribution of AV neutral stimuli in contrast to a negative misattribution of visual neutral stimuli. The authors suggest that the additional auditory information in crossmodal trials crucially impacted the misattribution valence, since auditory neutral stimuli were also positively misattributed (Thaler et al., 2013). This highlights the above mentioned potential role of modality effects on emotional MSI in psychosis. It is possible that interactions between modality effects and specific emotion categories, e.g. negative and neutral information, play a role in the manifestation of emotional MSI deficits in psychosis.

## 4.2 Associations between Emotional MSI and Psychotic Symptoms

### *4.2.1 Discussion and Comparison to Previous Evidence*

Our second aim consisted of investigating potential associations between emotional MSI and psychotic symptoms, such as perceptual anomalies and hallucinatory or paranoid experiences. We expected significant correlations between in-/congruency effects in emotional categorization and self-reported anomalous perceptual experiences, hallucinatory events, and paranoid convictions. We further expected correlations between in-/congruency

effects and negative schemata about the self and others, since negative beliefs have been discussed to influence the (psychotic) interpretation of aberrant perceptual experiences (Garety et al., 2001). After multiple comparison correction, we did not observe any significant correlation between IE difference scores and questionnaire scores in CAPS, LSHS-E, PCL and BCSS negative subscales. Thus, our analysis did not reveal any association between in-/congruency effects during emotional categorization of AV stimuli and self-reported experience of perceptual anomalies, hallucinations, paranoia and negative schemata.

In our exploratory correlation analysis between IE scores and questionnaire data, we observed a small, negative Pearson correlation between IE scores during categorization of emotionally congruent stimuli in the attend face condition and CAPS scores with a correlation coefficient $r$ = -.33 and an uncorrected $p$ = .005, which however did not survive multiple comparison, Bonferroni-corrected $p$ = .15. It is possible that our study was underpowered and therefore unable to detect a statistically significant association. However, we were only able to observe this one potentially meaningful association between IE scores and questionnaires instead of a broader pattern reflecting systematic associations between AV emotion categorization and psychotic symptoms in our sample. Our results do not support clear conclusions about association patterns between psychotic symptoms in our proneness sample and AV emotion processing. Future, larger-scale studies should develop hypotheses based on this hint in our results regarding associations between psychotic symptoms and processing of emotional faces with emotionally congruent vocal prosody. Given a well-designed future study, it could be speculated to potentially reveal that the higher the self-reported occurrence of perceptual anomalies, the better the categorization performance of emotional face expression with simultaneously presented emotionally congruent prosody. However, even if a larger-scale study would reveal a statistically significant correlation, it remains questionable which theoretical conclusions regarding AV emotion processing in psychosis proneness might be drawn from this isolated correlation. To date, it could be at best discussed to reflect a potential small processing benefit for emotionally congruent AV information in persons with higher occurrence of aberrant perceptual phenomena.

Only few studies on emotional MSI in psychosis investigated associations between multisensory emotion recognition and positive or negative symptoms. The majority of studies failed to observe correlations between emotional MSI and positive or negative symptoms (Giannitelli et al., 2015; Jeong et al., 2021; Mangelinckx et al., 2017; Müller et al., 2014;

Seubert, 2010), and only one study revealed an association to negative symptoms (Castagna et al., 2013). However, both Müller et al. (2014) and Portnova et al. (2021) reported associations between symptom severity and aberrant neural responses to AV emotional stimuli. This could suggest that deficient MSI might not show associations to specific symptoms of psychosis but rather to general symptom severity. It is possible that we failed to observe associations to psychotic symptoms, since we focused on positive symptoms rather than severity of symptoms associated with psychosis.

### *4.2.2 Associations to Positive vs. Negative Symptoms and a Role of Stimulus Complexity*

Previous studies on unisensory emotion processing in psychosis reported associations between facial emotion recognition deficits and negative symptoms (e.g. Edwards et al., 2001; Fakra et al., 2015; Turetsky et al., 2007). It is possible that we were unable to observe associations between AV emotion processing deficits and psychosis proneness because we focused on positive rather than negative symptoms. However, a recent review revealed associations between emotion recognition deficits of face expressions and negative symptoms whereas vocal prosody recognition showed associations with positive symptoms, suggesting a role of presentation modality (Lin et al., 2020). Further, Darke et al. (2021) observed associations between positive symptoms such as delusions and facial emotion processing deficits and argue that specific stimulus characteristics might have a crucial impact. Previous studies reporting associations to negative symptoms used static face stimuli, while Darke et al. (2021) and Johnston et al. (2010) observed associations to positive symptoms using dynamic face stimuli. Thus, association patterns between positive or negative symptoms and emotion processing deficits might differ depending on whether presented face stimuli are static or dynamic (Darke et al., 2021; Johnston et al., 2010).

It is possible that the usage of dynamic vs. static face stimuli has an impact on processing crossmodal emotional information in psychosis and on revealing associations between psychotic symptoms and deficient emotional MSI. The majority of previous studies on emotional MSI in patients with the diagnosis of a psychotic disorder used static rather than dynamic face stimuli (Castagna et al., 2013; de Gelder et al., 2005; de Jong et al., 2009, 2010; Giannitelli et al., 2015; Mangelinckx et al., 2017; Müller et al., 2014; Portnova et al., 2021; Seubert, 2010; Zvyagintsev et al., 2013), whereas only a subgroup of studies used dynamic voice stimuli (Fiszdon & Bell, 2009; Sestito et al., 2013; Simpson et al., 2013; Thaler et al., 2013). However, within both stimuli categories results of previous studies are inconsistent,

with patients showing either deficient or unimpaired emotional MSI and only very few studies reporting associations to symptoms of psychosis.

### 4.3 Limitations and Implications for Future Research

In addition to limitations of our study already mentioned earlier, e.g. the recruitment of a possibly high-functional sample and a potential lack of statistical power for correlation and emotion-specific analyses, following potential limitations and implications for future studies should be discussed.

Patients with the diagnosis of a psychotic disorder and persons with psychosis proneness showed a negativity bias in unisensory emotion processing (Pinkham et al., 2011; Seo et al., 2020), which might be associated with increased paranoia and greater difficulties in social situations (Pinkham et al., 2011). In our study, we focused on overall emotional MSI in psychosis proneness and explored emotion-specific patterns of crossmodal emotion recognition, analyzing correctly classified stimuli analogue to the analysis strategy by Föcker et al. (2011). As mentioned earlier, our experimental paradigm adopted by Föcker et al. (2011) restricted our possibilities for statistical testing regarding emotion-specific patterns, especially regarding statistically meaningful analyses of incorrect emotion classifications. It would be interesting to investigate whether our proneness sample showed a negativity bias in incorrectly classified emotional stimuli, i.e. misclassified stimuli as negatively valenced. Future studies should investigate whether a negativity bias can also be observed during multisensory emotion recognition both in patient and proneness samples.

A further potential limitation of our study might consist in the restriction to the emotion categories happy, angry, sad and neutral. Previous studies reported impaired processing of fearful stimuli and a bias to mislabel emotional cues as fearful in psychosis (e.g. Garcia-Leon et al., 2021; van Dijke et al., 2016). Further, the recognition of fearful face expressions seemed to be already particularly impaired in early stages of disorder progression - contrary to the recognition of other emotions, which rather seemed to deteriorate over the course of disorder development - and it has recently been suggested that impaired fear recognition in particular might be a vulnerability marker for psychosis (Pena-Garijo et al., 2023). Future implementations of emotional MSI paradigms in psychosis proneness should include fearful stimuli and investigate if psychosis proneness is associated with particular deficits in the recognition of fearful AV stimuli.

## 4.4 Conclusion

In conclusion, our findings do not offer support for the hypothesis that deficient emotional MSI might play a role in the development of psychotic disorders. Studies reporting unisensory emotion processing deficits in psychosis proneness assumed a potential predictive role of emotion processing deficits for symptom development and disorder progression (Corcoran et al., 2015; D. Martin et al., 2020; Seo et al., 2020; Tripoli et al., 2022; Tseng et al., 2016). Based on our findings and as mentioned throughout the discussion of the literature, it likely requires a more nuanced approach to study the interplay of both unisensory and multisensory emotion processing in the developmental trajectory from psychosis proneness to psychosis. Specifically, future studies on emotion processing in psychosis should take additional influential factors such as specific emotion categories, modality effects, symptom severity (Lin et al., 2020), or stimulus characteristics such as static vs. dynamic cues (Darke et al., 2021) and their potential interactions into account.

# Chapter III: Temporal Multisensory Integration & Crossmodal Temporal Recalibration in Psychosis (Study 2)

# 1 Introduction

## 1.1 Temporal Multisensory Processing

In everyday life, we receive a multitude of environmental and self-related information at any time point, with the temporal characteristics of incoming cues forming a temporal relationship. The temporal dynamics of stimuli, such as their temporal synchrony, are a central feature determining whether crossmodal stimuli should be perceptually integrated or not (Meredith et al., 1987; Stein & Meredith, 1993). The appropriate temporal ordering of events and judgements whether events are synchronous or not are crucial for the construction of a coherent percept and subsequent adaptive behaviour.

Uni- and crossmodal stimuli have been found to be integrated with a higher likelihood, if their stimulus onset asynchrony (SOA), i.e. the temporal delay between their respective onset, falls within in a so-called temporal binding window, TBW (Chen & Vroomen, 2013; Vroomen & Keetels, 2010; Wallace & Stevenson, 2014). Stimuli with a SOA smaller than this TBW are judged to likely belong to the same event due to their temporal alignment and perceptually integrated, a process called temporal MSI, and stimuli with a SOA larger than the TBW as belonging to different events (Wallace et al., 2020; Wallace & Stevenson, 2014). The TBW varies per subject (Stevenson et al., 2012; Wallace & Stevenson, 2014) and is narrower for simple stimuli such as arbitrary flashes and tones and wider for complex information such as speech (Vroomen & Keetels, 2010; Wallace & Stevenson, 2014).

Apart from the on-line integration of crossmodal temporal stimuli, the perceptual system also is able to adapt to temporal asynchronies of crossmodal stimuli, a phenomenon called crossmodal temporal recalibration (for reviews, see Chen & Vroomen, 2013; Vroomen & Keetels, 2010). For example, repeated exposure to a fixed temporal asynchrony between AV stimuli shifts the point of subjective simultaneity (PSS) of a following temporal order judgement (TOJ) or SJ of the AV stimuli in direction of the temporal asynchrony (Fujisaki et al., 2004; Vroomen et al., 2004). This cumulative crossmodal recalibration has been observed between various modality combinations such as AV, audio-tactile or visuo-tactile stimuli as well as in sensorimotor cues and both in simple and complex stimuli such as speech (Chen & Vroomen, 2013). It has been discussed that the perceptual system adapts to crossmodal asynchronies, caused e.g. by different physical signal transmission and arrival time or neural processing speed, due to a tendency to reduce discrepancies between the various sensory modalities (Simon et al., 2018; Vroomen & Keetels, 2010). In addition to this cumulative temporal recalibration building up by continuous exposure to a crossmodal asynchrony,

previous studies have observed rapid temporal recalibration occurring already after a single exposure to a crossmodal asynchrony (e.g. Van der Burg et al., 2013, 2015). For example, the PSS during a SJ task of an AV trial has been observed to shift in direction of the leading modality on the immediately preceding AV trial, indicating an immediate temporal recalibration effect induced by the immediate stimulus history (Van der Burg et al., 2013). Immediate temporal recalibration might benefit multisensory processing in a highly dynamic environment with continuously changing temporal relations between stimuli: a rapid adjustment to crossmodal asynchronies might facilitate MSI processes, since crossmodal gain is expected to be maximal during crossmodal synchrony (Van der Burg et al., 2015).

Further, the cumulative and immediate temporal recalibration effects have been discussed to reflect separate processes. A study by Van der Burg et al. (2015) successfully observed both effects in a SJ paradigm with prior exposure to AV asynchronies. The authors reported that cumulative recalibration was maximal directly after repeated asynchrony exposure and declined over time, whereas immediate recalibration changed trial-by-trial and seemed unaffected by the magnitude of cumulative recalibration. Van der Burg et al. (2015) argued that this suggested a relative independence of both recalibration processes and potentially indicated different mechanisms.

Studies in the spatial domain suggested that, although MSI and crossmodal recalibration are separate processes, both seemed to be at least partly related. Park & Kayser (2019) observed shared neural substrates in medial parietal regions of spatial MSI and immediate crossmodal recalibration, suggesting a link between both multisensory processes. Further, a study by Rohlf et al. (2020) observed MSI and immediate recalibration in children of seven years, whereas cumulative recalibration was only observed in later age groups. The authors proposed that MSI and immediate recalibration processes develop earlier in life compared to cumulative recalibration and speculated that especially MSI might be a process necessary for the development of cumulative recalibration (Rohlf et al., 2020). Bruns et al., (2022) demonstrated that congenital-cataract subjects, who were born blind and regained vision through surgery later in life, showed intact MSI and immediate recalibration, but atypical cumulative recalibration. Crucially, MSI correlated significantly with immediate recalibration (Bruns et al., 2022), a finding which was also observed in the study by Rohlf et al. (2020). In sum, previous evidence suggested that MSI and immediate crossmodal recalibration might be related multisensory processes, whereas cumulative recalibration develops later and potentially relies on functional development of especially MSI but also

immediate recalibration (Bruns et al., 2022; Rohlf et al., 2020). It is important to note, however, that the evidence suggesting such links (Bruns et al., 2022; Park & Kayser, 2019; Rohlf et al., 2020) focused on the spatial rather than the temporal domain. More research is needed to investigate if analogue links between temporal MSI and crossmodal temporal recalibration might exist.

In sum, the brain uses information about the temporal dynamics of environmental and self-related stimuli to constantly order perceptual events, judge whether stimuli stem from the same source and recalibrate for small temporal discrepancies (Chen & Vroomen, 2013; Meredith et al., 1987; Stein & Meredith, 1993). A failure in appropriate temporal integration and recalibration of crossmodal stimuli might drastically impair the construction of a coherent percept and the ability to adapt to a constantly changing environment, which might facilitate perceptual-cognitive disintegration phenomena in neurodevelopmental disorder such as psychosis (Hornix et al., 2019; Postmes et al., 2014; Uhlhaas & Mishara, 2006).

### 1.2 Psychosis, Time Perception and Temporal Processing

There has been considerable work on time perception and temporal processing in psychosis. Numerous studies reported that psychosis is associated with a wide range of time processing impairments, with patients with the diagnosis of a psychotic disorder showing altered time perception, such as overestimation of time periods or difficulties in rhythm detection, and deficient temporal processing, such as poorer performance in SJ or TOJ tasks, compared to healthy controls (for reviews, see Amadeo et al., 2022; Ciullo et al., 2016; Thoenes & Oberfeld, 2017).

It has been reported that time perception was accelerated in psychosis, possibly due to a hyperfunction of an internal pace maker (Ueda et al., 2018). This pace maker is thought to regulate time perception and processing by emitting pulses used for e.g. estimation of time periods (Hartcher-O'Brien et al., 2016). Crucially, the internal pace maker is thought to be regulated via the dopamine system (Cheng et al., 2007), which has been found to be dysregulated in psychosis (Howes & Kapur, 2009; Kapur, 2003), suggesting a connection between hypervigilance in psychosis and aberrant time processing mechanisms (Ueda et al., 2018).

It has been proposed that a dysfunction in the perception and processing of temporal information constitutes a core deficit in psychosis (e.g. Amadeo et al., 2022; Andreasen et al., 1999) and that psychotic symptoms might be the consequence of impaired temporal organization of information processing (Thoenes & Oberfeld, 2017). The model of cognitive

dysmetria proposes that dysfunctional temporal information processing in psychosis might trigger a detrimental cascade of effects, such as inappropriate temporal linking of external and internal cues, hampering the formation of coherent percepts and facilitating the emergence of psychotic symptoms (Andreasen et al., 1999). Support for a connection between psychotic symptoms and impaired time processing comes from a recent meta-analysis reporting an association between positive symptoms in psychosis and impaired performance on timing tasks (Ueda et al., 2018). Further, dysfunctional time perception and temporal processing have also been observed in individuals with psychosis proneness and high genetic risk for psychosis, suggesting a potential endophenotype for psychosis (Amadeo et al., 2022).

### 1.3 Temporal Multisensory Processing in Psychosis

Aberrances in temporal MSI in psychosis have been reported in previous studies, with patients with the diagnosis of a psychotic disorder showing an enlarged TBW during temporal multisensory processes compared to healthy controls (Noel et al., 2018; for a review and meta-analysis, see Zhou et al., 2018; Zhou, Lai, et al., 2022; Zopf et al., 2021). This enlargement of the TBW in psychosis has also been reported within unisensory modalities, with an enlargement in the range of tens of milliseconds (Amadeo et al., 2022). For crossmodal stimuli however, the enlargement of the TBW increases up to hundreds of milliseconds (Amadeo et al., 2022), suggesting pronounced impairments in multisensory compared to unisensory temporal processing. This is supported by a study of Stevenson, Park et al. (2017), who reported that AV SJ performance deficits in patients with the diagnosis of a psychotic disorder were larger than impaired performance within visual or auditory modalities. Further, enlargements of the TBW in psychosis might be pronounced in speech compared to simple stimuli (Foucher et al., 2007; Noel et al., 2018), indicating a possible influence of stimulus complexity and suggesting potentially detrimental consequences for communication.

Impaired temporal multisensory processing in terms of an enlarged TBW has been discussed to represent a potential core deficit in psychosis and already in stages prior to disorder onset, e.g. in psychosis proneness (Amadeo et al., 2022). This is supported by findings of Stevenson, Park et al. (2017), who observed that an enlarged TBW predicted hallucination severity in patients with the diagnosis of a psychotic disorder. Further, several studies reported associations between an enlarged TBW and higher schizotypy as well as symptoms such as disorganization and hallucinations in psychosis proneness (Dalal et al., 2021; Di Cosmo et al., 2021; Ferri et al., 2017, 2018; Marsicano et al., 2022).

An enlarged TBW might form a sensory-perceptual basis for psychosis and facilitate the development of psychotic symptoms (Zhou et al., 2018). An enlarged TBW in psychosis might lead to integration of crossmodal stimuli, which would not be integrated in healthy controls. Inappropriate temporal integration of sensory input might facilitate insufficient information filtering and ambiguous perception (Sartorato et al., 2017), leading to deficits in organizing sensory information into coherent percepts (Uhlhaas & Mishara, 2006) and to psychotic symptoms such as hallucinations (Stevenson, Park, et al., 2017) and self-disorders (Postmes et al., 2014).

## 1.4 Crossmodal Temporal Recalibration in Psychosis and other Neurodevelopmental Disorders

While there has been considerable work on unisensory and multisensory temporal processing in psychosis (for reviews, see Amadeo et al., 2022; Ciullo et al., 2016; Thoenes & Oberfeld, 2017; Wallace et al., 2020; Wallace & Stevenson, 2014; Zhou et al., 2018), to our knowledge studies investigating crossmodal temporal recalibration in psychosis are extremely scarce.

Initial evidence regarding crossmodal temporal recalibration in patients with the diagnosis of a psychotic disorder can be found in a study by Zhou, Cui, et al. (2022). Using an SJ paradigm with simple AV stimuli and AV speech stimuli, Zhou, Cui, et al. (2022) observed that adolescent patients with the diagnosis of EOS showed a similar magnitude of immediate temporal recalibration compared to age-matched controls. While this suggested that early-onset psychosis might be associated with intact recalibration of SJ to the immediate stimulus history, generalizability to adult-onset psychosis[4] might be drastically limited. EOS is a particularly rare form of psychosis (Driver et al., 2013) and has been found to exhibit significantly higher symptom load as well as premorbid cognitive impairments (e.g. Vyas et al., 2011), suggesting that the disorder trajectory and perceptual-cognitive deficits might not be entirely comparable to adult-onset psychosis. Further, multisensory temporal processing performance varies with age (Noel et al., 2016), with reports of immediate temporal recalibration being observed in adults both in simple and speech stimuli but not adolescents, who only showed recalibration in speech stimuli (Zhou et al., 2020). Thus, conducting a crossmodal temporal recalibration paradigm in adult patients with the diagnosis of adult-onset

[4] Whenever the terms „psychosis“ or “diagnosis of a psychotic disorder” are mentioned throughout this chapter, they refer to adult-onset psychosis. EOS/early onset psychosis will be mentioned as such to distinguish both disorder manifestations.

psychotic disorder is necessary to a) eliminate age-related confounding factors on multisensory temporal processing and b) collect evidence on both cumulative and immediate temporal recalibration in the more commonly encountered adult-onset psychosis as opposed to the highly rare EOS.

Although autism spectrum disorders (ASD) and psychosis compose distinct disorders with different clinical manifestations, it has been discussed that both disorders might be partly related (e.g. King & Lord, 2011) with ASD and psychosis sharing similarities e.g. in neural connectivity impairments (Friston et al., 2016; Just, 2004) or genetic factors (Carroll & Owen, 2009). Further, previous work suggested that ASD and psychosis share dysfunctions in uni- and multisensory processes (Wallace et al., 2020; Wallace & Stevenson, 2014; Zhou et al., 2018). Patients with the diagnosis of an ASD have been found to show comparable impairments in temporal multisensory processes as patients with the diagnosis of a psychotic disorder, such as an enlarged TBW in crossmodal stimuli (Wallace & Stevenson, 2014; Zhou et al., 2018). Given these overlaps between both neurodevelopmental disorders, studies investigating crossmodal temporal recalibration in ASD might give further hints for hypothesis generation of potential crossmodal temporal recalibration alterations in psychosis.

A study by Turi et al. (2016) investigated immediate temporal recalibration in autism and observed that adult patients with the diagnosis of an ASD showed little to no immediate recalibration in a SJ task compared to healthy controls. Both Noel, De Niear et al. (2017) and Zhou, Cui, et al. (2022) extended these findings to samples consisting of children and adolescents with the diagnosis of an ASD, reporting reduced immediate temporal recalibration in SJ tasks to immediately preceding asynchronies in ASD compared to age-matched control groups. Another study by Stevenson, Toulmin, et al. (2017) observed an association between increased autistic traits and a reduced size of the cumulative recalibration effect in a SJ task in a healthy student sample. This suggested that symptoms of ASD might be associated with decreased cumulative temporal recalibration (Stevenson, Toulmin, et al., 2017). In sum, prior evidence indicated that ASD might be associated with deficient immediate and cumulative temporal recalibration, which adds to previously published evidence regarding deficits in temporal MSI in ASD (Zhou et al., 2018).

In sum, previous studies reported dysfunctional unisensory and multisensory temporal processing in psychosis (e.g. Amadeo et al., 2022; Thoenes & Oberfeld, 2017; Zhou et al., 2018) and altered crossmodal temporal recalibration in ASD (Noel, De Niear, et al., 2017; Stevenson, Toulmin, et al., 2017; Turi et al., 2016; Zhou, Cui, et al., 2022), a disorder

demonstrating comparable deficits in temporal MSI compared to psychosis (Zhou et al., 2018). Further, it has been discussed that MSI and crossmodal recalibration might be partly related processes at least in the spatial domain (Bruns et al., 2022; Park & Kayser, 2019; Rohlf et al., 2020). Thus, it can be speculated that crossmodal temporal recalibration is altered in psychosis. Altered or dysfunctional cumulative and immediate temporal recalibration processes might have detrimental effects for the perceptual adaptation to natural temporal asynchronies and impair the formation of coherent percepts of the environment and the self.

### 1.5 Research Questions

The aim of this study was to investigate if psychosis is associated with impaired crossmodal temporal recalibration during an SJ task. Further, we intended to replicate earlier findings of an increased TBW window in simple AV stimuli in psychosis (Zhou et al., 2018; Zhou, Lai, et al., 2022). Finally, we aimed to reveal associations between impaired multisensory temporal processing (i.e. both integration and recalibration) and psychotic symptoms, thereby extending previous findings which had focused on associations between an enlarged TBW and psychotic symptoms in patients and psychosis proneness (Dalal et al., 2021; Di Cosmo et al., 2021; Ferri et al., 2017, 2018; Foucher et al., 2007; Marsicano et al., 2022; Stevenson, Park, et al., 2017).

Patients with the diagnosis of psychotic disorder and healthy controls judged if AV stimuli presented with varying SOAs appeared simultaneously in several SJ test phases. Before each SJ test phase, the same AV stimuli were presented with a fixed temporal lag - half of the phases a visual lead adaptation phase, the other half auditory lead - to induce cumulative temporal recalibration, i.e. a change in PSS depending on the modality order of the previous adaptation phase. Further, immediate temporal recalibration was assessed by examining the PSS of SJ responses depending on the modality order of the immediately preceding SJ trial (paradigm adopted from Van der Burg et al., 2015).

We expected the patient group to showed a wider distribution of positive SJ responses compared to healthy controls, i.e. to judge AV stimuli as simultaneous more often in larger SOAs, indicating an enlarged TBW. Further, based on findings in ASD (Noel, De Niear, et al., 2017; Turi et al., 2016; Zhou, Cui, et al., 2022), we expected altered crossmodal temporal recalibration in psychosis: We expected the patient group to show a smaller difference in PSS between visual lead adapted SJ trials and auditory lead adapted SJ trials compared to controls, indicating altered cumulative temporal recalibration. We also expected the patient group to show a smaller PSS-difference between SJ trials immediately preceded by a visual first SJ

trial and preceded by an auditory first SJ trial compared to healthy controls, indicating altered immediate temporal recalibration. Finally, we expected the TBW, the cumulative, and the immediate temporal recalibration effect to show associations with self-reports about anomalous perceptual experiences, hallucinations, delusions, passivity experiences and negative schemata about the self and others.

# 2 Methods

## 2.1 Participants

28 patients with the diagnosis of a psychotic disorder and 27 healthy controls originally enrolled in the study. Patients were recruited via flyers in outpatient clinics and social care centers as well as contact data bases from earlier studies with consent for further contact, healthy controls via online recruitment services from the University of Hamburg. Two patients discontinued the participation. We further had to exclude one patient from statistical analysis due to random response behavior in the experimental task (see **2.4 Data Analysis**). The final sample therefore consisted of 25 patients with the diagnosis of a psychotic disorder and 27 healthy controls. Groups did neither differ in age, $t(49.4) = 1.02$, $p = .31$, Cohen's $d = -0.28$, nor regarding the distribution of gender, $X^2(1) = 1.90$, $p = .168$, $\varphi = 0.19$, but differed in distribution of education level, $X^2(3) = 9.51$, $p = .023$, $\varphi = 0.43$ (see Table 3 for demographic data).

The inclusion criteria for patients were defined as the Diagnostic and Statistical Manual for Mental Disorders (DSM-V; American Psychiatric Association, 2022) criteria fulfillment for schizophrenia or schizoaffective disorder. Patient-specific exclusion criteria consisted of a comorbid autism spectrum disorder. Exclusion criteria for healthy controls consisted of a past or present psychiatric disorder or a family history of first degree relatives with the diagnosis of a psychotic disorder. General exclusion criteria were defined as dementia or other neurological disorders, acute suicidality, substance abuse disorder in the last 6 months, uncorrected sight or hearing impairment as well as inability to give consent or understand study instructions due to health status or language barriers.

Prior to participation, informed consent was discussed with participants and obtained in written form. Further, participation was compensated by 10€/hour or course credits. The study was performed by the standards of the Declaration of Helsinki and was approved by the Local Ethics Committee of the Faculty of Psychology and Human Movement Science, University of Hamburg.

***Table 3***

*Demographic and Diagnostic Data Separated by Group*

| | Patients | Healthy Controls |
|---|---|---|
| *n* | 25 | 27 |
| Age, *M* (range) | 41.2 (23-63) | 37.9 (21-63) |
| Gender, *n* (%) | | |
| female | 10 (40) | 17 (63.0) |
| male | 15 (60) | 10 (37.0) |
| diverse | 0 | 0 |
| Education, *n* | | |
| 9 years | 0 | 1 |
| 10 years | 7 | 4 |
| 12-13 years[a] | 15 | 9 |
| university degree[b] | 3 | 13 |
| CAPS, *M* (*SD*) | 11.6 (5.8) | 1.4 (2.0) |
| LSHS-E, *M* (*SD*) | 24.3 (11.9) | 3.2 (4.1) |
| PCL, *M* (*SD*) | 94.2 (52.3) | 26.0 (26.2) |
| Passivity experiences, *M* (*SD*) | 7 (4.5) | 0.1 (0.5) |
| BCSS n.S., *M* (*SD*) | 6.6 (4.5) | 2.0 (2.5) |
| BCSS n.O., *M* (*SD*) | 9.8 (5.6) | 3.5 (3.8) |

*Notes.* $N = 52$. CAPS = Cardiff Anomalous Perception Scale, total score; LSHS-E = Launey-Slade Hallucination Scale-Extended, total score; PCL= Paranoia Checklist, total score; BCSS n.S. = Brief Core Schema Scales, negative self subscale; BCSS n.O. = Brief Core Schema Scales, negative other subscale.

[a] due to changes in German education law in 2007, the years of education for acquire the Abitur (A-level/high school equivalent) vary between 12-13 years depending on school and Bundesland; [b] university degree: Bachelor level or higher.

## 2.2 Stimuli and Experimental Paradigm

The experimental paradigm was adopted from Van der Burg et al. (2015). Stimuli were presented and participants' behavioural data recorded by means of the Neurobs Presentation® Software (Version 22; Neurobehavioural Systems Inc., Berkeley, California, USA).

The experiment took place in a dimly-lit room. Participants were seated in front of a computer screen (61cm diameter, screen resolution: 1920 x 1200), resting their head on a chin rest at 40cm distance to the screen. Over the course of the whole experiment, a white fixation dot was continuously presented in the center of the black screen. Visual stimuli consisted of a white ring (radius 2.6° and width 0.4° visual angle, duration 50ms), which was presented in the screen center. The auditory stimulus consisted of a sinusoidal tone (500Hz, duration 50ms with 0.5ms rise and fall time) and was presented via headphones (Sennheiser HD 65TV, Sennheiser electronic GmbH & Co. KG, Wedemark, Germany).

The paradigm consisted of five blocks: a pretest block and four experimental blocks, each with alternating adaptation and SJ test phases. In adaptation phases, both the visual and auditory stimulus were presented with a fixed SOA of ±200ms, followed by an Inter-Trial-Interval (ITI), which randomly varied from 650 - 850ms in 50ms steps. AV adaptation trials were repeated 235 times per adaptation phase, resulting in a duration of approximately 3min for each adaptation phase. In half of the adaptation phases, the visual stimulus was leading (visual lead, SOA +200ms), in the other half the auditory stimulus was leading (auditory lead, SOA -200ms). Within each of the four experimental blocks, one visual lead and one auditory lead adaptation phase was presented. For half of the participants, each experimental block first contained a visual lead and then an auditory lead adaptation phase and vice versa for the other half of participants. Participants were asked to pay attention to both stimuli while fixating the white fixation dot.

Within each experimental block, each adaptation phase was followed by a SJ test phase, resulting in two adaptation conditions: visual lead adapted test trials and auditory lead adapted test trials. In test phases, both the visual and auditory stimuli were presented with a varying SOA (0, ±64, ±128, ±256 or ±512ms, with positive SOAs reflecting visual stimulus first trials). Each SOA condition was presented multiple times, with the smaller SOA conditions (i.e. 0, ±64) being presented 28 times each and the larger (i.e. ±128, ±256 or ±512ms) 14 times each, resulting in 98 test trials per test phase. Immediately following stimulus presentation, participants were asked to indicate per button press if AV stimuli were

synchronous or not, i.e. give a SJ. Responses were given by means of response buttons, which were placed left and right of the chin rest in comfortable position to be reached by the participant. Each button was assigned one answer, i.e. "synchronous" vs. "not synchronous" and the allocation to the left and right button was counterbalanced over participants. In SJ test phases, the respective answer for each button was continuously presented at the bottom left and bottom right of the computer screen. The visual stimulus always remained on screen until the participant pressed a button, while the auditory stimulus was presented for 50ms. Following the participant's response, the next test trial started after an inter-trial-interval, which randomly varied from 650 - 850ms in 50ms steps. The pretest block had the exact same structure as one test phase.

At the beginning of each phase, the instruction of the phase type (adaptation or SJ test phase) was presented at the center of the screen and participants started the phase by pressing Enter on a custom keyboard. After each experimental block consisting of two adaptation and two test phases in alternating order, participants could take a short break of 2-3 min before the experimenter started the next experimental block. Breaks within a block were not allowed.

### 2.3 General Procedure

The study was split into two sessions, which took place on two different days. Each session had a duration of about 1.25 to 1.5 hours.

In the session one, informed consent was discussed with participants, followed by diagnostic procedures enquiring both in- as well exclusion criteria and data relevant for analysis. Participants gave information about age, gender, education level, potential sight or hearing impairment and respective corrections, potential psychiatric diagnoses and their current status, illness duration of psychotic disorders, potential family history of psychotic disorder, psychopharmacological medication, wakefulness, motivation, stress level and substance consumption in the last 24 hours. Further, the Edinburgh Handedness Inventory (EHI; Oldfield, 1971) was conducted measuring handedness, followed by the TMT Version A and B (Reitan & Wolfson, 1985) for attention and processing speed. Following these shorter diagnostic procedures, in- and exclusion criteria regarding psychiatric diagnoses were checked by means of the Structural Clinical Interview for DSM-V (SCID-V; Beesdo-Baum et al., 2019). For healthy controls, the sections A "Affective Disorders" and B "Psychotic Disorders" were conducted, followed by a screening for the remaining sections. If the screening question for a section was scored with "yes", the respective section was conducted

in full. In patients with the diagnosis of a psychotic disorder, sections A and B were also conducted. Instead of the screening, however, the Psychotic Symptom Rating Scales (PSYRATS; Haddock et al., 1999) was administered to obtain further details regarding acoustic verbal hallucinations and delusions. Finally, vigilance was measured by means of the Continuous Performance Test Identical Pairs (CPT-IP; Nuechterlein et al., 2008)).

Participants were asked to provide self-reports about psychosis-related constructs in-between sessions. At the end of session one, participants received a questionnaire set consisting of questions regarding usual substance consumption and the following questionnaires: the BCSS (Fowler et al., 2006) for schemata about the self and others, the CAPS (Bell et al., 2006) for anomalous perceptual experiences, the LSHS-E (Lincoln et al., 2009; Siddi et al., 2019) for hallucinations, the PCL (Freeman et al., 2005) for paranoia and questions for the factor *bizarre experiences* of the Community Assessment of Psychic Experiences (Schlier et al., 2015; Stefanis et al., 2002) for passivity experiences (a glossary of the instruments and questionnaires used in this dissertation project can be found in Table A.1 in Appendix A). Participants were asked to fill in the questionnaires at home and hand them back at the second session. Potential questions regarding the questionnaire set were addressed at the end of session two.

Depending on the participant's individual schedule, session two took place within 1-10 days after session one. During session two, the experimental paradigm (see **2.2 Stimuli and Experimental Paradigm**) was conducted. Participants were comfortably seated in a chair in front of the experimental computer placing their chin on the chin rest. Experimental instructions were read to participants and potential questions clarified. The pretest block was then started, followed by blocks one to four with short breaks in-between each block. The experimenter remained in the experimental room and sat in a chair with a distance about 2.5 meters to the participant surveying the experimental progress. Following the experimental run, participants were de-briefed and received participation compensation.

### 2.4 Data Analysis

The preprocessing steps of raw behavioral data were adapted from Van der Burg et al. (2015). Data preprocessing and analysis was conducted using R Version 4.2.2 (R Core Team, 2022).

Prior to statistical analysis, we visualized each participant's SJ response ratio relative to each SOA to judge if the SJ responses relative to each SOA approximated a Gaussian

distribution. This was done for each of the following preprocessing steps with Gaussian fits: for each participant, the actual SJ responses relative to each SOA and the curve of the Gaussian fit were visualized in the same plot to detect random response behaviour. Based on this procedure, one patient with random response behaviour was excluded from further analysis.

To investigate potential group differences in overall SJ performance, we calculated the percentage of correct SJ per participant over all SJ test trials and compared both groups by means of a Welch two sample *t*-Test. Further, we fit a Gaussian distribution over all SJ test trials with the parameters mean, bandwidth, and height per participant and defined the bandwidth as an estimation of the general TBW. The TBW was compared between groups by means of a Welch two sample *t*-test.

To investigate cumulative temporal recalibration, i.e. the effect of AV adaptation phases on following SJ responses, we fit a Gaussian distribution with the parameters mean, bandwidth, and height over the first 50 trials of SJ test phases per participant. Per adaptation condition (visual lead adapted vs. auditory lead adapted), we pooled the first 50 test trials over all 4 blocks, resulting in 200 trials per adaptation condition. We calculated a *Group* (patients vs. healthy controls) * *Adaptation Condition* (visual lead adapted vs. auditory lead adapted) mixed ANOVA on the mean of the Gaussian fit over the first 50 test trials as an estimation of the PSS, followed by Bonferroni-corrected post-hoc *t*-tests. The Gaussian fit over the first 50 test trials per adaptation condition was chosen since we expected the cumulative recalibration effect to be largest at the beginning of each test phase based on the results of Van der Burg et al. (2015). The Greenhouse-Geisser correction (Greenhouse & Geisser, 1959) was applied in case of sphericity violations.

To investigate immediate temporal recalibration, we fit a Gaussian distribution with the parameters mean, bandwidth, and height over all SJ test trials per participant, separately for test trials immediately preceded by a visual first (i.e. with an SOA > 0ms) vs. by an auditory first (i.e. with an SOA < 0ms) test trial. The mean of the Gaussian fits was defined as an estimate for the PSS and entered in a *Group* (patients vs. healthy controls) * *Modality order on t-1* (prior visual first vs. prior auditory first) mixed ANOVA, followed by Bonferroni-corrected post-hoc *t*-tests. The Gaussian distribution was fit over all test trials since we expected the immediate recalibration effect to be stable over course of the experiment based on the results of Van der Burg et al. (2015).

We further investigated the time course of the cumulative and immediate recalibration effects by fitting a Gaussian distribution with the parameters mean, bandwidth and height over the respective SJ test trials for both recalibration effects (s.a.) in moving trial window bins of 50 trials each, starting with trials 1 to 50, fitting the Gaussian distribution, shifting the trial window by one trial (i.e. 2 to 51), fitting and so on, resulting in 49 Gaussian fits. For each of the 49 Gaussian fits, i.e. per trial window, we calculated the difference between the mean of the fit, as an estimate of the PSS, over visually lead adapted minus the mean of the fit over auditory lead adapted test trials for the cumulative recalibration effect, and between the mean of the fit over test trials preceded by a visual first minus the mean of the fit over by an auditory first test trial for the immediate recalibration effect. The respective PSS-difference scores were then entered in FDR-corrected Welch two sample *t*-tests per trial window to compare the groups at each time point, for both recalibration effects separately. Further, we calculated FDR-corrected one-sample *t*-tests against zero per trial window separately for each group for both recalibration effects.

Demographic data, i.e. age, gender, and education, was statistically compared between groups by means of a Welch two sample *t*-test for age and Pearson $X^2$-tests for gender and education distribution, respectively.

TMT-A and -B completion times (in seconds) were each compared between groups by means of Welch two sample *t*-tests to check for group differences in attention and processing speed. Further, a *Group* (patients vs. healthy controls) * *Digit Load* (2 vs. 3 vs. 4 digits) mixed ANOVA on CPT-IP *d'* scores was conducted to check for group differences in vigilance, followed by Bonferroni-corrected post-hoc *t*-tests. If group differences in TMT or CPT-IP were revealed, a potential influence of the respective control variable on the TBW, the cumulative, and the immediate recalibration effect was estimated. For the cumulative recalibration effect, we calculated the PSS-difference in the first 50 visual lead minus auditory lead adapted test trials. For the immediate recalibration effect, we calculated the PSS-difference in test trials preceded by a visual first minus by an auditory first test trial. Spearman's rank correlation coefficients ($\rho$) with Bonferroni-corrected *p*-values were then calculated between the TMT completion time scores and CPT-IP *d'*, respectively, and the general TBW (defined as bandwidth over all trials), the cumulative recalibration effect in the first 50 test trials, and the immediate recalibration effect.

Questionnaire scores, i.e. CAPS, LSHS-E, PCL, passivity experiences and BCSS scores, were compared between groups by means of Welch two sample t-tests, separately for each questionnaire.

Spearman rank correlation coefficients $\rho$ with Bonferroni-corrected $p$-values were calculated between experimental data (i.e. general TBW, the cumulative recalibration effect in the first 50 test trials, and the immediate recalibration effect) and questionnaire scores to investigate associations between TBW as well as cumulative and immediate temporal recalibration and self-reports of psychosis-related constructs. For all correlation analyses, Spearman rank correlation coefficients $\rho > .3$ were marked in correlation matrices and regarded potentially considerable for further discussion.

Additionally, Bayesian hypothesis testing using standard priors in JASP Version 0.16.4 (Wagenmakers et al., 2018) was conducted analogue to the frequentist statistical tests for experimental and questionnaire data described above, with Bayes Factors (BF) indicating evidence in favor of the null or alternative hypothesis. BF are classified per convention as follows: a BF of 1 is discussed to reflect no evidence, between 3-10 moderate, between 10-100 strong, and >100 extreme evidence in favor of the alternative hypothesis. Accordingly, a BF between 1/3-1/10 is discussed to reflect moderate, between 1/10-1/100 strong, and <1/100 extreme evidence in favor of the null hypothesis (an extensive BF classification scheme can be found in Wagenmakers et al., 2018). We reported $BF_{10}$ for two-tailed tests and correlations, the BF describing evidence for the alternative relative to the null hypothesis, and $BF_{incl}$ for mixed ANOVAS, the BF reflecting the posterior probability for model inclusion for a factor or interaction term averaged over all candidate models (Wagenmakers et al., 2018).

# 3 Results

## 3.1 Experimental data

### *3.1.1 Percent of Correct SJ*

Patients ($M$ = 50.82, $SD$ = 8.45) and healthy controls ($M$ = 50.18, $SD$ = 7.46) did not differ in percentage of correct SJ, $t(48.03) = 0.28$, $p = .773$, Cohen's $d = 0.08$, $BF_{10} = 0.29$.

### *3.1.2 TBW – Bandwidth over all Test Trials*

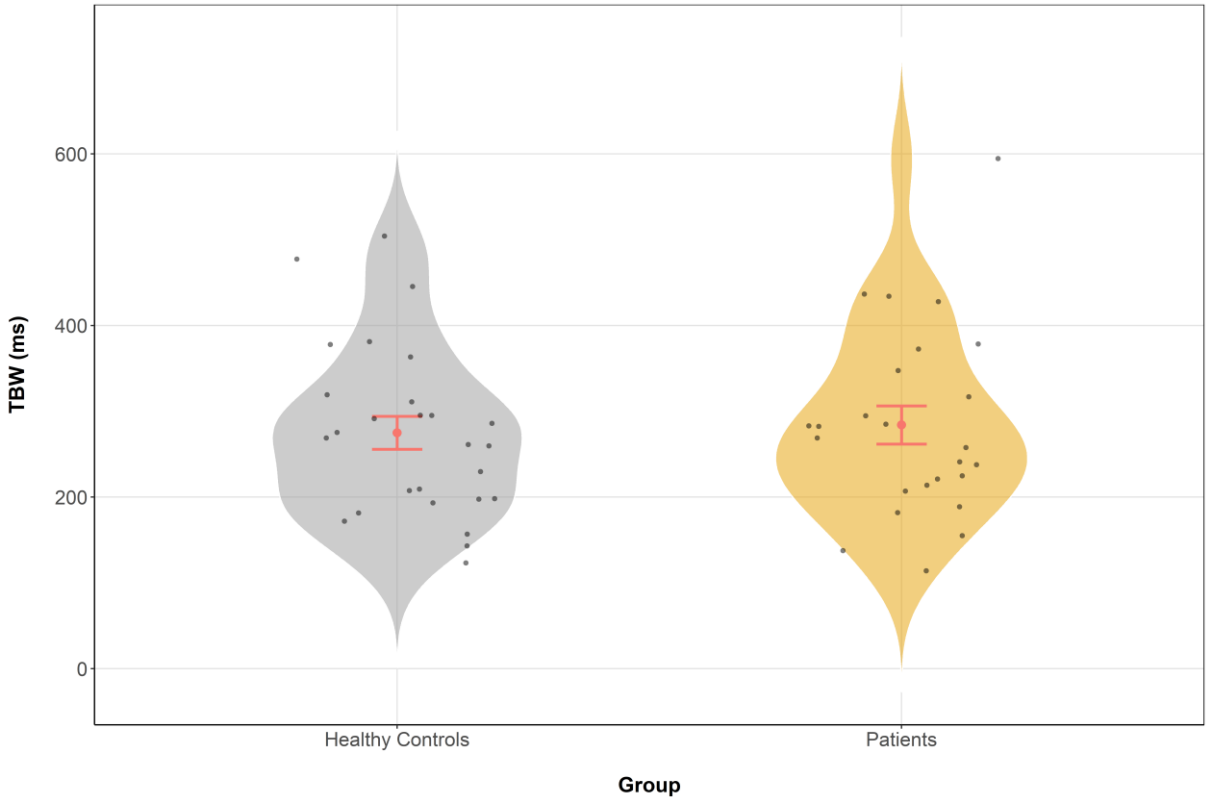
The general TBW (defined as the bandwidth of the Gaussian fit over all test trials) did not statistically differ between patients ($M$ = 284.06, $SD$ = 110.79) and healthy controls ($M$ = 274.96, $SD$ = 100.04), $t(48.44) = 0.31$, $p = .758$, Cohen's $d = 0.09$, $BF_{10} = 0.29$ (see Figure 5 for the distribution of the TBW in both groups).

### *3.1.3 Cumulative Temporal Recalibration*

A *Group* (patients vs. healthy controls) * *Adaptation* (visual lead adapted vs. auditory lead adapted) mixed ANOVA revealed a significant main effect of *Group*, $F(1, 50) = 4.71$, $p = .035$, $\eta^2_p = 0.09$, $BF_{incl} = 1.48$, indicating that the overall PSS in the first 50 trials (defined as the mean of the Gaussian fit over the first 50 test trials) differed between patients ($M$ = 67.21, $SD$ = 48.33) and healthy controls ($M$ = 42.7, $SD$ = 32.09). Further, there was a significant main effect of *Adaptation*, $F(1, 50) = 5.04$, $p = .029$, $\eta^2_p = 0.09$, $BF_{incl} = 1.63$, indicating that the PSS in the first 50 test trials differed between visual lead adapted ($M$ = 63.31, $SD$ = 47.59) and auditory lead adapted ($M$ = 45.66, $SD$ = 53.57) test trials. There was no *Group* * *Adaptation* interaction, $F(1, 50) = 0.23$, $p = .631$, $\eta^2_p < 0.01$, $BF_{incl} = 0.53$.

***Figure 5***

*Distribution and Mean of Temporal Binding Window Values per Group*




*Notes.* $n_{PAT} = 25$; $n_{HC} = 27$. *TBW* = temporal binding window. The general TBW was defined as the bandwidth of a Gaussian fit over all test trials. Error bars denote ±1 standard error of the mean.

### *3.1.4 Immediate Temporal Recalibration*

A *Group* (patients vs. healthy controls) * *Modality order on t-1* (prior visual first vs. prior auditory first) mixed ANOVA revealed a significant main effect of *Group*, $F(1, 50) = 4.45$, $p = .040$, $\eta^2_p = 0.08$, $BF_{incl} = 1.33$, indicating that the overall PSS over all test trials (defined as the mean of the Gaussian fit over all test trials) differed between patients ($M = 61.74$, $SD = 42.27$) and healthy controls ($M = 38$, $SD = 38.86$) . Further, there was a significant main effect of *Modality order on t-1*, $F(1, 50) = 10.91$, $p = .002$, $\eta^2_p = 0.18$, $BF_{incl} = 21.26$, indicating that the PSS (defined as the mean of the Gaussian fits over all test trials) differed between test trials preceded by a visual first ($M = 63.64$, $SD = 50.02$) and an auditory first ($M = 35.20$, $SD = 53.81$) test trial. There was no *Group* * *Modality order on t-1* interaction, $F(1, 50) = 0.02$, $p = .876$, $\eta^2_p < 0.01$, $BF_{incl} = 0.61$.

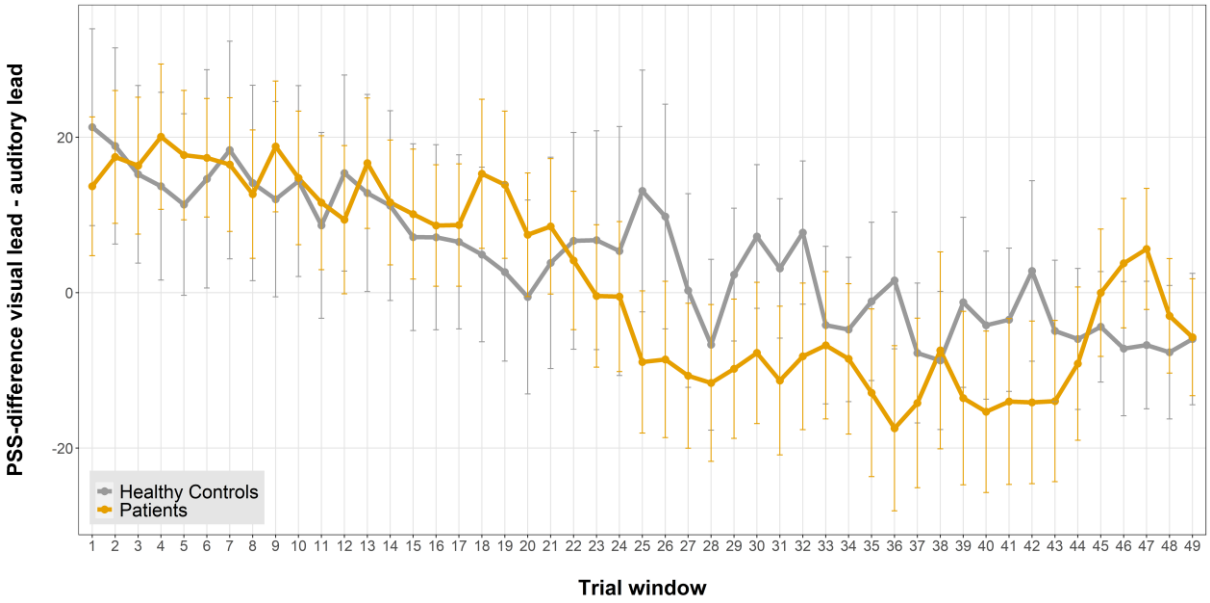
### *3.1.5 Time Course of Cumulative and Immediate Recalibration*

**Time Course of Cumulative Recalibration.** There were no statistically significant differences between groups in the PSS-difference between visual lead vs. auditory lead

adapted test trials, i.e. the cumulative recalibration effect, in moving trial windows (i.e. 1-50, 2-51, and so on), all uncorrected $p > .05$, all FDR-corrected $p = .984$, all $BF_{10} < 1$ (see Figure 6 for the time course of the cumulative temporal recalibration effect for both groups). Further, one-sample *t*-Tests against zero per group and trial window revealed no significant difference to zero in any trial window for healthy controls, all uncorrected $p > .05$, all FDR-corrected $p = .906$, all $BF_{10} < 1$. Before FDR-correction, patients showed a PSS-difference different from zero in trial window 4, uncorrected $p = .042$, $BF_{10} = 1.47$, trial window 5, uncorrected $p = .044$, $BF_{10} = 1.42$, trial window 6, uncorrected $p = .032$, $BF_{10} = 1.82$, and trial window 9, uncorrected $p = .034$, $BF_{10} = 1.71$, which did not survive FDR-correction, all FDR-corrected $p = .464$. All other comparisons were not statistically different from zero, all uncorrected $p > .05$, all FDR-corrected $p \geq .464$, all $BF_{10} < 1$[5].

***Figure 6***

*Time Course of the Cumulative Recalibration Effect per Group*



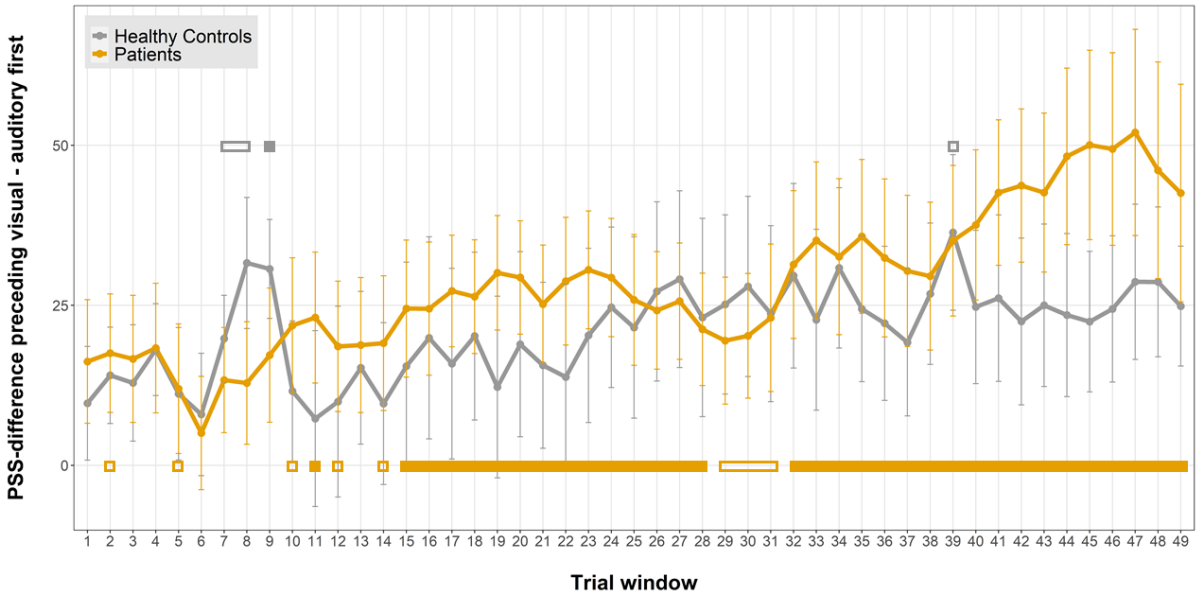

*Notes.* $n_{PAT} = 25$; $n_{HC} = 27$. Depicted is the PSS-difference between visual lead adapted minus auditory lead adapted test trials. The respective PSS was estimated by means of the mean of a Gaussian fit over the respective test trials in moving trial windows of 50 trials each, i.e. trials 1-50, 2-51 etc. *p*-values for each comparison are FDR-corrected. Error bars denote ±1 standard error of the mean.

[5] With the exception of trial window 13, which was not significant in the frequentist analysis, uncorrected $p = .058$, FDR-corrected $p = .464$, $BF_{10} = 1.13$.

**Time Course of Immediate Recalibration.** There were no statistically significant differences between groups in the PSS-difference between test trials preceded by a visual first vs. auditory first test trials, i.e. the immediate recalibration effect, in moving trials windows, all uncorrected $p > .05$, all FDR-corrected $p = .985$, all $BF_{10} < 1$. (see Figure 7 for the time course of the immediate temporal recalibration effect for both groups). For healthy controls, the PSS-difference was significantly different from zero in trial window 9, $t(26) = 3.97$, uncorrected $p = < .001$, FDR-corrected $p = .025$, $BF_{10} = 62.33$. Before, FDR-correction, the PSS-difference was also significantly different from zero in trial windows 4, 7, 8, 27, 34, 35, 38, 39, 40, 46, 47, 48, 49, i.e. in some earlier but mostly middle to late trial windows, all uncorrected $p < .05$. Out of these, only trial windows 7, 8, and 39 showed a non-significant trend for a difference to zero after FDR-correction, all three FDR-corrected $p = .086$, all $6 < BF_{10} < 10$; the rest did not survive FDR-correction, all FDR-corrected $p > .1$, all $1 < BF_{10} < 4$. All other comparisons were not significant, all uncorrected $p > .05$, all FDR-corrected $p > .1$, all $BF_{10} < 1$. For patients, the PSS-difference was significantly different from zero in trial windows 11, 15-28, 32-49, i.e. in some earlier but mostly middle and late trial windows, all uncorrected $p < .05$, all FDR-corrected $p < .05$, all $1 < BF_{10} < 36$. Trial windows 2, 4, 10, 12 ,14, 29, 30 and 31 showed a non-significant trend for a difference to zero after FDR-correction, all FDR-corrected $p < .1$, all $0.5 < BF_{10} < 1.5$. All other comparisons did not survive FDR-correction, all FDR-corrected $p > .1$, all $BF_{10} < 1$.

***Figure 7***

*Time Course of the Immediate Recalibration Effect per Group*




*Notes.* $n_{PAT} = 25$; $n_{HC} = 27$. Depicted is the PSS-difference between test trials preceded by a visual first minus by an auditory first test trial. The respective PSS was estimated by means of the mean of a Gaussian fit over the respective test trials in moving trial windows of 50 trials each, i.e. trials 1-50, 2-51 etc. Filled rectangles denote a significant difference from zero for the respective trial window per group, $p < .05$; hollow rectangles denote a non-significant trend for a difference from zero, $p < .1$. *p*-values for each comparison are FDR-corrected. Error bars denote ±1 standard error of the mean.

### *3.1.6 Correlations of Experimental Data*

There were no significant Spearman rank correlations between the general TBW, the cumulative recalibration effect (defined as the PSS-difference between the first 50 visual lead minus auditory lead adapted test trials) and the immediate recalibration effect (defined as the PSS-difference between test trials preceded by a visual first minus auditory first test trial), all $\rho < .3$, all $p > .05$, all $BF_{10} < 1$ (see Table 4 for Spearman rank coefficients $\rho$ over all participants. Exploratory Spearman rank coefficients $\rho$ split by group can be found in Table C.1 and Table C.2 in Appendix C).

***Table 4***

*Spearman Rank Correlation Coefficients ρ for Associations Between Experimental Data and Questionnaire Scores over all Participants*

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. |
|---|---|---|---|---|---|---|---|---|
| 1. TBW | | | | | | | | |
| 2. Cumulative Recalibration | .03 | | | | | | | |
| 3. Immediate Recalibration | .12 | .04 | | | | | | |
| 4. CAPS | -.08 | .08 | .07 | | | | | |
| 5. LSHS-E | -.12 | .03 | -.04 | **.87*** ** | | | | |
| 6. PCL | -.05 | .06 | .14 | **.75*** ** | **.70*** ** | | | |
| 7. Pass. Exp | .01 | -.11 | -.01 | **.79*** ** | **.78*** ** | **.68*** ** | | |
| 8. BCSS n.S. | -.16 | -.02 | .09 | **.66*** ** | **.66*** ** | **.58*** ** | **.56*** ** | |
| 9. BCSS n.O. | -.19 | -.12 | .13 | **.64*** ** | **.63*** ** | **.61*** ** | **.55*** ** | **.70*** ** |

*Notes.* $N = 52$. Spearman rank correlation coefficients $\rho > .3$ are marked in bold. *p*-values are Bonferroni-corrected. TBW = temporal binding window, defined as the bandwidth of the Gaussian fit over all test trials; Cumulative Recalibration = cumulative recalibration effect, defined as the difference in mean of the Gaussian fits, i.e. PSS, over the first 50 visual lead vs. auditory lead adapted test trials; Immediate Recalibration = immediate recalibration effect, defined as the difference in mean of the Gaussian fits, i.e. PSS, over test trials preceded by a visual first vs. auditory first test trial; CAPS = Cardiff Anomalous Perception Scale, total score; LSHS-E = Launey-Slade Hallucination Scale-Extended, total score; PCL = Paranoia Checklist, total score; Pass. Exp. = total score in passivity experiences items; BCSS n.S. = Brief Core Schema Scales, negative self subscale; BCSS n.O. = Brief Core Schema Scales, negative other subscale.

* $p < .05$, ** $p < .01$, *** $p < .001$

## 3.2 Control Variables and Questionnaire Data

### *3.2.1 Control Tests for Attention, Processing Speed, and Vigilance*

**TMT.** Groups differed significantly in completion time (in seconds) of the TMT-A version, with patients ($M = 32.79$, $SD = 13.21$) requiring more time to complete the test compared to healthy controls ($M = 23.12$, $SD = 8.90$), $t(41.6) = 3.07$, $p = .004$, Cohen's $d = 0.87$, $BF_{10} = 12.49$. Further, patients ($M = 78.12$, $SD = 30.34$) showed higher completion times in the TMT-B version compared to healthy controls ($M = 61.03$, $SD = 26.24$), $t(47.65) = 2.16$, $p = .036$, Cohen's $d = 0.6$, $BF_{10} = 1.88$.

**CPT-IP.**[6] A 2 (*Group*) * 3 (*Digit Condition*) mixed ANOVA on CPT-IP *d'* scores revealed a significant main effect of *Group*, $F(1, 50) = 9.03$, $p = .004$, $\eta^2_p = 0.15$, $BF_{incl} = 6.45$, indicating that patients ($M = 2.84$, $SD = 0.83$) showed overall lower d' scores compared to healthy controls ($M = 3.46$, $SD = 0.6$). There was a highly significant main effect of *Digit Condition*, $F(1.67, 83.48) = 129.57$, $p < .001$, $\eta^2_p = 0.72$, $BF_{incl} > 1000$. Bonferroni-corrected post-hoc comparisons revealed that d' scores were significantly larger in the 2-digit ($M = 3.57$, $SD = 0.65$) compared to the 3-digit ($M = 2.96$, $SD = 0.91$), $t(51) = 6.86$, $p < .001$, and the 4-digit condition ($M = 1.85$, $SD = 1.04$), $t(51) = 13.41$, $p < .001$, and larger in the 3-digit compared to the 4-digit condition, $t(51) = 10.81$, $p < .001$. There was no *Group* * *Digit Condition* interaction, $F(1.67, 83.48) = 0.53$, $p = .557$, $\eta^2_p = 0.01$, $BF_{incl} = 0.63$.

**Associations Between Experimental Data and Control Variables.** Bonferroni-corrected Spearman rank correlations revealed significant associations between the general TBW and both the TMT-B version completion time, $\rho = .40$, $p = .027$, $BF_{10} = 2.91$, and the CPT overall *d'* score, $\rho = -.52$, $p < .001$, $BF_{10} = 24.97$. There were no other significant associations between control tests and experimental data, all $p > .05$, all $BF_{10} < 1$[7] (see Table C.3 in Appendix C for Spearman rank correlation coefficients $\rho$).

### *3.2.2 Questionnaires*

**CAPS.** Groups differed in CAPS scores, with patients ($M = 11.6$, $SD = 5.8$) reporting a significantly higher number of anomalous perceptual experiences compared to healthy controls ($M = 1.4$, $SD = 2.0$), $t(29.21) = 8.4$, $p < .001$, Cohen's $d = 2.41$, $BF_{10} > 1000$.

**LSHS.** LSHS scores differed significantly between groups, with patients ($M = 24.3$, $SD = 11.9$) showing a higher occurrence of hallucinatory events compared to healthy controls ($M = 3.2$, $SD = 4.1$), $t(29.37) = 8.44$, $p < .001$, Cohen's $d = 2.41$, $BF_{10} > 1000$.

**PCL.** Groups differed significantly in PCL scores, indicating significantly higher levels of paranoia in patients ($M = 94.2$, $SD = 52.3$) compared to healthy controls ($M = 26.0$, $SD = 26.2$), $t(34.72) = 5.88$, $p <. 001$, Cohen's $d = 1.67$, $BF_{10} > 1000$.

[6] One patient did not complete the CPT-IP due to fatigue at the end of both sessions. Their CPT *d'* scores were replaced by the respective group mean to not lose the subject for further analysis.
[7] With the exception of TBW ~ TMT-A completion time, which was not significant in the frequentist analysis: Bonferroni-corrected $p = .216$, $BF_{10} = 1.71$

**Passivity Experiences.** Patients ($M$ = 7, $SD$ = 4.5) reported a significantly higher frequency of passivity experiences compared to healthy controls ($M$ = 0.1, $SD$ = 0.5), $t(24.45) = 7.55$, $p < .001$, Cohen's $d = 2.18$, $BF_{10} > 1000$.

**BCSS.** Patients ($M$ = 6.6, $SD$ = 4.5) showed a significantly higher level of negative schemata towards the self compared to healthy controls ($M$ = 2, $SD$ = 2.5), $t(36.68) = 4.5$, $p < .001$, Cohen's $d = 1.28$, $BF_{10} = 653.69$. Further, patients ($M$ = 9.8, $SD$ = 5.6) reported a significantly higher level of negative schemata towards others compared to healthy controls ($M$ = 3.5, $SD$ = 3.8), $t(42.09) = 4.77$, $p < .001$, Cohen's $d = 1.34$, $BF_{10} > 1000$.

**Correlations Between Questionnaire Scores.** CAPS, LSHS, PCL, passivity experiences and both BCSS negative subscales all showed highly significant, strong positive Spearman rank correlations with each other, all $\rho > .5$, all Bonferroni-corrected $p < .001$, all $BF_{10} > 30$[8] (see Table 4 for Spearman rank coefficients $\rho$. Exploratory Spearman rank coefficients $\rho$ split by group can be found in Table C.1 and Table C.2 in Appendix C).

### 3.3 Correlations Between Experimental Data and Questionnaire Scores

There was no significant Spearman rank correlation between the general TBW, the cumulative recalibration effect, the immediate recalibration effect and questionnaires scores, all $\rho < .3$, all Bonferroni-corrected $p > .05$, all $BF_{10} < 1$ (see Table 4 for Spearman rank coefficients $\rho$. Exploratory Spearman rank coefficients $\rho$ split by group can be found in Table C.1 and Table C.2 in Appendix C).

## 4 Discussion

In this study, we investigated temporal MSI and crossmodal temporal recalibration in psychosis. While considerable work on temporal MSI in patients with the diagnosis of a psychotic disorder has been published (Noel et al., 2018; Zhou et al., 2018; for a review and meta-analysis, see Zhou, Lai, et al., 2022; Zopf et al., 2021), to our knowledge no published study investigated cumulative and immediate crossmodal temporal recalibration in psychosis. Only one recent study provided initial findings on immediate temporal recalibration in extraordinarily rare EOS (Zhou, Cui, et al., 2022), but generalizability to adult-onset psychosis might be limited. Based on previous evidence reporting deficits in time perception and temporal processing in psychosis (Amadeo et al., 2022; Ciullo et al., 2016; Thoenes &

[8] The majority of $BF_{10}$ for correlations between questionnaire scores was > 1000, with following exceptions: LSHS ~ BCSS negative other $BF_{10} = 992.48$; PCL ~ BCSS negative self $BF_{10} = 288.32$; passivity experiences ~ BCSS negative self $BF_{10} = 30.95$; passivity experiences ~ BCSS negative other $BF_{10} = 136.02$.

Oberfeld, 2017; Wallace et al., 2020; Wallace & Stevenson, 2014; Zhou et al., 2018) as well as prior findings of impaired crossmodal temporal recalibration in ASD (Noel, De Niear, et al., 2017; Turi et al., 2016; Zhou, Cui, et al., 2022), a neurodevelopmental disorder sharing several aspects with psychosis including an increased TBW during temporal MSI (Zhou et al., 2018), we expected to observe altered crossmodal temporal recalibration in psychosis.

Patients with the diagnosis of a psychotic disorder and healthy controls concluded several phases of a SJ task with simple AV stimuli. Before each SJ phase, the same AV stimuli were presented with a fixed SOA, half visual lead and half auditory lead (paradigm adopted from Van der Burg et al., 2015). Further, all subjects completed neuropsychological tests for processing speed, attention, and vigilance as well as questionnaires for psychotic symptoms and negative schemata. We analyzed the SJ data of patients and healthy controls and compared the TBW over all SJ trials, the PSS-difference between visual lead adapted vs. auditory lead adapted SJ trials, i.e. the cumulative recalibration effect, and the PSS-difference in SJ trials immediately preceded by a visual first vs. auditory first SJ trial, i.e. the immediate recalibration effect, between both groups.

Patients and healthy controls did not differ in the TBW over all trials, the PSS-difference between visual lead adapted vs. auditory lead adapted SJ trials, and the PSS-difference in SJ trials with a preceding visual vs. auditory first SJ trial. This indicates that both groups showed comparable temporal MSI, cumulative temporal recalibration, and immediate temporal recalibration. In the following, these results will be discussed in more detail with respect to our hypotheses and previous findings on temporal multisensory processing in psychosis.

### 4.1 Hypothesis 1: Psychosis and the TBW

Our first aim was to replicate previous findings on an increased TBW in psychosis (Noel et al., 2018; Zhou et al., 2018; Zhou, Lai, et al., 2022; Zopf et al., 2021). Contrary to our expectation, we did not observe a difference in the bandwidth of the Gaussian fit over all SJ test trials between patients and healthy controls. While the mean bandwidth was descriptively slightly larger in patients than in controls, i.e. pointed in the expected direction, this difference was not significant with a small effect size. Further, our additional Bayesian analysis revealed a $BF < .333$, reflecting moderate evidence for the null hypothesis. This indicates that the general TBW did not differ between both groups, reflecting a similar distribution of positive SJ responses between patients and healthy controls. Thus, our sample

of patients with the diagnosis of a psychotic disorder did not show an enlarged TBW compared to controls.

Our findings are consistent with a study by de Boer-Schellekens et al. (2014), who observed impaired temporal acuity in patients compared to controls during a TOJ task in a visual only condition, but comparable TOJ performance between both groups when two sounds were presented in addition to the visual stimuli. This indicated that additional presentation of auditory cues ameliorated deficits in visual temporal processing, reflecting intact AV temporal integration (de Boer-Schellekens et al., 2014) and suggesting potential compensatory effects of crossmodal gains in psychosis. This is consistent with previous studies reporting intact multisensory facilitation in psychosis using AV target detection (Wynn et al., 2014) or AV far-near judgement paradigms (Stephen et al., 2013; Stone et al., 2011), suggesting a potential influence of task type or task demands on crossmodal gains in psychosis.

However, our findings are inconsistent with the majority of previous studies on multisensory temporal processing in psychosis. An enlarged multisensory TBW has repeatedly been observed in psychosis and it has been discussed that the enlargement of the TBW might be more pronounced in speech compared to simple AV stimuli (Noel et al., 2018; Zhou et al., 2018; Zhou, Lai, et al., 2022; Zopf et al., 2021). Several other studies investigated the TBW in visual, i.e. unimodal stimuli. The majority of those reported an enlarged visual TBW in psychosis (for a review and meta-analysis, see Zhou et al., 2018), with only one study failing to observe impaired visual temporal processing (Grimsen et al., 2013). Crucially, impairments in AV temporal processing were observed to go beyond unisensory processing deficits (Stevenson, Park, et al., 2017; Zhou, Lai, et al., 2022), highlighting the importance of investigating multisensory temporal processing in psychosis. In sum, while previous evidence indicated an impairment in visual and AV asynchrony detection in psychosis reflected by an enlarged TBW (Noel et al., 2018; Zhou et al., 2018; Zhou, Lai, et al., 2022; Zopf et al., 2021), our findings are in contrast to previous work.

An enlarged multisensory TBW has been discussed as a potential core deficit in psychosis (Amadeo et al., 2022). As mentioned above (see **1.3 Temporal Multisensory Processing in Psychosis**), an enlarged multisensory TBW might lead to erroneous integration of stimuli originating from distinct sources, facilitating anomalous and confusing percepts and potentially leading to psychotic symptoms (Zhou et al., 2018). A meta-analysis by Zhou et al.

(2018) on the uni- and multisensory TBW in psychosis reported a large effect size of the enlarged TBW, i.e. a Hedge's $g > 0.8$, in both unimodal visual and AV stimuli. However, only two out of the above mentioned multisensory studies could be included in the meta-analysis by Zhou et al. (2018). This implies that more studies and a more current meta-analytic approach are needed to obtain robust findings on a potentially large effect size of an enlarged multisensory TBW in psychosis and to gather experimental support for the hypothesis on a potential role of impaired temporal MSI in the development of psychosis (Amadeo et al., 2022; Zhou et al., 2018).

Importantly, the designs of previous studies on AV temporal processing in patients with the diagnosis of a psychotic disorder varied, including SJ (Foucher et al., 2007; Noel et al., 2018; Stevenson, Park, et al., 2017; Zhou, Lai, et al., 2022), TOJ (de Boer-Schellekens et al., 2014), double-flash illusion (Haß et al., 2017), stream-bounce illusion (Zvyagintsev et al., 2017) as well as McGurk paradigms (B. Martin et al., 2013) and – moreover – some investigated simple and others speech stimuli. Taken together, these differences in study designs might limit comparability between studies and it could be argued that different complexity of stimulus material and task might partially account for diverging results such as the findings from de Boer-Schellekens et al. (2014) and our study.

It is important to note that the general TBW, reflected by the mean bandwidth over all SJ trials, showed a high variability in both our patient and control groups, minimizing the possibility to statistically detect small differences. Crucially, this variability of the TBW was comparable between our patient and control samples, whereas previous studies reported a larger variability in patients with the diagnosis of a psychotic disorder compared to healthy controls (de Boer-Schellekens et al., 2014; Foucher et al., 2007; B. Martin et al., 2013; Noel et al., 2018; Stevenson, Park, et al., 2017; Zhou, Lai, et al., 2022). This further suggests that AV temporal processing was similar in our patient and control samples in contrast to previous studies.

It is possible that we failed to observe an enlarged TBW in simple AV stimuli in psychosis, since the TBW enlargement in psychosis is discussed to be larger in speech compared to simple stimuli (Zhou et al., 2018). Previous evidence suggested that - beyond the temporal domain - MSI deficits might be pronounced in psychosis in socially relevant information, i.e. stimuli with linguistic or emotional content, compared to simple cues (Tseng et al., 2015). However, an enlarged TBW in psychosis has repeatedly been observed using

simple AV stimuli (Foucher et al., 2007; Haß et al., 2017; Stevenson, Park, et al., 2017; Zhou, Lai, et al., 2022; Zvyagintsev et al., 2017), so it could be argued the findings in our patient sample might due to methodological differences or specific for our sample. Nevertheless, comparing our findings in our patient sample of intact temporal MSI with simple AV stimuli to an additional condition with AV speech stimuli could be informative for further hypothesis generation regarding factors, which potentially influence temporal MSI in simple stimuli vs. speech in psychosis.

Temporal processing impairments in psychosis have been observed both in SJ and TOJ tasks (Noel et al., 2018; Zhou et al., 2018; Zhou, Lai, et al., 2022; Zopf et al., 2021). However, Capa et al. (2014) reported that deficient temporal processing in patients with the diagnosis of a psychotic disorder was pronounced in a visual TOJ compared to a visual SJ tasks, suggesting a specific impairment in temporal ordering compared to asynchrony detection (Capa et al., 2014). It is possible that we failed to observe group differences in AV temporal processing since we administered a SJ rather than a TOJ task. However, most of previous studies on multisensory temporal processing in psychosis administered SJ tasks, indicating a deficit in multisensory asynchrony detection. Nevertheless, it might be possible that we would have observed significant group differences using a multisensory TOJ task. It has been discussed that TOJ tasks require additional cognitive resources compared to SJ tasks (García-Pérez & Alcalá-Quintana, 2012), which might be deficient in psychosis. Patients with the diagnosis of a psychotic have been observed to perceive stimuli individually rather than in sequence, i.e. fail to structure events relative to each other in time. It has been discussed that this impairment might root in a deficit in predicting future events while the present event is still in focus, leading to a disintegration of temporal information processing in psychosis (Amadeo et al., 2022). Thus, it might be possible that our patient sample might have shown impaired performance during a TOJ task, while the cognitive resources were intact enough for typical SJ performance. Future research should address this and directly compare multisensory SJ and TOJ tasks to investigate temporal MSI and crossmodal temporal recalibration (see **4.2 Hypothesis 2: Psychosis and Crossmodal Temporal Recalibration**) in psychosis and thus to study if specific impairments in temporal ordering of events might influence both aspects of multisensory temporal processing in psychosis.

We conclude that further research is needed to disentangle possible reasons for the contradictory findings of de Boer-Schellekens et al. (2014) and our study. Future studies should specifically examine where the potential effects of impaired temporal MSI in

psychosis might lie. Possible candidate factors might be found in e.g. stimulus and task complexity, such as specific impairments in multisensory speech processing (Zhou et al., 2018) or temporal ordering of sensory events (Capa et al., 2014). Further, investigating temporal MSI in patient samples with larger variability of cognitive impairments or symptom severity might be beneficial to elucidate the potential effects of cognitive abilities and symptom load, which has been suggested in previous (Foucher et al., 2007; Stevenson, Park, et al., 2017; Zhou, Lai, et al., 2022). Moreover, by directly comparing temporal MSI in samples on different stages of the psychosis spectrum such as subjects with psychosis proneness, first-episode patients, patients in an acute psychotic episode or patients with chronic psychosis, future studies might succeed in gaining further support for theoretical accounts suggesting a role of an enlarged multisensory TBW for the development of psychosis (Amadeo et al., 2022; Zhou et al., 2018).

## 4.2 Hypothesis 2: Psychosis and Crossmodal Temporal Recalibration

### *4.2.1 Discussion and Comparison to Previous Evidence*

Our second aim was to provide first experimental evidence on cumulative and immediate crossmodal temporal recalibration in psychosis. We expected to observe altered crossmodal recalibration in our patient sample, since psychosis has been reported to be associated with marked deficits in time perception and temporal processing (Amadeo et al., 2022; Ciullo et al., 2016; Thoenes & Oberfeld, 2017; Wallace et al., 2020; Wallace & Stevenson, 2014; Zhou et al., 2018). Further, previous evidence suggested deficient crossmodal temporal recalibration in ASD (Noel, De Niear, et al., 2017; Stevenson, Toulmin, et al., 2017; Turi et al., 2016; Zhou, Cui, et al., 2022), a neurodevelopmental disorder sharing deficits in temporal processing with psychosis (Zhou et al., 2018). We regarded these findings as hints for possibly deficient crossmodal temporal recalibration in psychosis.

We overall successfully replicated the findings of Van der Burg et al. (2015). In the first 50 SJ trials, the PSS in visual lead adapted SJ trials was shifted towards the visual lead adapted SOAs compared to the auditory lead adapted SJ trials. This PSS shift in direction of the respective SOA during preceding adaptation phases reflects the cumulative recalibration effect first reported by Fujisaki et al. (2004) and Vroomen et al. (2004), which has been successfully replicated since (e.g. Di Luca et al., 2009; Stevenson, Toulmin, et al., 2017; Uno & Yokosawa, 2022; Van der Burg et al., 2015). Further, the PSS in SJ trials with an immediately preceding visual first SJ trial was significantly shifted towards the previous trial's visual first SOA compared to SJ trials preceded by an auditory first SJ trial. This PSS

shift in direction of the immediately preceding SOA reflects the immediate temporal recalibration effect first reported by Van der Burg et al. (2013) and indicates rapid recalibration processes depending on the immediate stimulus history.

However, contrary to our expectation the effect of adaptation, i.e. if SJ trials were visual lead vs. auditory lead adapted, on the PSS over the first 50 SJ trials did not differ between groups, as reflected by a lack of a significant interaction. This indicates that the cumulative recalibration effect in the first 50 SJ trials following consistent AV adaptation did not differ between patients and healthy controls. Thus, our patient sample showed intact cumulative temporal recalibration in the first 50 SJ trials directly after prolonged adaptation, i.e. in the time frame in which the cumulative temporal recalibration effect has been observed to be largest (Van der Burg et al., 2015).

To our knowledge, our study is the first to investigate cumulative temporal recalibration in psychosis. We therefore compare our results to one previous study by Stevenson, Toulmin et al. (2017) on cumulative temporal recalibration and autistic traits. Healthy participants performed a SJ task with preceding visual lead vs. auditory lead adaptation, analogue to our design. The authors observed a significant association between a decrease in the cumulative recalibration effect and autistic traits, suggesting a link between symptoms of ASD and impaired cumulative temporal recalibration (Stevenson, Toulmin, et al., 2017). Given the scarcity of evidence regarding cumulative temporal recalibration in neurodevelopmental disorders such as psychosis and ASD, interpretations and conclusions should be drawn with caution. It is possible that future studies might confirm that ASD is associated with a deficit in cumulative temporal recalibration, i.e. an impaired ability to perceptually adjust to consistent asynchronies. Further, it also is possible that psychosis is not or weaker associated with impaired cumulative recalibration compared to ASD, as our findings might suggest. To date, further research on cumulative recalibration both in psychosis and ASD is urgently needed to confirm or reject this speculation of potentially different impairments in cumulative temporal recalibration in psychosis and ASD.

Our results on the immediate recalibration effect also contrast our expectations: the effect of modality order on the immediately preceding SJ trial on the PSS over all test trials did not differ between groups, as reflected by a lack of a significant interaction. This indicates that the immediate recalibration effect did not differ between patients and healthy controls. Thus, our findings suggest that patients showed intact immediate temporal recalibration over

all SJ trials, i.e. recalibrated their SJ responses depending on the immediate stimulus history similarly as healthy controls.

Our results of an unimpaired immediate temporal recalibration effect are consistent with initial results of intact immediate recalibration in children and adolescents with the diagnosis of EOS (Zhou, Cui, et al., 2022). Crucially, the transferability of the findings of Zhou, Cui et al. (2022) to adult patients with the diagnosis psychosis might be limited due to marked differences in premorbid cognitive impairment as well as symptom severity between early- and adult-onset psychosis (Vyas et al., 2011) and potential age-related confounds in temporal recalibration processes (Noel et al., 2016; Zhou et al., 2020) in the study of Zhou, Cui, et al. (2022). Nonetheless, further support might be found in an earlier study by Zhou et al. (2020). In their study, the authors investigated immediate temporal recalibration during a SJ task in healthy participants and expected associations between a reduction of the immediate recalibration effect and increased schizotypal and autistic traits. However, Zhou et al. (2020) failed to observe the expected correlation to schizotypal traits, suggesting that there might be no direct link between subclinical psychotic symptoms and immediate AV temporal recalibration. It could be argued that this finding of Zhou et al. (2020) matches our results of unimpaired immediate recalibration in our patient sample. In sum, previous findings in healthy individuals with schizotypal traits and in minors with the diagnosis of early-onset psychosis could - along with our initial findings in adult patients with the diagnosis of a psychotic disorder - indicate that psychosis might not be associated with impaired immediate recalibration, suggesting that patients with the diagnosis of a psychotic disorder might be able to rapidly recalibrate to AV asynchronies in the immediate stimulus history.

However, our findings in psychosis are in contrast with studies on immediate temporal recalibration in ASD. Turi et al. (2016) investigated immediate temporal recalibration in adults with the diagnosis of an ASD and observed that patients hardly recalibrated their SJ responses to AV asynchronies in the immediate stimulus history compared to controls. This reduction in immediate temporal recalibration in ASD was replicated in children and adolescents with the diagnosis of an ASD by Noel, De Niear et al. (2017) and Zhou, Cui, et al. (2022), suggesting a deficit to rapidly recalibrate to preceding AV asynchronies in the developmental trajectory of ASD. Interestingly, while Noel, De Niear et al. (2017) observed an intact immediate recalibration effect in speech and deficits in simple and complex non-speech stimuli, Zhou, Cui, et al. (2022) observed the opposite pattern, i.e. deficits in speech but not in simple stimuli. Thus, the findings of Zhou, Cui. et al. (2022)

matched the pattern of previous studies on pronounced temporal processing deficits in speech in ASD (e.g. Stevenson et al., 2014), while the study of Noel, De Niear et al. (2017) contrasted them. As mentioned earlier, deficits in multisensory processing in psychosis have been discussed to be pronounced in speech compared to simple stimuli (Tseng et al., 2015; Zhou et al., 2018). While our study provides initial results suggesting intact immediate recalibration in psychosis, it is possible that a different pattern might emerge when investigating immediate recalibration in psychosis using speech compared to simple stimuli.

In sum, our study provides first evidence on cumulative and immediate crossmodal temporal recalibration. While our results suggest intact crossmodal recalibration in patients with the diagnosis of a psychotic disorder, future studies should aim to replicate our findings with larger samples to investigate if our findings might be e.g. specific for our patient sample, which potentially showed relatively spared temporal processing capabilities.

### *4.2.2 Deficits in Perceptual Prediction as an Underlying Mechanism*

Stevenson, Toulmin et al. (2017) argued that a possible cumulative temporal recalibration deficit in individuals with autistic traits might be rooted in perceptual inference dysfunctions, which have been previously reported in ASD (for reviews, see Pellicano & Burr, 2012; Sinha et al., 2014). Stevenson, Toulmin et al. (2017) proposed that decreased cumulative temporal recalibration to statistically regular AV asynchronies reflects an impaired ability to weigh priors, i.e. mental representations or models of the environment acquired and updated through experience, in relation to incoming sensory input during perceptual inference. This might decrease the impact of repeatedly presented AV asynchronies on perception and inhibit cumulative crossmodal recalibration (Stevenson, Toulmin, et al., 2017). Further, both Noel, De Niear et al. (2017) and Zhou, Cui, et al. (2022) proposed an analogue argumentation for impaired immediate temporal recalibration in ASD and argued that a deficit in weighing perceptual priors and thus an overweighing of sensory input, i.e. the currently perceived AV asynchrony, leads to insufficient recalibration processes to AV asynchronies in the immediate stimulus history (Noel, De Niear, et al., 2017; Zhou, Cui, et al., 2022).

Importantly, dysfunctions in perceptual inference, as described by the predictive coding framework, have also been reported in psychosis and it has been proposed that altered weighing of priors and sensory input might be a core deficit in psychosis, potentially even contributing to the development of psychotic symptoms (for reviews, see Heinz et al., 2019;

Sterzer et al., 2018). While ASD has been discussed to show specific impairments in weighing of priors (Pellicano & Burr, 2012), psychosis might be associated with altered weighing of both priors and sensory input (Sterzer et al., 2018). On the one hand, previous studies reported impaired weighing of priors in psychosis (e.g. Schmack et al., 2013, 2015; Stuke et al., 2018), leading to a weakened impact of prior knowledge on perception and volatile belief updating. This incites the speculation that psychosis could be associated with decreased crossmodal temporal recalibration, i.e. analogue to mechanisms in ASD as proposed by Noel, De Niear et al. (2017), Stevenson, Toulmin, et al. (2017) and Zhou, Cui, et al. (2022). On the other hand, some studies reported overweighing of priors in psychosis, shaping perception based on overly fixed prior beliefs (e.g. Davies et al., 2018; Limongi et al., 2018; Powers et al., 2017; Schmack et al., 2017; Teufel et al., 2015). This encourages the question, if overweighing of priors in psychosis might be associated with an increased impact of statistically regular AV asynchronies or asynchronies in the immediate stimulus history on recalibration processes, potentially resulting in excessive crossmodal temporal recalibration. However, our findings do not offer support for either hypothesis, with our patient sample showing intact cumulative and immediate temporal recalibration.

Importantly, Noel et al. (2018) revealed potential differences between psychosis and ASD regarding underlying inference mechanisms of altered multisensory temporal processing. Using a causal inference modeling approach, Noel et al. (2018) showed that an enlarged TBW might be rooted in changes in binding priors in ASD, whereas the enlarged TBW in psychosis might be rather a consequence of changes in both binding priors and sensory input. This suggested that although both disorders show comparable phenomenological manifestations of altered temporal MSI, i.e. an enlarged TBW, the underlying mechanism might be different (Noel et al., 2018). Thus, future studies on perceptual inference mechanisms in both temporal MSI and crossmodal temporal recalibration in psychosis and ASD might differentiate the underlying mechanisms in both disorders. This could contribute to answering the question if psychosis and disorder might be differently impaired in multisensory temporal processes, as our data on crossmodal temporal recalibration in psychosis in contrast to previous findings in ASD might suggest.

Crucially, perceptual inference impairments as described by the predictive coding account of psychosis are discussed to be influenced by a multitude of factors such as the hierarchical level of inference processes or disorder progression status, implying a complex framework for understanding psychosis in need of further research (Sterzer et al., 2018). This

suggests that more studies on crossmodal temporal recalibration in psychosis are needed, targeting diverse cognitive hierarchical levels, e.g. by using differently complex stimuli, or recruiting patients in various stages of disorder progression, such as first-episode psychosis, chronic psychosis or patients in an acute vs. non-acute psychotic phase, and by this investigating crossmodal temporal recalibration in psychosis and its potential relations to dysfunctions in perceptual inference mechanisms.

### *4.2.3 Larger Bias Towards Visual Lead Asynchronies in Psychosis*

In both the analyses for the cumulative and immediate recalibration effect, respectively, a main effect of group indicated that the PSS differed between patients and healthy controls. This reflected a shift towards visual leading SOAs in the patient compared to the healthy control group. While a bias to perceive AV stimuli more often as synchronous in visual leading compared to auditory leading stimuli has been reported in healthy participants (Vroomen & Keetels, 2010), our findings suggest that this bias towards visual leading SOAs might be larger in our patient sample compared to healthy controls. It could be argued that this might indicate that the TBW in patients might have been not entirely similar compared to our control group. It could be possible that we did not find TBW differences (see **4.1 Hypothesis 1: Psychosis and the TBW**) because we did not differentiate between auditory and visual leading stimuli but averaged over both. Potentially, the enlargement of the TBW in psychosis might not be symmetrical, but might possibly be primarily driven by a larger bias towards visual lead compared to auditory lead stimuli, i.e. a skewed TBW enlargement favoring visual leading stimulus pairs. This might facilitate the perception of asynchronous AV stimuli as synchronous, especially when the visual stimulus is first. To our knowledge however, no published study so far has compared PSS biases in psychosis in visual vs. auditory lead AV stimuli during SJ. This highlights the importance for future studies to not only investigate the TBW on multisensory temporal processing in psychosis, but also possible PSS shifts towards visual leading stimulus pairs.

### *4.2.4 Time Course of Cumulative and Immediate Recalibration*

Additionally, we ran an exploratory analysis for the time course of both the cumulative and the immediate recalibration effect and compared the time course of both effects between groups as well as within each group against zero. There was no statistically significant difference between patients and healthy controls in any trial window for neither effect, indicating that the time course of both the cumulative and the immediate temporal recalibration effect did not differ between groups.

Regarding the time course of cumulative recalibration, it descriptively seemed that the cumulative recalibration effect for both groups was largest directly after prolonged adaptation and decayed over time, which would be consistent with the findings of Van der Burg et al. (2015). However, FDR-corrected comparisons against zero separately for both groups did not reach statistical significance in any time window. It is likely that the large variability in the cumulative recalibration effect per trial window (see Figure 6) within each group limited the possibility of observing significant differences against zero within in each group.

For the time course of immediate recalibration, the effect seemed rather stable in the control group, which would be consistent with the findings of Van der Burg et al. (2015), but it descriptively seemed to increase over time in the patient group. FDR-corrected comparisons against zero separately for healthy controls indicated (trends for) differences against zero in two early and one later trial window. For the patient group, however, most of the trial windows showed FDR-corrected (trends for) differences against zero, which were especially observable in middle to late trial windows. This supported the descriptive impression of an increase of the immediate recalibration effect over time in our patient sample.

Seen from a perceptual inference perspective, this impression of a potential increase of the immediate recalibration effect in our patient sample allows the question, if psychosis might be associated with specific changes in weighing prior knowledge and sensory input over time during immediate recalibration. Dysfunctions in perceptual inference mechanisms, as described by the predictive coding account, have been reported in psychosis (Sterzer et al., 2018) and it has been discussed that psychosis proneness might be associated with a decreased binding tendency, i.e. a decreased prior to integrate crossmodal stimuli (Odegaard & Shams, 2017). To our knowledge, however, the time course of perceptual inference mechanisms in psychosis has been scarcely studied. It might be interesting to investigate if psychosis is associated with altered learning mechanisms during perceptual inference, e.g. if patients require more time, i.e. more stimulus exposure, to form or update priors or if weighing prior knowledge and sensory input might differently change over time in psychosis. Support for altered learning mechanism might be found in a study by Powers et al. (2017), who used a modeling approach to investigate perceptual inference mechanisms in AV association learning during AV stimulus detection. Powers et al. (2017) observed a decreased volatility parameter during AV association learning, suggesting that perceptual priors in AV stimulus processing might be resistant to updating in psychosis. Future studies should aim to

investigate if this resistance of perceptual prior updating in psychosis might be associated with altered time courses of multisensory temporal processing.

In sum, our exploratory analysis of the time courses of crossmodal temporal recalibration partly replicated the findings of Van der Burg et al. (2015), who reported a decrease of the cumulative recalibration effect after adaptation phases but a rather stable immediate recalibration effect. However, our data might give first hints for a potential difference in time course of immediate temporal recalibration in psychosis. Future research should specifically investigate the time courses of cumulative and immediate temporal recalibration in psychosis and study if this is associated with changes over time in the accumulation of prior knowledge or in weighing of priors and sensory input during inference mechanisms.

### *4.2.5 A Theoretical Approach for Altered Temporal Recalibration as a Mechanism Underlying Psychotic Symptoms*

To our knowledge, only one review aimed to discuss a potential link between altered temporal recalibration mechanisms and the development of psychotic symptoms. In their review, Riemer (2018) argued that altered sensorimotor temporal recalibration might form a potential basis for the emergence of delusions of control, i.e. the conviction that one's own actions are controlled by external forces (Sass & Parnas, 2003). In healthy subjects, it has been discussed that the conscious intention of an action follows the perception of the action itself (with the intention being unconscious up to that time point). The sense that the action was self-caused is then inferred by internal temporal recalibration processes, which subjectively contract the timings of the conscious intention and the sensory perception of the action, leading to a temporal order of intention and action making sense to us (Haggard et al., 2002; Wenke & Haggard, 2009). Riemer (2018) proposed that impaired temporal recalibration of the perception of own actions and their conscious intention is a potential key mechanism behind delusions of control in psychosis. They argue that a deficit in temporal shifting the conscious intention before the perceived action results in a more veridical temporal perception of the time structure of behavior and intention, which however is confusing and does not make sense (Riemer, 2018). This might facilitate the feeling that the action was not self-intended, leading to the conclusion that external forces must have caused the behavior.

While the proposition of Riemer (2018) discussed a connection between altered crossmodal temporal recalibration and the development of psychotic symptoms, it only described a specific mechanism of altered sensorimotor temporal recalibration in the development of a specific symptom of psychosis, namely delusions of control, and thus offers no generalized framework. To our knowledge, no theoretical approach trying to connect sensory-sensory temporal recalibration and the development of psychotic symptoms exists to date. This lack of a theoretical account might lead to poor hypothesis generation and testing, possibly resulting in a rather result-driven study approach failing to connect experimental findings with phenomenological constructs of psychosis. We argue that a crucial next step is trying to connect existing work on multisensory temporal processing deficits in psychosis with disorder models of psychosis, such as investigating potential links between impaired multisensory temporal processing and the experience of anomalous percepts (see **4.2 Anomalous Perceptual Experiences as a Potential Link between Multisensory Processing and Psychosis** in the **General Introduction**) or altered inference mechanisms (see **5.1 Causal Inference** in the **General Discussion**). Such attempts of linking previous findings with concepts of disorder models might be beneficial for generating hypotheses on a potential key role of altered multisensory temporal integration and (potentially) crossmodal temporal recalibration for psychosis.

### 4.3 Hypothesis 3: Associations between Temporal MSI, Temporal Recalibration and Psychotic Symptoms

Finally, we aimed to replicate previous findings of correlations between multisensory temporal processing and psychotic symptoms (Dalal et al., 2021; Di Cosmo et al., 2021; Ferri et al., 2017, 2018; Foucher et al., 2007; Marsicano et al., 2022; Stevenson, Park, et al., 2017) and further extend these findings to crossmodal temporal recalibration. We expected associations between the general TBW, the cumulative as well as the immediate recalibration effect and self-reports on psychotic symptoms and negative schemata.

Contrary to our expectation, we did not observe any significant correlation between experimental and questionnaire data. Further, no Spearman rank correlation coefficient $\rho$ between the TBW and questionnaire scores approached our threshold of interest, i.e. $\rho > .3$. This indicated that the general TBW, the cumulative as well as the immediate recalibration effect showed no associations to anomalous perceptual experiences, hallucinations, paranoia, passivity experiences and negative schemata about the self and others, suggesting that

temporal MSI as well as crossmodal temporal recalibration might not be directly linked to psychotic symptoms.

### *4.3.1 Associations between the TBW and Psychotic Symptoms*

Our findings are inconsistent with previous work on the multisensory TBW in psychosis reporting significant associations between an enlarged TBW and psychotic symptoms. Stevenson, Park, et al. (2017) observed that an enlarged AV TBW predicted hallucination severity in patients with the diagnosis of a psychotic disorder. Foucher et al., (2007) observed an association between the AV TBW and symptoms of disorganization. Further, associations between an enlarged TBW and higher schizotypy as well as symptoms of hallucinations and disorganization in psychosis proneness have been reported, suggesting a potential endophenotype (Dalal et al., 2021; Di Cosmo et al., 2021; Ferri et al., 2016, 2017, 2018; Marsicano et al., 2022; Zhou, Cui, et al., 2022).

Given the lack of differences in the TBW between our patient and control samples, it is possible that we failed to observe correlations because the symptom load and in turn perceptual-cognitive impairments of our outpatient sample might have been too low to observe an enlarged TBW, which might have reduced the variance needed to detect statistically significant correlations between the TBW and psychotic symptoms. However, our patient sample reported significantly higher scores than healthy controls in all questionnaires, therefore we argue that this minimizes that possibility.

However, our findings are consistent with studies reporting no associations between the AV TBW and psychotic symptoms (Haß et al., 2017; B. Martin et al., 2013; Zhou, Lai, et al., 2022) or schizotypal traits in healthy subjects (Muller et al., 2020; Zhou et al., 2020). Further, several studies on the TBW in visual, i.e. unimodal stimuli also reported no associations between the TBW and psychotic symptoms (Capa et al., 2014; de Boer-Schellekens et al., 2014; Giersch et al., 2009; Grimsen et al., 2013; Tenckhoff et al., 2002). This suggests that an enlarged TBW might not be directly related to psychotic symptoms, especially in conditions where only visual stimuli are presented. This is however not consistent with the findings of a meta-analysis by Ueda et al. (2018) reporting a significant correlation between psychotic symptoms and altered temporal processing, which the authors argue might be rooted in timing acceleration (Ueda et al., 2018). Crucially, only four studies could be included in the meta-analysis of Ueda et al. (2018), we thus argue that their findings should be interpreted with caution. Further research with larger sample sizes for robust

correlational testing is needed to investigate specific association patterns between psychotic symptoms and uni- and multisensory temporal processing and to gather empirical support for accounts of a potential role of altered temporal processing for disorder development (e.g. Amadeo et al., 2022; Andreasen et al., 1999; Thoenes & Oberfeld, 2017).

### *4.3.2 Associations of Crossmodal Temporal Recalibration and Psychotic Symptoms*

Our findings regarding a lack of associations between crossmodal temporal recalibration and psychotic symptoms are consistent with a study in healthy subjects by Zhou et al. (2020). Administering a SJ paradigm, Zhou et al. (2020) expected significant associations between a reduced immediate recalibration effect and schizotypal and autistic traits, but their results failed to support their expectation. These and our findings in patients with the diagnosis of a psychotic disorder might suggest that (immediate) temporal recalibration is not directly associated with psychotic symptoms, but more studies corroborating this are needed. To our knowledge, the study by Zhou et al. (2020) is the only published study investigating associations between psychotic symptoms/schizotypy and crossmodal temporal recalibration. Reviewing findings of crossmodal temporal recalibration in ASD/samples with autistic traits could possibly give hints for future hypothesis generation in psychosis.

Noel et al. (2017) reported reduced immediate recalibration in ASD, suggesting a link between ASD and recalibration processes. However, the authors did not investigate correlations between the size of the immediate recalibration effect and symptoms of ASD (Noel, De Niear, et al., 2017). Both Turi et al. (2016) and Zhou, Cui, et al. (2022) reported correlations between impaired immediate temporal recalibration and increased autistic symptoms, which however is not consistent with findings of Zhou et al. (2020), who failed to observe associations between autistic traits in healthy individuals and immediate recalibration. Moreover, Stevenson, Toulmin et al. (2017) reported significant correlations between increased autistic traits and a reduced cumulative recalibration effect in a healthy student sample.

In sum, while prior evidence suggested a link between crossmodal temporal recalibration and symptoms of ASD (Noel, De Niear, et al., 2017; Stevenson, Toulmin, et al., 2017; Turi et al., 2016; Zhou, Cui, et al., 2022), our findings did not reveal associations between psychotic symptoms and both cumulative and immediate crossmodal temporal recalibration. It is possible that impaired crossmodal temporal recalibration shows specific

associations with autistic symptoms but not so much to psychotic symptoms. Given the to-date scarcity of findings on crossmodal temporal recalibration in neurodevelopmental disorders – especially in psychosis – and at least party contradictory findings on associations between ASD/autistic traits and immediate temporal recalibration (Turi et al., 2016; Zhou, Cui, et al., 2022; Zhou et al., 2020), further research on both cumulative and immediate temporal recalibration in psychosis and ASD is needed to clarify if there are specific links between crossmodal temporal recalibration and symptoms of ASD, psychosis or even both.

### 4.4 Limitations and Implications for Future Research

In addition to potential study limitations already mentioned earlier, i.e. the recruitment of a potentially rather functional outpatient sample compared to previous studies, small sample sizes for correlation analyses, and a large variability in experimental data in both groups, we propose the following possible limitations and implications for further research.

Our patient and control samples differed significantly in education, with fewer patients reporting higher education levels compared to controls. This is not surprising, since previous studies reported that individuals with the diagnosis of a psychotic disorder were less likely to achieve higher education levels (for a review and meta-analysis, see Dickson et al., 2020). While recent findings suggested that multisensory processes might predict the development of cognitive abilities in school children likely via benefits of multisensory learning environments (Denervaud et al., 2020), how the highest achieved level of education in adults influences multisensory processing in an experimental setting might be debatable – a question, which requires future experimental investigation. Importantly, it can be argued that this group difference in education did not impact our findings, since both groups showed similar levels of temporal multisensory processing.

Our patient sample showed lower processing speed and vigilance compared to healthy controls, as indicated by group differences in TMT and CPT-IP. Importantly, while vigilance levels were overall lower in patients compared to controls, it did not decrease faster than in healthy controls. When evaluating the potential influence of TMT and CPT-IP performance on our behavioural data, we observed correlations between TMT-B completion times, CPT-IP overall *d'* score and the general TBW, indicating that longer TMT-B completion time and lower CPT-IP d' over *d'* scores were associated with a larger TBW. This link between processing speed, vigilance and the TBW is consistent with reports of impaired processing speed and vigilance (Gebreegziabhere et al., 2022; Mesholam-Gately et al., 2009) and of an

enlarged TBW (Noel et al., 2018; Zhou et al., 2018; Zhou, Lai, et al., 2022; Zopf et al., 2021) in psychosis, suggesting that psychosis could offer a link between our observed association between TMT, CPT-IP and the general TBW. In contrast, we did not observe any association of TMT and CTP-IP to our measures for crossmodal temporal recalibration. Crucially, since no group differences in the general TBW and crossmodal temporal recalibration could be observed in our study, it can be argued that group differences in TMT and CPT-IP might have had limited impact on our behavioural findings.

During adaptation phases, we presented AV stimuli with a fixed SOA of 200ms, which can be assumed to lie within the TBW of both the patient and control group. Presenting AV stimuli with SOAs larger than the TBW decreases the likelihood of perceiving the stimuli as belonging to the same event (Wallace et al., 2020; Wallace & Stevenson, 2014), potentially reducing cumulative recalibration processes (Vroomen et al., 2004). As reported earlier, an enlarged multisensory TBW in psychosis might lead to erroneous integration of AV stimuli stemming from distinct events (Zhou et al., 2018). It might be interesting to present AV stimuli during adaptation phases with an SOA inside the enlarged TBW reported in psychosis but outside the typical TBW of healthy controls and to investigate if this impacts cumulative temporal recalibration differently in patients and healthy controls. In such an experimental design, it could be a plausible speculation that patients with the diagnosis of a psychotic disorder might show a cumulative temporal recalibration effect whereas healthy control might not. Future studies should investigate the hypothesis if psychosis is associated with maladaptive cumulative temporal recalibration in situations in which cumulative recalibration should not occur.

## 4.5 Conclusion

In this study on temporal MSI and crossmodal temporal recalibration in psychosis, we were unable to replicate previous findings of an enlarged TBW in psychosis (Noel et al., 2018; Zhou et al., 2018; Zhou, Lai, et al., 2022; Zopf et al., 2021). Further, our study is the first to investigate cumulative and immediate crossmodal temporal recalibration in psychosis. Our findings indicate both intact cumulative and immediate AV recalibration and contrast prior findings of impaired crossmodal recalibration in subjects with autistic traits and in ASD (Noel, De Niear, et al., 2017; Stevenson, Toulmin, et al., 2017), a neurodevelopmental disorder sharing deficits in temporal MSI with psychosis (Zhou et al., 2018). Finally, we did not observe associations between temporal MSI as well as crossmodal temporal recalibration and psychotic symptoms, providing no support for accounts of a potential role of impaired

temporal multisensory processing for the development of psychotic symptoms (e.g. Amadeo et al., 2022; Andreasen et al., 1999; Thoenes & Oberfeld, 2017). Further research on crossmodal temporal recalibration both in psychosis and ASD is needed to investigate a) if our findings of intact crossmodal recalibration in psychosis can be replicated, b) if factors such as symptom severity, cognitive impairments or task complexity influence temporal multisensory processing in psychosis, and c) if patterns of impaired crossmodal temporal recalibration potentially differ between psychosis and ASD, which would offer support for previous findings on different underlying mechanisms for impaired multisensory temporal processing in psychosis and ASD (Noel et al., 2018).

# Chapter IV: Spatial Multisensory Integration & Crossmodal Spatial Recalibration in Psychosis (Study 3)

# 1 Introduction

## 1.1 Spatial Multisensory Processing

In everyday life, a multitude of sensory information originating from various spatial positions needs to be continuously processed. Spatial relations between crossmodal stimuli, such as the position of a ball bouncing at the pavement and the accompanying sound, are crucial determinants for multisensory processing. Spatially close stimuli are more likely to perceived as originating from the same source than spatially distant stimuli (Meredith & Stein, 1986; Spence, 2013). Spatial MSI describes the integration of two or more crossmodal sensory signals alongside spatial dimensions, which often appear spatially disparate to a certain extent due to e.g. differences in spatial resolution of the respective sensory organs (Alais & Burr, 2004).

A well-known effect discussed to reflect spatial MSI is the VE. It describes the spatial capture of the perceived location of a sound by a spatially disparate visual stimulus (Alais & Burr, 2004; Radeau & Bertelson, 1974). It has been discussed that the higher spatial resolution of the visual system compared to that of auditory system might be a central characteristic leading to the sound location capture by the visual stimulus (Alais & Burr, 2004). Further, it has been reported that this crossmodal capture process depends on stimulus reliability of the respective stimuli: If the reliability of the sound increases in relation to the reliability of the visual stimulus, the extent of AV integration decreases. The VE has been even found to reverse when the reliability of the sound drastically exceeded the visual reliability, leading to a visual stimulus capture by the sound (Alais & Burr, 2004). While the VE has been investigated most in AV stimuli, spatial integration has also been reported in other crossmodal stimulus combinations such as audio-tactile or visuo-tactile stimuli (e.g. Bruns & Röder, 2010; Caclin et al., 2002; Samad & Shams, 2016).

Analogue to multisensory processes in the temporal domain (see **1.1 Temporal Multisensory Processing** in **Chapter III**), exposure to a crossmodal spatial conflict does not only induce "on-line" integration, but can also lead to crossmodal spatial recalibration. A prominent example is the so called cumulative ventriloquist aftereffect (VAE), which describes a shift in unisensory sound localization induced by previous, prolonged exposure to a consistent AV spatial conflict. This prolonged exposure leads to a shift in unisensory sound localization in the direction of the visual stimulus while the visual stimulus is no longer present (for a review, see Chen & Vroomen, 2013; Radeau & Bertelson, 1974). This cumulative crossmodal recalibration is thought to reflect adjustment processes of auditory

cortical maps to shifts in spatial correspondences between crossmodal signals. These shifts can occur as consequences of e.g. changes in the environment such as stepping outside of a building or changes in spatial relations between sensory organs due to bodily growth. As such, it has been argued that crossmodal recalibration reflects the aim of the perceptual system to reduce the discrepancy between the senses and maintain a coherent percept (Chen & Vroomen, 2013). The cumulative VAE has been studied extensively in AV (Chen & Vroomen, 2013), but also e.g. in audio-tactile (Bruns, Liebnau, et al., 2011; Bruns, Spence, et al., 2011) or visuo-tactile stimuli (Samad & Shams, 2018). Apart from the cumulative spatial recalibration effect induced by repeated exposure to an crossmodal spatial conflict, a rapid crossmodal spatial recalibration effect after a single exposure to an AV spatial conflict has also been observed, the so called immediate VAE (Wozny & Shams, 2011). This finding indicated that the perceptual system is able to rapidly adapt to crossmodal spatial discrepancies and suggests a high plasticity of the perceptual system (Wozny & Shams, 2011). Crucially, Bruns & Röder (2015) observed cumulative and immediate spatial recalibration in the same experiment, with both effects occurring on partly different time scales. Based on their findings, Bruns & Röder (2015) proposed that cumulative and immediate recalibration reflect distinct underlying processes with both effects influencing sound localization.

Previous evidence has suggested that although spatial MSI and crossmodal spatial recalibration are separate processes, they might partly be linked to each other. As mentioned earlier (see **1.1 Temporal Multisensory Processing** in **Chapter III**), recent studies have indicated that spatial MSI might be associated with immediate spatial recalibration with both processes being based on common underlying neural substrates, whereas cumulative spatial recalibration might be a relatively distinct process (Bruns et al., 2022; Park & Kayser, 2019; Rohlf et al., 2020). Further, recent evidence has suggested that both spatial MSI and immediate spatial recalibration develop earlier than cumulative spatial recalibration, supporting the notion of a relative dissociation of spatial MSI and cumulative spatial recalibration (Rohlf et al., 2020).

In sum, the perceptual system continuously processes crossmodal information and utilizes the spatial relations of crossmodal cues to infer they belong to a unitary event and to recalibrate for crossmodal spatial conflicts (Chen & Vroomen, 2013; Spence, 2013). Analogue to the previously mentioned proposition in the temporal domain (see **1.1 Temporal Multisensory Processing** in **Chapter III**), it could be argued that erroneous spatial

integration and/or recalibration of crossmodal stimuli might impact the creation of holistic and stable representations of the environment and the self and is potentially associated with disintegration of sensory information in psychosis (Postmes et al., 2014; Uhlhaas & Mishara, 2006).

### 1.2 Psychosis and Spatial Multisensory Integration

Psychosis has been found to be associated with unisensory processing deficits such as impaired visuospatial perception (Hardoy et al., 2004) or deficient auditory pitch detection and sound localization (Gold et al., 2012; Perrin et al., 2010; Sardari et al., 2022). This raises the question if impaired unisensory spatial processing might impair multisensory spatial processing in psychosis. In the temporal domain, deficits in AV MSI in psychosis have been reported to be not entirely explainable by auditory or visual processing deficits, suggesting specific deficits in multisensory compared to unisensory processing in psychosis (Stevenson, Park, et al., 2017). To date, it however remains unclear if and how unisensory processing deficits reported in patients with the diagnosis of a psychotic disorder (Carter et al., 2017; Gold et al., 2012; Hardoy et al., 2004; Javitt, 2009a; Javitt & Freedman, 2015) might influence spatial MSI in psychosis.

While there has been a considerable amount of studies on multisensory processing of social, i.e. linguistic and emotional (Lin et al., 2020; Tseng et al., 2015) and temporal (Zhou et al., 2018) information in psychosis, spatial multisensory processing has been scarcely investigated in patients with the diagnosis of a psychotic disorder.

An earlier study by de Gelder et al. (2003) investigated spatial MSI in patients with the diagnosis of a psychotic disorder via a ventriloquist paradigm with short sequences of simple AV stimuli. In this first experiment of the study of de Gelder et al. (2003), auditory localization responses were similarly captured by concurrently presented visual stimuli in patients compared to healthy controls, suggesting an intact VE. In contrast, the second experiment in the study of de Gelder et al. (2003) indicated significantly reduced integration of AV speech stimuli in patients compared to controls in a McGurk (McGurk & MacDonald, 1976) paradigm. de Gelder et al. (2003) proposed that MSI of linguistic information is impaired in psychosis, but basic spatial MSI might be intact. This proposition of intact spatial MSI in psychosis by de Gelder et al. (2003) is supported by three further studies using an AV far-near judgement paradigm, which reported similar crossmodal facilitation in patients with the diagnosis of a psychotic disorder compared to healthy controls. Importantly, unisensory

performance in visual and auditory trials was impaired in patients compared to controls, suggesting that crossmodal stimulus presentation might ameliorate unisensory processing deficits (Stephen et al., 2013; Stone et al., 2011, 2014).

A recent study by Noel et al. (2020) investigated visuo-tactile spatial integration in psychosis and ASD. Both clinical groups showed similar tactile localization performance in spatially congruent and incongruent crossmodal conditions compared to healthy controls, suggesting intact spatial visuo-tactile integration with simple stimuli. In their second experiment, Noel et al. (2020) used visual stimuli either moving towards or away from the participant to investigate visuo-motor spatial integration depending on the individual's peripersonal space. Contrary to their expectation, visuo-tactile integration was not differently affected in psychosis compared to controls depending on whether stimuli were within vs. outside the peripersonal space, with only the ASD group showing a smaller peripersonal space in which visuo-tactile integration occurred. While this suggested intact visuo-tactile integration in psychosis in the peripersonal space (Noel et al., 2020), this contrasted earlier findings of altered sensorimotor integration and weakened borders of peripersonal space in psychosis (for a review, see Noel, Cascio, et al., 2017).

While the above mentioned studies suggested intact spatial MSI in simple AV and visuo-tactile stimuli in psychosis (de Gelder et al., 2003; Noel et al., 2020; Stephen et al., 2013; Stone et al., 2011, 2014), AV target detection studies in psychosis reported inconsistent findings. Wynn et al. (2014) reported similar AV target detection in patients compared to controls, whereas Williams et al. (2010) observed a reduced crossmodal facilitation effect during target detection in patients compared to controls. Further, Williams et al. (2010) observed significant associations between reduced crossmodal facilitation and negative symptoms, suggesting a link between MSI of AV information and symptoms of psychosis.

Given the scarcity of studies on spatial MSI using simple stimuli in psychosis, the existing evidence suggesting intact spatial integration in psychosis (de Gelder et al., 2003; Noel et al., 2020; Stephen et al., 2013; Stone et al., 2011, 2014) to date should be interpreted cautiously. The comparability of the few published studies might be limited due to variations in experimental paradigms, study designs, and involved sensory modalities. Further, the only published study on spatial AV integration in psychosis administering a VE paradigm (de Gelder et al., 2003), one of the most frequently applied operationalization of spatial integration in healthy participants (Bruns, 2019), deviated from classical VE paradigms by

presenting series of sound and flash stimuli rather than single pairs of AV stimuli. In typical VE experiments, crossmodal stimuli are usually presented with randomized degrees and directions of spatial conflicts to prevent cumulative recalibration (Bruns, 2019). It is possible that localization responses in the study by de Gelder et al. (2003) were influenced by recalibration processes due to the presentation of sequential AV stimuli with fixed spatial disparities. Thus, we aimed to replicate earlier findings of intact spatial MSI in patients with the diagnosis of a psychotic disorder by administering a VE paradigm with single pairs of AV stimuli.

As mentioned throughout this dissertation as an overarching open issue, the question to date remains if deficient MSI in psychosis might be domain-specific, i.e. observable in linguistic, emotional, temporal and sensorimotor but not necessarily simple spatial stimuli (Lin et al., 2020; Noel, Cascio, et al., 2017; Tseng et al., 2015; Zhou et al., 2018), or rather reflect a global impairment. A global deficit in MSI, i.e. impairments across all domains and already observable in simple stimulus conditions and tasks, could e.g. be rooted in an altered global crossmodal binding tendency (Odegaard & Shams, 2017). Crucially, temporal and spatial MSI have been reported to be stable but dissociated processes in a study by Odegaard & Shams (2016). This speaks against any assumption that a deficit in temporal MSI in psychosis (Zhou et al., 2018) might imply that spatial MSI could also be impaired. With our study on spatial multisensory processing in psychosis outlined in this chapter of the dissertation, we aimed to add to the body of evidence on MSI in psychosis necessary to address the above mentioned overarching issue of impairment globality vs. specificity.

### 1.3 Psychosis and Crossmodal Spatial Recalibration

While previous studies suggested intact AV (de Gelder et al., 2003; Stephen et al., 2013; Stone et al., 2011, 2014) and visuo-tactile (Noel et al., 2020) spatial MSI in patients with the diagnosis of a psychotic disorder, to our knowledge no published study investigated crossmodal, i.e. sensory-sensory spatial recalibration in psychosis.

Previous studies on spatial recalibration in psychosis focused on sensorimotor rather than sensory-sensory processes and suggested impaired sensorimotor recalibration. A study by Bartolomeo et al. (2020) showed deficient visuomotor recalibration in patients with the diagnosis of a psychotic disorder compared to healthy controls when reaching a target in a prism adaptation paradigm. Crucially, impaired performance was observed during and after prism adaptation but not during baseline, suggesting deficits specific for visuomotor learning

in psychosis (Bartolomeo et al., 2020). Rösler et al. (2015) observed impaired saccadic remapping in psychosis during a task, in which the target stimulus shifted in position after saccade initiation. Further, impaired remapping in patients with the diagnosis of a psychotic disorder was significantly associated with severity of psychotic symptoms, suggesting a link between visuomotor recalibration processes and psychosis (Rösler et al., 2015). Ferri et al. (2016) showed altered proprioceptive processing and touch remapping in psychosis proneness in a tactile TOJ paradigm, during which two tactile stimuli were presented on the hands in a crossed and uncrossed condition.

While the above discussed studies suggested impaired sensorimotor recalibration processes in psychosis, the findings of two other studies however showed a different alteration pattern. Bansal et al. (2019) observed similar perceptual motor learning in patients with the diagnosis of a psychotic disorder and healthy controls. Subjects had to reach to a visual target with a cursor in a baseline condition, under an adaptation condition with rotated movement behaviour of the cursor and a generalization condition with rotated cursor movement and novel target positions. Patients and controls performed similar on baseline and during adaptation, but patients failed to generalize to novel target positions, suggesting impaired transfer of perceptual motor learning in psychosis (Bansal et al., 2019). Lencer et al. (2017) investigated saccadic remapping with shifting targets after saccade initiation and observed similar adaptation amplitude in patients with the diagnosis of a psychotic disorder and healthy controls. While this contrasted the findings of previous saccade remapping studies in psychosis (Rösler et al., 2015), Lencer et al. (2017) observed a significantly slower adaptation speed and higher amplitude variability in patients compared to controls.

In sum, previous studies suggested altered sensorimotor and perceptual motor learning processes in psychosis, which might however manifest in diverse parameters of motor learning. The extent of e.g. visuomotor recalibration might be impaired or intact depending on the experimental task, while deficits might also manifest in learning rate, precision or transfer. Crucially, all of the above studies indicating altered spatial recalibration processes during perceptual motor learning in psychosis, but evidence whether deficient spatial recalibration might also be observed in non-proprioceptive crossmodal, i.e. sensory-sensory stimuli is lacking.

While prior evidence suggested intact spatial MSI in psychosis (de Gelder et al., 2003; Noel et al., 2020; Stephen et al., 2013; Stone et al., 2011, 2014), this does not allow direct

inferences about whether crossmodal spatial recalibration might be intact, especially in light of previous evidence suggesting at least a partial dissociation between spatial MSI and crossmodal (cumulative) spatial recalibration (Bruns et al., 2022; Park & Kayser, 2019; Rohlf et al., 2020). We argue that it might be theoretically possible that the integration of simple AV information could be intact in psychosis, but the remapping of auditory representations depending on prior exposure to AV conflicts, i.e. crossmodal spatial recalibration, could be impaired – analogue to findings in sensorimotor recalibration processes (e.g. Bartolomeo et al., 2020; Ferri et al., 2016; Rösler et al., 2015). If sensory-sensory recalibration processes might be impaired in psychosis, we argue that this could reflect an impairment to create and maintain a coherent and stable representation of the environment, similar to discussions on representations of the own body in psychosis (Postmes et al., 2014). As such, this could drastically alter the experience of the surrounding world up to potentially confusing and distressing percepts. Since no published study investigated AV spatial recalibration in psychosis, we aim to fill this gap to follow up on the research question if a) both spatial MSI and recalibration are intact in psychosis or b) if spatial recalibration is impaired or c) if even both spatial multisensory processes are impaired. The former would rather support the assumption of domain-specificity of impaired multisensory processes, i.e. deficits in linguistic, emotional (Lin et al., 2020; Tseng et al., 2015), temporal (Zhou et al., 2018), and sensorimotor (Bartolomeo et al., 2020; Ferri et al., 2016; Noel, Cascio, et al., 2017; Rösler et al., 2015) but not in spatial processes. The latter would offer support for an assumption of a global deficit and of an a universally impaired underlying mechanism of multisensory processing in psychosis.

### 1.4 Research Question

In this study, we investigated spatial MSI and crossmodal spatial recalibration in psychosis using simple AV stimuli. We aimed to investigate whether psychosis is associated with intact spatial ventriloquism, which would replicate earlier findings by de Gelder et al. (2003), or if spatial ventriloquism is altered in psychosis. Further, we investigated if crossmodal spatial recalibration, measured via the cumulative and immediate VAE, is altered or intact in psychosis. To our knowledge, our study is the first to investigate the VAE in psychosis. We therefore aimed to extend previous findings on altered spatial recalibration in psychosis in the sensorimotor domain (e.g. Bartolomeo et al., 2020; Ferri et al., 2016; Rösler et al., 2015) to AV stimuli. Finally, we investigated associations between spatial MSI,

cumulative as well as immediate spatial recalibration and self-reports of psychotic symptoms and negative schemata.

Patients with the diagnosis of a psychotic disorder and healthy controls localized auditory stimuli presented with one of two frequencies either in unimodal auditory (UA) or AV trials. During AV trials, the visual stimulus was either presented with a fixed spatial distance to the left or right of the sound depending on the sound frequency of the auditory stimulus, resulting in two sound frequency-specific AV spatial conflicts (paradigm adopted from Bruns & Röder, 2015). For each frequency, we calculated the deviance of sound localization responses in AV trials from the veridical sound position, i.e. the VE reflecting spatial MSI. Further, per frequency we calculated the deviance of sound localization responses in UA trials from the veridical sound position as a function of repeated exposure to AV spatial conflicts, i.e. the cumulative VAE reflecting cumulative spatial recalibration, and as a function of AV spatial conflicts during AV trials in the immediate stimulus history, i.e. the immediate VAE reflecting immediate spatial recalibration.

Given the scarcity of findings in AV spatial MSI in psychosis, especially regarding its potentially most frequently applied operationalization via the VE (Bruns, 2019), we aimed to replicate the findings of de Gelder et al. (2003) and - in a broader context - those of other studies on spatial MSI in psychosis (Noel et al., 2020; Stephen et al., 2013; Stone et al., 2011, 2014). Thus, we expected our patient sample to show intact spatial MSI, reflected by a comparable VE size in patients compared to healthy controls. Based on findings of impaired sensorimotor recalibration in psychosis (e.g. Bartolomeo et al., 2020; Ferri et al., 2016; Rösler et al., 2015), we expected to observe altered spatial recalibration in patients with the diagnosis of a psychotic disorder. However, in light of previous findings on a dissociation of cumulative spatial recalibration and spatial MSI as well as evidence that immediate spatial recalibration and spatial MSI might be based on the same underlying processes (Bruns et al., 2022; Park & Kayser, 2019; Rohlf et al., 2020), we expected different patterns for cumulative and immediate spatial recalibration. We expected to observe altered cumulative recalibration, indicated by a difference in the size of the cumulative VAE, but intact immediate spatial recalibration, indicated by a comparable size of the VAE, in patients with the diagnosis of a psychotic disorder compared to healthy controls. Finally, we explored associations between the VE, the cumulative as well as the immediate VAE and anomalous perceptual experiences, hallucinations, paranoid delusions, passivity experiences as well as negative schemata.

# 2 Methods

## 2.1 Participants

We recruited 35 patients with the diagnosis of a psychotic disorder via flyers in outpatient clinics and social care centers as well as contact data bases from earlier studies with consent for further contact and 32 healthy controls via online recruitment services from the University of Hamburg. 3 patients cancelled and 3 additional patients discontinued the participation. Further, we excluded 1 participant from the healthy control group and 2 from the patient group for not being able to sufficiently execute the experimental task (see **2.4 Data Analysis**). The final sample therefore consisted of 27 patients with the diagnosis of a psychotic and 31 healthy controls (see Table 5 for demographic data). Participants differed significantly in age, with patients showing a higher mean age ($M$ = 41.0, $SD$ = 10.6) than healthy controls ($M$ = 32.3, $SD$ = 14.6), $t(59) = 2.54$, $p = .013$, Cohen's $d = 0.67$. Groups did not differ regarding the distribution of gender, $X^2(2) = 2.22$, $p = .330$, $\varphi = 0.20$, or education, $X^2(3) = 4.47$, $p = .215$, $\varphi = 0.28$.

Patients were included if they fulfilled the DSM-V (American Psychiatric Association, 2022) criteria for schizophrenia or schizoaffective disorder. Patients were excluded if they had a comorbid autism spectrum disorder. Healthy participants were excluded if they had a past or present psychiatric disorder or a family history of first degree relatives with the diagnosis of a psychotic disorder. General exclusion criteria consisted of dementia or other neurological disorders, acute suicidality, substance abuse disorder in the last 6 months, uncorrected sight or hearing impairment as well as inability to give consent or understand study instructions due to health status or language barriers.

Participants gave written informed consent prior to participation and received 10€/hour or course credits as compensation. The study was performed by the standards of the Declaration of Helsinki and was approved by the Local Ethics Committee of the Faculty of Psychology and Human Movement Science, University of Hamburg.

***Table 5***

*Demographic and Diagnostic Data Separated by Group*

| | Patients | Healthy Controls |
|---|---|---|
| *n* | 27 | 31 |
| Age, *M* (range) | 41.0 (20-58) | 32.3 (18-65) |
| Gender, *n* (%) | | |
| female | 13 (48.1) | 19 (61.3) |
| male | 14 (51.9) | 11 (35.5) |
| diverse | 0 | 1 (3.2) |
| Education, *n* | | |
| 9 years | 0 | 1 |
| 10 years | 8 | 3 |
| 12-13 years[a] | 12 | 18 |
| university degree[b] | 7 | 9 |
| CAPS, *M* (*SD*) | 9.9 (6.7) | 2.2 (2.4) |
| LSHS-E, *M* (*SD*) | 22.3 (11.3) | 4.4 (4.4) |
| PCL, *M* (*SD*) | 81.7 (56.4) | 21.3 (18.7) |
| Passivity experiences, *M* (*SD*) | 5.5 (4.4) | 0.3 (0.5) |
| BCSS n.S., *M* (*SD*) | 6.9 (4.9) | 2.0 (2.1) |
| BCSS n.O., *M* (*SD*) | 7.9 (6.1) | 3.3 (3.9) |

*Notes.* $N = 58$. CAPS = Cardiff Anomalous Perception Scale, total score; LSHS-E = Launey-Slade Hallucination Scale-Extended, total score; PCL= Paranoia Checklist, total score; BCSS n.S. = Brief Core Schema Scales, negative self subscale; BCSS n.O. = Brief Core Schema Scales, negative other subscale.

[a] due to changes in German education law in 2007, the years of education for acquire the Abitur (A-level/high school equivalent) vary between 12-13 years depending on school and Bundesland; [b] university degree: Bachelor level or higher.

## 2.2 Stimuli and Experimental Paradigm

The experimental paradigm was adapted from Bruns & Röder (2015). Stimulus presentation was controlled and participant responses were recorded by means of the Neurobs Presentation® Software (Version 22; Neurobehavioural Systems Inc., Berkeley, California, USA).

Stimuli were presented in a dark and sound-attenuated room using a custom built set-up consisting of a semi-circular metal rack of 85cm radius, covered in an acoustically transparent black curtain. Participants were seated at the center of the semi-circle and rested their head on a chin rest at 90cm distance to the semi-circle. Auditory stimuli were presented via speakers (ConceptC Satellit, Teufel GmbH, Berlin, Germany), positioned on the semi-circle at ±4.5, ±13.5 and ±22.5 degrees from the center and slightly above eye level. Auditory stimuli consisted of 750Hz and 3000Hz sinusoidal tones with a duration of 200ms (5ms rise and fall time) and were presented at 65dB(A), measured at the subject's head position. For each presentation, auditory stimuli randomly varied within a range of 4dB(A) to account for any potential variation in speaker transformation function. Visual stimuli were presented by means of LED-panels (APA102C RGB Full Color LED control IC, iPixel LED, Shiji Lighting, Shiyan, China) with 256 color LEDs in a horizontal row (LED diameter = 0.5 cm, spacing between LEDs = 0.5 cm, 2.54 ppi), arranged in front of the curtain and at eye-level and covered the entire horizontal length of the semi-circle. Visual stimuli consisted of red flashes of single LEDs with a duration of 200ms. For sound localization response, participants used a movable pointer attached below the chin rest. The pointer was equipped with two potentiometers, measuring the X- and Y-coordinates (in degrees relative to the pointer origin) of the pointer position. The X and Y coordinate of the current pointer orientation as well as the response time was recorded as soon as the participant pressed a button on the tip of the pointer.

Prior to the experimental runs with participants, precise AV timing was ensured via calibrating the presentation times of speakers and LEDs using a photometer and adjusting for technically induced latencies.

The experimental paradigm consisted of 240 UA and 240 AV trials, divided in three blocks of 160 trials. Each combination of frequency, speaker position and trial type (UA or AV) was presented 20 times over course of the experiment, resulting in 480 trials. UA and AV trials were presented in pseudorandom order. The trial order was preset so that for each

frequency half of the UA trials were preceded by an AV trial with the opposite frequency and the other half by an AV trial of the same frequency, not counting UA trials in between. Further, for the 60 UA trials of one frequency preceded by an AV trial of the same frequency, 15 trials each were preceded with a same-frequency AV trial in the last 1-4 trials.

Each trial started with a white fixation point of a single LED at the center. To start each trial with the approximately same pointer orientation, participants were instructed to point at the fixation point and press the button. If the pointer orientation was within a horizontal range of ±10 degrees relative to the fixation, the fixation point disappeared, followed by a randomized ISI of 750-1500ms (uniform distribution). Otherwise participants had to re-orient the pointer and press again. Depending on trial type, subsequently an auditory stimulus was either presented alone at one of the speaker positions or together with a synchronous visual stimulus, horizontally displaced at ±13.5 degrees relative to the auditory stimulus. For each participant, the visual stimulus was always presented 13.5 degrees to the left relative to sounds of one of the two frequencies (either 750Hz or 3000Hz) and 13.5 degrees to the right relative to the other, resulting in two frequency-specific AV adaptation directions: leftward-adaptation and rightward-adaptation. We balanced the allocation of frequency and leftward/rightward-adaptation over participants in dependence of the participant's subject number in the Neurobs Presentation® script, allocating each participant automatically to one of two adaptation direction groups: even subject numbers were allocated to the 750R/3000L and odd numbers to 750L/3000R adaptation direction subgroup.

Participants were instructed to indicate the perceived position of the tone by orienting the tip of the pointer towards the position and button press while ignoring any visual stimulus. Following the button press, the next trial started with the fixation point. After 160 trials each, participants could take a short break of 2-3 minutes. The start and end of each block was announced by five flashes of the white fixation point of 500ms duration each.

### 2.3 General Procedure

The study consisted of two sessions, which each lasted about 1.25 to 1.5 hours and took place on two different days.

The first session covered informed consent and diagnostic procedures. After discussing informed consent, participants provided information about age, gender, education level, potential sight or hearing impairment and respective corrections, potential psychiatric diagnoses and their current status, illness duration of psychotic disorders, potential family

history of psychotic disorder, psychopharmacological medication, wakefulness, motivation, stress level and substance consumption in the last 24 hours. Handedness was assessed using the EHI (Oldfield, 1971), followed by the TMT Version A and B (Reitan & Wolfson, 1985) to measure attention and processing speed. Subsequently, the in- and exclusion criteria regarding psychiatric diagnoses were checked using the SCID-V (Beesdo-Baum et al., 2019). Healthy controls were interviewed using the sections A "Affective Disorders" and B "Psychotic Disorders", followed by a screening for the remaining sections. If any of the screening questions was scored with "yes", the respective section was conducted. Patients with the diagnosis of a psychotic disorder were also interviewed using sections A and B, but not with the screening. Instead, the PSYRATS (Haddock et al., 1999) was administered to further enquire more details about acoustic verbal hallucinations and delusions. Finally, the CPT-IP (Nuechterlein et al., 2008) was administered to measure vigilance.

Between session, participants were asked to provide self-reports about usual substance consumption, schemata about the self and others by means of the BCSS (Fowler et al., 2006), anomalous perceptual experiences by means of the CAPS (Bell et al., 2006), hallucinations by means of the LSHS-E (Lincoln et al., 2009; Siddi et al., 2019), paranoia by means of the PCL (Freeman et al., 2005) and passivity experiences by means of the factor "bizarre experiences" of the CAPE (Schlier et al., 2015; Stefanis et al., 2002) (a glossary of the instruments and questionnaires used in this dissertation project can be found in Table A.1 in Appendix A). Participants completed the questionnaires at home and returned them at the second session. Potential questions could be clarified at the end of the second session.

The experimental paradigm (see **2.2 Stimuli and Experimental Paradigm**) was administered during the second session. The second session took place within 1-10 days after the first session, depending on the individual participant's schedule. Participants were comfortably seated in a chair in the center of the experimental set-up and placed their chin on the chin rest. Participants received experimental instructions and potential questions were answered. Afterwards, the first block was started, followed by block two and three with short breaks in-between. Over the whole course of the experiment, the experimenter was seated in the adjacent control room and surveyed the experiment progress on a screen. After the experimental run, participants were de-briefed and received participation compensation.

## 2.4 Data Analysis

Prior to data analysis, we checked whether participants were able to localize sounds in appropriate relative position to each speaker. We fit a regression line over localization responses depending on speaker location in UA trials separately for each frequency and per participant. The regression line slope of each participant was then compared to the mean slope per frequency and group. We chose UA trials for investigating the relationship between localization performance and speaker position, since sound localization response in AV trials can reasonably be expected to be more biased towards the concurrently presented visual stimulus compared to UA trials. Based on this slope comparison, three participants (two patients and one healthy control) were excluded for having a slope close to zero and more than 1.5 *SD* below the average. Further, we plotted the localization responses relative to the speaker positions in UA trials separately for both groups and frequencies to visualize localization performance.

After data collection, we noticed that the allocation of the two the adaptation direction subgroups (750R/3000L vs. 750L/3000R) did not result in a balanced distribution due to technical difficulties (37 vs. 21). Importantly, the adaptation direction ratio was comparable in patients (17 vs. 10) and healthy controls (20 vs. 11). Further, data analysis revealed that the effect pattern was similar in both adaptation direction subgroups (see ***3.1.7 Exploratory Analysis of Non-Merged Data***), so it can be assumed that this allocation imbalance did not impact the results.

Potential group differences regarding age and distribution of gender and education were analyzed by means of an independent t-test and Pearson $X^2$-tests, respectively.

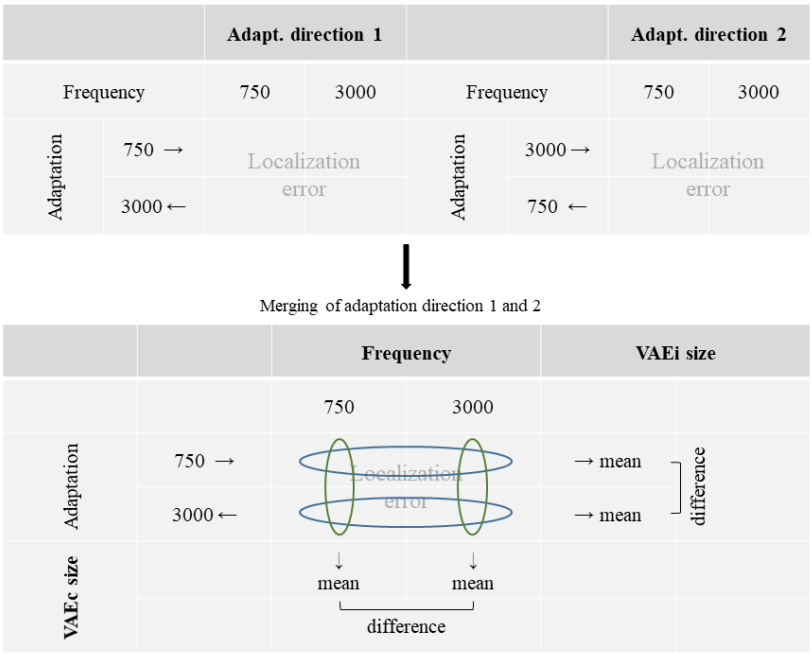
Experimental data was preprocessed and statistical analysis conducted using R Version 4.2.2 (R Core Team, 2022). The preprocessing steps were adapted from Rohlf et al., (2020).

To quantify the VE, the mean constant error, i.e. the mean deviation of the sound localization response from the true sound location, during AV trials was calculated per participant. Within each of the two experimental groups (patients and healthy controls), we merged the data from both adaptation direction groups (750R/3000L and 750L/3000R) by inverting the sign of the mean constant error of the second adaption direction group. This data merging was done for all of the following preprocessing steps of the experimental data (see Figure 8 for an illustration of the data merging process). A 2x2 mixed ANOVA with the

between-factor *Group* (patients vs. healthy controls) and the within-factor *Direction* (leftward-adaptation vs. rightward-adaptation) on the mean constant error in AV trials was computed to investigate group differences in VE. Further, we defined the size of the VE as the difference in mean constant error between the leftward-adaptation and the rightward-adaptation AV trials and quantified the VE size per participant accordingly.

***Figure 8***

*Illustration of the Data Merging Process*




*Notes*. VAEc = cumulative ventriloquist aftereffect; VAEi = immediate ventriloquist aftereffect. The constant error was calculated as the mean deviation of the localization response from the veridical sound location. To enable comparison over the entire groups of patients and healthy controls, the data from the second adaptation direction (750 left/3000 right) subgroup was inverted in signs and then merged with the data from the first adaptation (750 right/3000 left) subgroup. Afterwards, the cumulative VAE was calculated as the difference between the mean constant errors in unimodal auditory trials of the leftward- vs. rightward adapted frequency. The immediate VAE was calculated as the difference between the mean constant errors in unimodal auditory trials preceded by a leftward-adaptation vs. rightward-adaptation audiovisual trial. Adapted from "Multisensory Integration Develops Prior to Crossmodal Recalibration" by S. Rohlf, L. Li, P. Bruns & R. Röder, 2020, *Current Biology*, *30*(9), 1726-1732 (https://doi.org/10.1016/j.cub.2020.02.048).

To quantify the cumulative and immediate VAE, the mean constant error was calculated for each UA trial per participant. After data merging of the two adaptation direction groups (see paragraph above), both aftereffects were calculated accordingly. A 2x2 mixed ANOVA with the between-factor *Group* (patients vs. healthy controls) and within-factor *Adapted Direction* (leftward-adapted vs. rightward-adapted) on the mean constant error

during UA trials was computed to investigate group differences in cumulative VAE. Further, we defined the size of the cumulative VAE as the difference in mean constant error between UA trials with the leftward-adapted frequency and UA trials with the rightward-adapted frequency and calculated it per participant.

A 2x2 mixed ANOVA with the between-factor *Group* (patients vs. healthy controls) and the within-factor *Direction of Preceding AV Adaptation* (preceding leftward-adaptation vs. rightward-adaptation AV trial) on the mean constant error was computed to investigate group differences in immediate VAE. Additionally, we defined the size of the immediate VAE as the difference in mean constant error between UA trials preceded by a leftward-adaptation AV trial and UA trials preceded by a rightward-adaptation AV trial and calculated it per participant.

Additionally, the accumulation of the immediate VAE over the last one to four consecutive preceding AV trials with one frequency was calculated. The difference in mean constant error of the merged data (see above) between leftward- and rightward-adapted UA trials preceded by the same frequency and different frequency for the last one to four consecutive AV trials was calculated. Positive values indicated a shift of auditory localization performance in direction of the preceding AV mismatch, negative values a shift in the opposite direction of the preceding mismatch. Two separate mixed ANOVAs with the between-factor *Group* (patients vs. controls) and the within-factor *Consecutive Trial Number* (one, two, three, four) were calculated entering the size of the immediate VAE for UA trials preceded by an AV trial with the same frequency and with the other frequency, respectively, followed by Bonferroni-corrected post-hoc t-Tests. We applied the Greenhouse-Geisser correction (Greenhouse & Geisser, 1959) in case of sphericity violations.

To check if the mapping of frequency and adaptation direction had a systematic effect on localization performance, we additionally analyzed the non-merged data. For AV trials, we calculated an exploratory mixed ANOVA with the between-factors *Group* (patients vs. healthy controls) and *Adaptation Direction* (750R/3000L vs. 750L/3000R) and the within-factor *Frequency* (750Hz vs. 3000Hz) entering the mean constant error in AV trials. For UA trials, we calculated an exploratory mixed ANOVA with the between-factors *Group* (patients vs. healthy controls) and *Adaptation Direction* (750R/3000L vs. 750L/3000R) and the within-factors *Frequency* (750Hz vs. 3000Hz) and *Frequency of prior adaptation trail* (same vs. different) entering the mean constant error in UA trials.

Welch two sample t-tests were conducted to check for group differences in attention and processing speed, entering TMT-A and -B completion times (in seconds). Further, potential group differences regarding vigilance were investigated by means of a mixed ANOVA on CPT-IP *d'* scores with *Group* (patients vs. healthy controls) and *Digit Load* (2 vs. 3 vs. 4 digits) as factors, followed by post-hoc t-tests. The influence of any control variable on the main dependent variables was estimated by means of Spearman's rank correlation coefficients ($\rho$) with Bonferroni-corrected *p*-values, if group differences in TMT or CPT-IP were revealed.

CAPS, LSHS-E, PCL, passivity experiences and BCSS scores were compared between patients and healthy controls by means of Welch two sample t-tests, separately for each of the questionnaires.

To investigate associations between AV integration as well as recalibration and psychosis-related constructs, Spearman rank correlation coefficients $\rho$ with Bonferroni-corrected *p*-values were calculated between the size of the VE, cumulative VAE and immediate VAE and questionnaires scores. For all correlations, Spearman rank correlation coefficients $\rho > .3$ were defined as potentially meaningful for discussion and highlighted in correlation matrices.

The above mentioned mixed ANOVAS, t-tests, and correlations were additionally investigated by means of Bayesian hypothesis testing using standard priors in JASP Version 0.16.4 (Wagenmakers et al., 2018) with reported BF indicating evidence for either the null or alternative hypothesis. Per convention, a BF of 1 is discussed to reflect no evidence. A BF between 3-10 indicates moderate, between 10-100 strong, and >100 extreme evidence in favor of the alternative hypothesis. Accordingly, a BF between 1/3-1/10 reflects moderate, between 1/10-1/100 strong, and <1/100 extreme evidence in favor of the null hypothesis (for a complete BF classification scheme, see Wagenmakers et al., 2018). For two-tailed tests and correlations we reported $BF_{10}$, the BF describing evidence for the alternative relative to the null hypothesis, and for mixed ANOVAS $BF_{incl}$, the BF reflecting the posterior probability for model inclusion for a factor or interaction term averaged over all candidate models (Wagenmakers et al., 2018).

# 3 Results

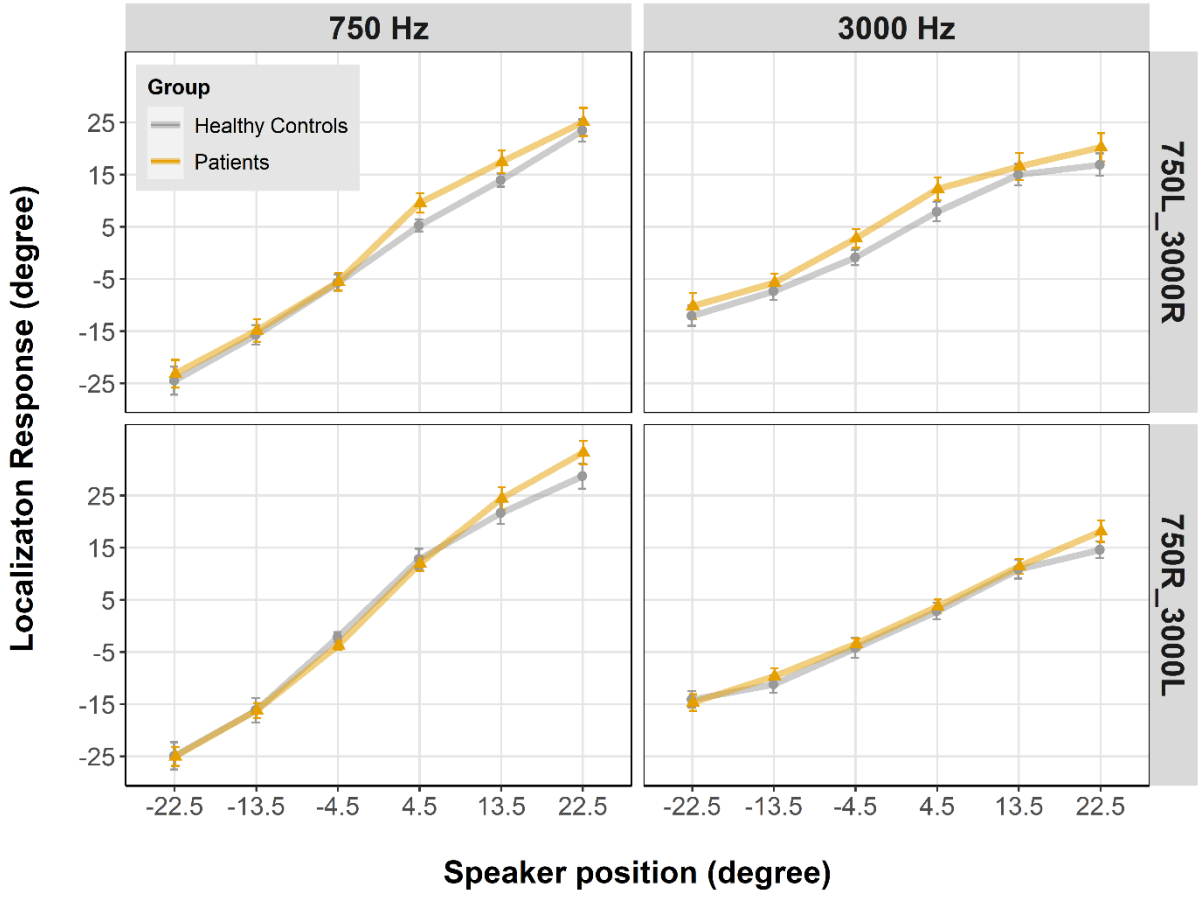
## 3.1 Experimental Data

### *3.1.1 Localization Responses Relative to Speaker Positions*

Figure 9 depicts the mean localization performance in UA trials relative to speaker position for each group and adaptation direction combination. For all four combinations, visual inspection implied a linear relationship between localization performance and veridical sound location.

***Figure 9***

*Localization Responses Relative to Speaker Position in Unimodal Auditory Trials*



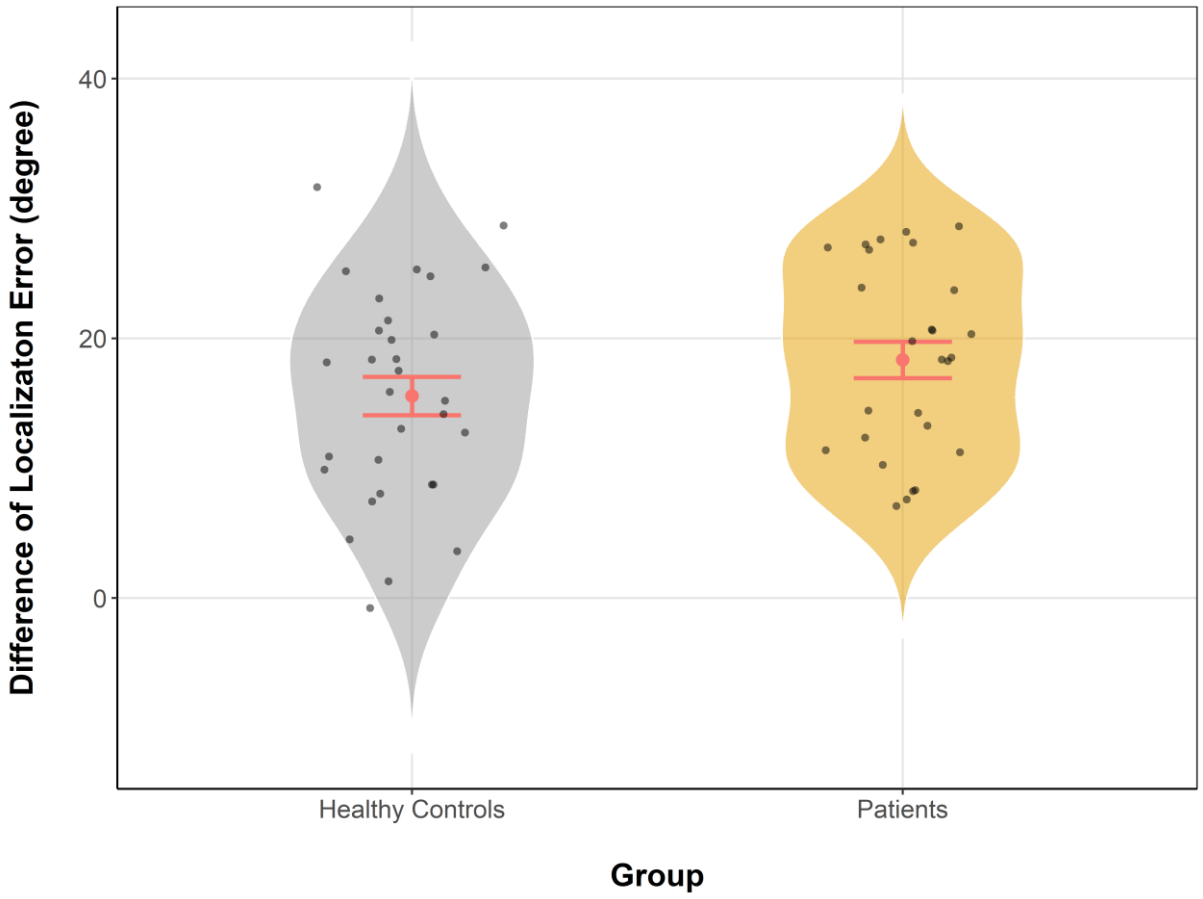

*Notes.* $n_{HC}$ = 31; $n_{PAT}$ = 27. Per group, the localization response in unimodal auditory trials relative to each speaker position is shown for all combinations of frequency (750 Hz & 3000 Hz) and adaptation direction (750 Hz leftward/3000 Hz rightward & 750 Hz rightward/3000 Hz leftward). Error bars denote the SEM.

### *3.1.2 VE*

A mixed ANOVA on the mean constant error in AV trials revealed a highly significant main effect of *Direction*, $F(1, 56) = 271.23$, $p < .001$, $\eta^2_p = 0.83$, $BF_{incl} > 1000$, indicating that the mean constant error differed between leftward-adaptation ($M = -8.73$, $SD = 4.69$) and rightward-adaptation ($M = 8.13$, $SD = 5.78$) AV trials. There was no main effect of *Group*, $F(1, 56) = 0.1$, $p = 0.757$, $\eta^2_p < 0.01$, $BF_{incl} = 0.26$, and no *Group* * *Direction* interaction, $F(1, 56) = 1.83$, $p = 0.182$, $\eta^2_p = 0.03$, $BF_{incl} = 0.53$. This indicated that the VE size, i.e. difference in mean constant error between leftward- and rightward-adaptation AV trials, did not differ between patients ($M = 15.35$, $SD = 7.27$) and healthy controls ($M = 15.57$, $SD = 8.23$). Two separate one-tailed t-Tests confirmed that each group independently showed a significant VE, both $p < .001$ (see Figure 10 for the mean and distribution of VE size per group).

***Figure 10***

*Mean and Distribution of Ventriloquist Effect Size*




*Notes.* $n_{HC} = 31$; $n_{PAT} = 27$. The difference in the constant error between leftward- and the rightward-adapted audiovisual trials is shown. Per group, the mean, error bars denoting the SEM, the value distributions as well as the individual data points are depicted.

### *3.1.3 Cumulative VAE*

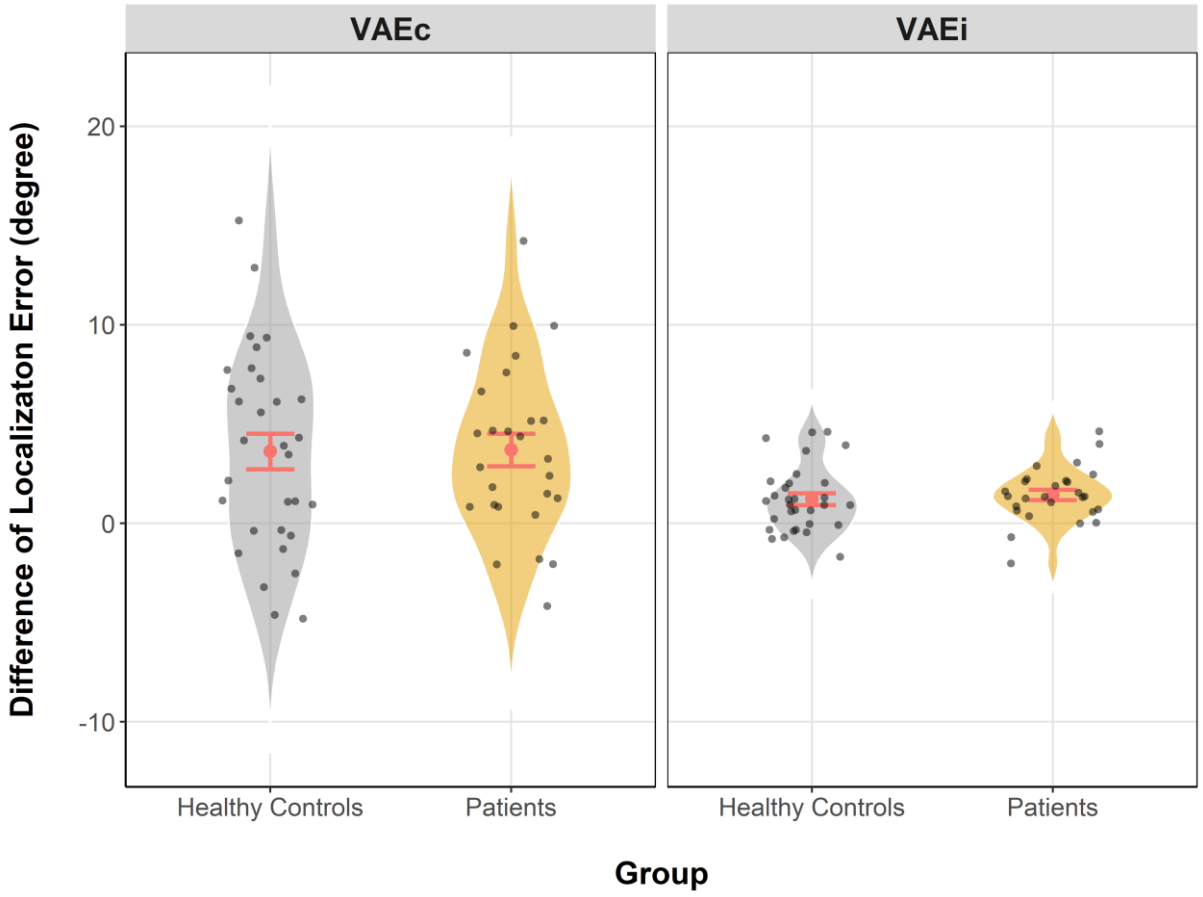
A mixed ANOVA on the mean constant error in UA trials revealed a highly significant main effect of *Adapted Direction*, $F(1, 56) = 35.47$, $p < .001$, $\eta^2_p = 0.39$, $BF_{incl} > 1000$, indicating that the mean constant error differed between leftward-adapted ($M = -1.45$, $SD = 5.0$) and rightward-adapted ($M = 2.21$, $SD = 3.57$) UA trials. There was no main effect of *Group*, $F(1, 56) = 0.11$, $p = 0.743$, $\eta^2_p < 0.01$, $BF_{incl} = 0.25$, and no *Group * Adapted Direction* interaction, $F(1, 56) < 0.01$, $p = .953$, $\eta^2_p < 0.01$, $BF_{incl} = 0.25$. This indicated that the size of the cumulative VAE, i.e. the difference in mean constant error between leftward- and rightward-adapted UA trials, did not differ between patients ($M = 3.70$, $SD = 4.28$) and healthy controls ($M = 3.62$, $SD = 5.0$). Two separate one-tailed t-Tests confirmed that each group independently showed a significant cumulative VAE, both $p < .001$ (see Figure 11 left panel for the mean and distribution of cumulative VAE size per group).

### *3.1.4 Immediate VAE*

A mixed ANOVA on the mean constant error in UA trials with a preceding leftward- vs. rightward-adaptation AV trial revealed a highly significant main effect of *Direction of Preceding AV Adaptation*, $F(1, 56) = 43.21$, $p < .001$, $\eta^2_p = 0.44$, $BF_{incl} > 1000$, indicating that the mean constant error differed between UA trials with an preceding leftward-adaptation ($M = -0.28$, $SD = 3.74$) vs. a preceding rightward-adaptation ($M = 1.04$, $SD = 3.75$) AV trial. There was no main effect of *Group*, $F(1, 56) = 0.11$, $p = 0.743$, $\eta^2_p < 0.01$, $BF_{incl} = 0.52$, and no *Group * Direction of Preceding AV Adaptation* interaction, $F(1, 56) = 0.28$, $p = 0.597$, $\eta^2_p = 0.01$, $BF_{incl} = 0.48$. This indicated that the size of the immediate VAE, i.e. the difference in mean constant error between UA trials with a preceding leftward- vs. rightward-adaptation AV trial, did not differ between patients ($M = 1.43$, $SD = 1.37$) and healthy controls ($M = 1.22$, $SD = 1.65$). Two separate one-tailed t-Tests confirmed that each group independently showed a significant immediate VEA, both $p < .001$ (see Figure 11 right panel for the mean and distribution of immediate VAE size per group).

***Figure 11***

*Mean & Distribution of Cumulative and Immediate Ventriloquist Aftereffect Size*




*Notes.* $n_{HC}$ = 31; $n_{PAT}$ = 27. VAEc = cumulative ventriloquist aftereffect; VAEi = immediate ventriloquist aftereffect. The left panel shows the difference in constant error between the leftward- and rightward-adapted sound frequency in unimodal auditory trials. The right panel shows the difference in constant error between unimodal auditory trials preceded by an audiovisual trial with leftward- vs. rightward-adaptation. Per group, the mean, error bars denoting the SEM, the value distributions as well as the individual data points are depicted for both the cumulative and immediate VAE.

### *3.1.5 Accumulation of the Immediate VAE*

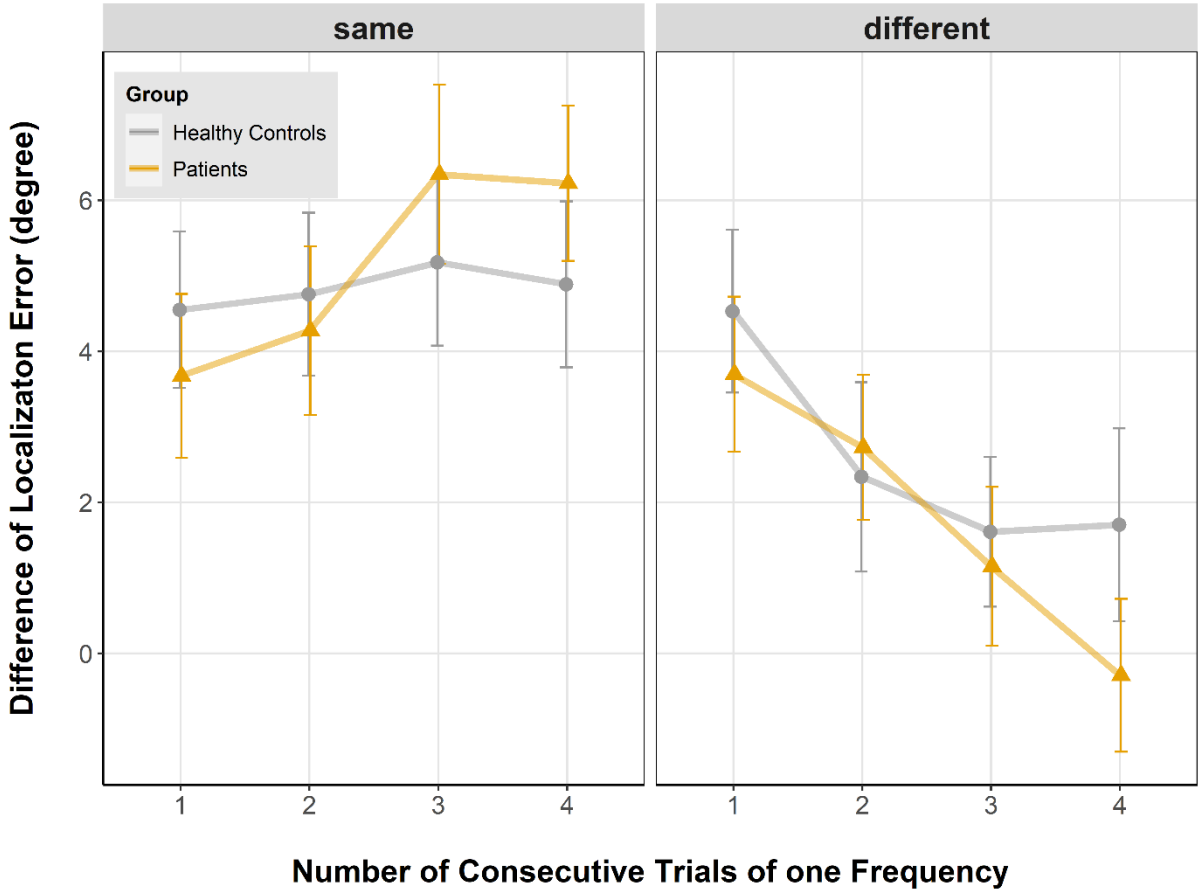
**Consecutive Preceding Trials with Same Frequency.** A mixed ANOVA for the accumulation of the immediate VAE in UA trials preceded by AV trials with the same frequency revealed a significant main effect of *Consecutive Trial Number*, $F(3, 168) = 2.90$, $p = .037$, $\eta^2_p = 0.5$, $BF_{incl} = 0.53$. However, Bonferroni-corrected post-hoc t-Tests for pairwise comparisons did not reveal any significant difference in immediate VAE size between UA trials with one, two, three or four consecutive preceding AV same frequency trials, all $p > .05$. There was no significant main effect for *Group*, $F(1, 56) = 0.05$, $p = .832$, $\eta^2_p < 0.01$, $BF_{incl} = 0.26$, nor a significant *Group* * *Consecutive Trial Number* interaction, $F(3, 168) = 1.59$, $p =$

.193, $\eta^2_p$ = 0.03, $BF_{incl}$ = 0.14 (see Figure 12 for the accumulation of immediate VAE size over the last one to four adaptation trials).

**Consecutive Preceding Trials with Different Frequency.** A mixed ANOVA for the accumulation of the immediate VAE in UA trials preceded by AV trials with the other frequency revealed a highly significant main effect of *Consecutive Trial Number*, $F(3, 168) = 8.84$, $p < .001$, $\eta^2_p = 0.14$, $BF_{incl} = 688.32$. Bonferroni-corrected post-hoc t-Tests for pairwise comparisons revealed a significant difference in immediate VAE size between UA trials with one ($M = 4.14$, $SD = 5.66$) and with three ($M = 1.40$, $SD = 5.45$) consecutive preceding AV different frequency trials, $t(57) = 4.45$, $p < .001$, Cohen's $d = 0.49$, as well as between UA trials with one and with four ($M = 0.78$, $SD = 6.34$) consecutive preceding AV different frequency trials, $t(57) = 5.08$, $p < .001$, Cohen's $d = 0.56$, reflecting that the VAE induced by the frequency - adaptation direction mapping decreased with increasing numbers of preceding adaptation trials of the different frequency. All other post-hoc comparisons did not reach statistical significance, all $p > .05$. There was no significant main effect for *Group*, $F(1, 56) = 0.31$, $p = .580$, $\eta^2_p < 0.01$, $BF_{incl} = 0.28$, nor a significant *Group * Consecutive Trial Number* interaction, $F(3, 168) = 0.99$, $p = .400$, $\eta^2_p = 0.02$, $BF_{incl} = 0.15$ (see Figure 12 for the accumulation of immediate VAE size over the last one to four adaptation trials).

***Figure 12***

*Accumulation of the Immediate Ventriloquist Aftereffect over Consecutive Preceding Audiovisual Trials with one Frequency*




*Notes.* $n_{HC}$ = 31; $n_{PAT}$ = 27. The left panel shows the difference in constant error between unimodal auditory (UA) trials preceded by an audiovisual (AV) trial with leftward- vs. rightward-adaptation relative to the number of consecutive trials with the same frequency as the UA trials. The right panel shows the difference in constant error between UA trials preceded by an AV trial with leftward- vs. rightward-adaptation relative to the number of consecutive trials with the different frequency as the UA trials. Error bars denote the SEM.

### *3.1.6 Correlations Between VE, Cumulative VAE, and Immediate VAE*

The VE showed significant, moderate Spearman rank correlations with both the cumulative VAE, Spearman's $\rho$ = .44, $p$ = .003, $BF_{10}$ = 261.81, and the immediate VAE, Spearman's $\rho$ = .36, $p$ = .018, $BF_{10}$ = 7.56. Cumulative VAE and immediate VAE showed no significant correlation, Spearman's $\rho$ = .15, $p$ = .753, $BF_{10}$ = 0.26. $p$-values are Bonferroni-corrected (see Table 6 for Spearman rank correlation coefficients $\rho$. Exploratory Spearman rank coefficients $\rho$ split by group can be found in Table D.1 and Table D.2 in Appendix D).

***Table 6***

*Spearman Rank Correlation Coefficients ρ for Associations Between VE, Cumulative VAE, Immediate VAE and Questionnaire Scores over all Participants*

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. | 9. |
|---|---|---|---|---|---|---|---|---|---|
| 1. VE | | | | | | | | | |
| 2. VAEc | **.44**** | | | | | | | | |
| 3. VAEi | **.36*** | .15 | | | | | | | |
| 4. CAPS | .21 | .18 | .14 | | | | | | |
| 5. LSHS-E | .22 | .10 | .11 | **.79***** | | | | | |
| 6. PCL | .18 | .05 | .11 | **.68***** | **.66***** | | | | |
| 7. Pass. Exp. | .19 | -.06 | .12 | **.67***** | **.79***** | **.59***** | | | |
| 8. BCSS n.S. | .14 | -.03 | .07 | **.63***** | **.63***** | **.63***** | **.60***** | | |
| 9. BCSS n.O. | .21 | .02 | -.13 | **.39*** | **.49***** | **.55***** | **.47***** | **.55***** | |

*Notes.* N = 58. Spearman rank correlation coefficients $\rho > .3$ are marked in bold. *p*-values are Bonferroni-corrected. VE = ventriloquist effect size; VAEc = cumulative ventriloquist aftereffect size; VAEi = immediate ventriloquist aftereffect size; CAPS = Cardiff Anomalous Perception Scale, total score; LSHS-E = Launey-Slade Hallucination Scale-Extended, total score; PCL = Paranoia Checklist, total score; Pass. Exp. = total score in passivity experiences items; BCSS n.S. = Brief Core Schema Scales, negative self subscale; BCSS n.O. = Brief Core Schema Scales, negative other subscale.

* $p < .05$, ** $p < .01$, *** $p < .001$

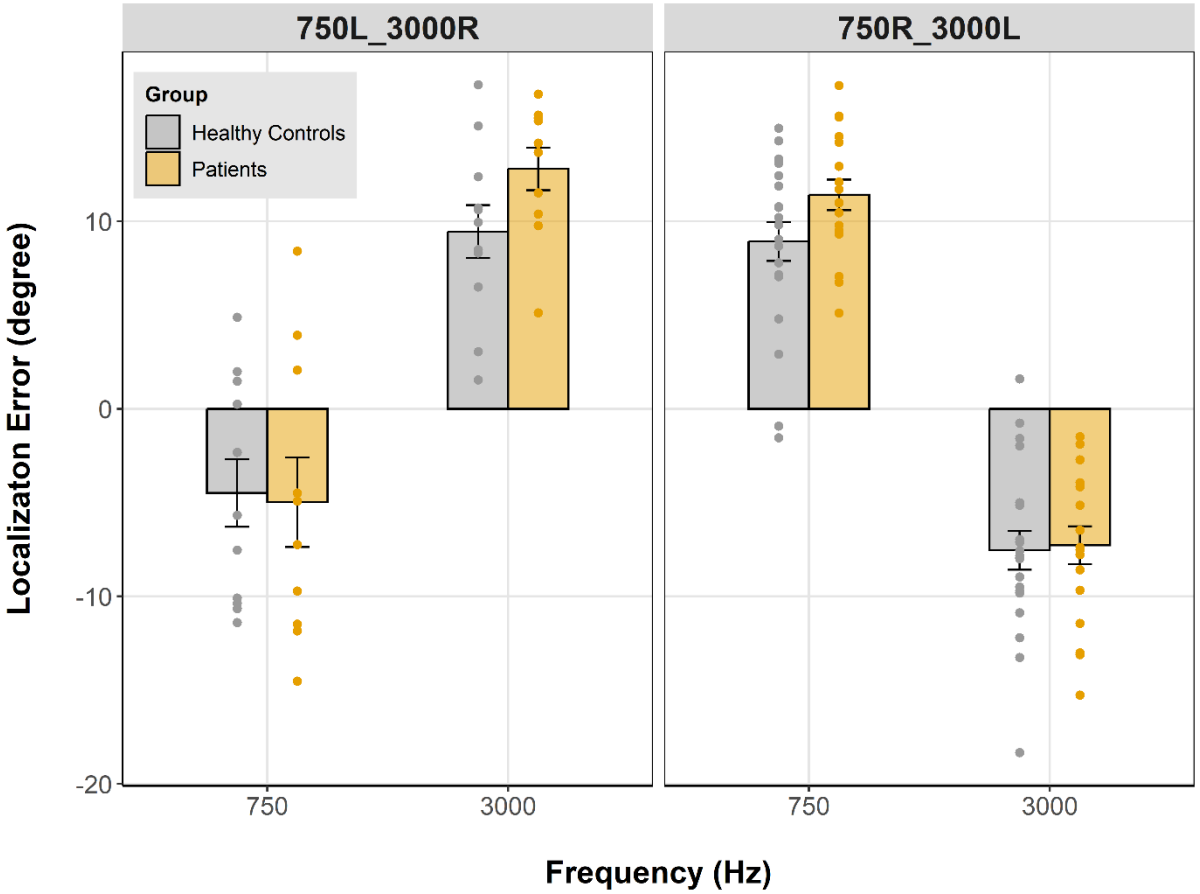
### *3.1.7 Exploratory Analysis of Non-Merged Data*

**AV Trials.** An exploratory mixed ANOVA with the between-factors *Group* (patients vs. healthy controls) and *Adaptation Direction* (750R/3000L vs. 750L/3000R) and the within-factor *Frequency* (750Hz vs. 3000Hz) on mean constant error in AV trials revealed a highly significant main effect of *Frequency*, $F(1, 54) = 27.68$, $p < .001$, $\eta^2_p = 0.34$, $BF_{incl} > 1000$, and a significant main effect of *Adaptation Direction*, $F(1, 54) = 5.95$, $p = .018$, $\eta^2_p = 0.10$, $BF_{incl} > 1000$. Crucially, the *Frequency* * *Adaptation Direction* interaction was highly significant, $F(1, 54) = 238.38$, $p < .001$, $\eta^2_p = 0.82$, $BF_{incl} > 1000$, indicating a bias in sound localization due to the concurrently presented visual stimulus consistent with the respective AV spatial mismatch, i.e. reflecting the VE. The main effect of *Group* only showed a non-significant trend, $F(1, 54) = 3.80$, p = .056, $\eta^2_p = .07$, $BF_{incl} = 0.35$, indicating that patients ($M = 2.75$, $SD = 2.93$) and healthy controls ($M = 1.33$, $SD = 2.72$) displayed a non-significant trend towards localization error bias differences during AV trials. All other effects did not reach statistical

significance, all $p > .05$ and all $BF_{incl} < 1$ (see Figure 13 for localization responses in AV trials).

***Figure 13***

*Localization Error in Audiovisual Trials*



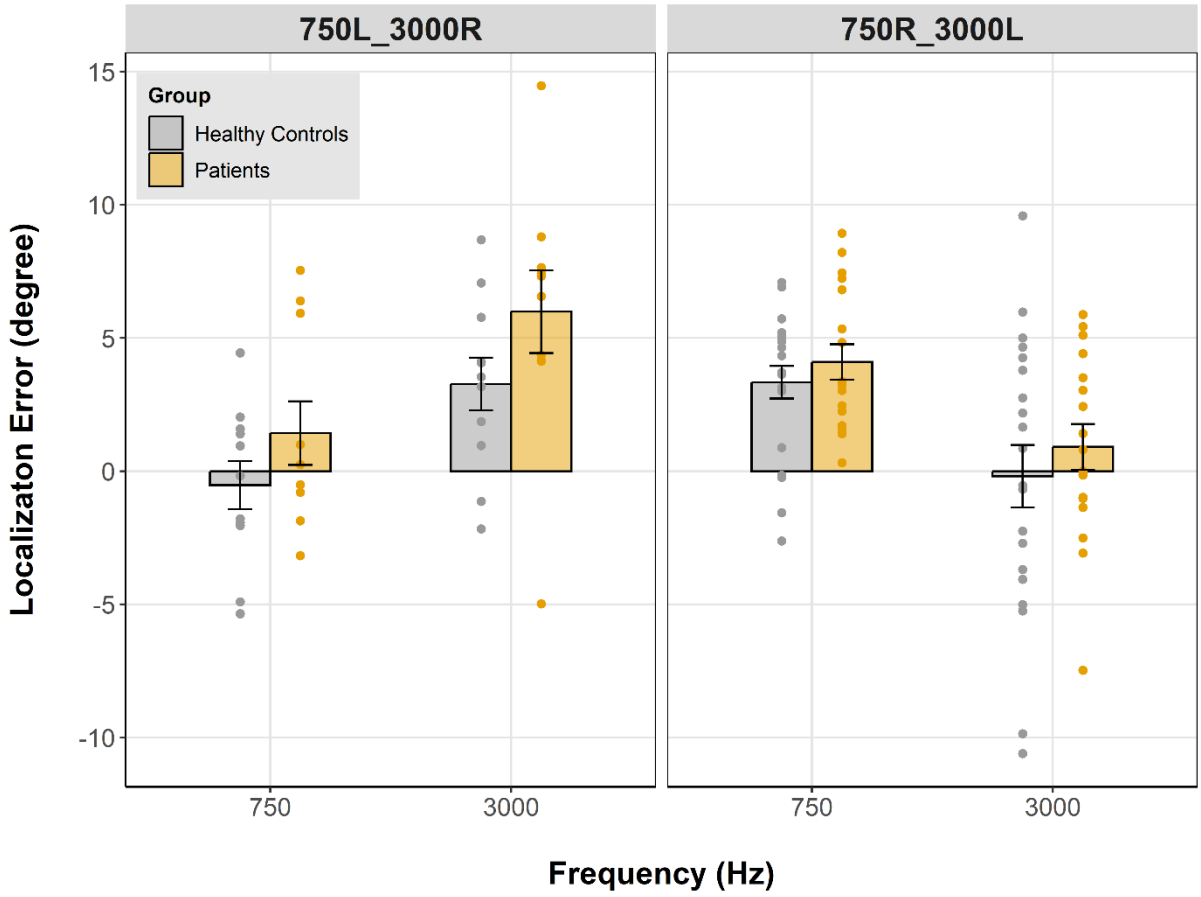

*Notes.* $n_{HC} = 31$; $n_{PAT} = 27$. The mean constant error in audiovisual trials is depicted per sound frequency (750Hz & 3000 Hz) and group for the non-merged data, i.e. separately for both adaptation directions (750 Hz leftward/3000 Hz rightward & 750 Hz rightward/3000 Hz leftward). Depicted are the mean, error bars denoting the SEM, and the individual data points.

**UA Trials.**[9] An exploratory mixed ANOVA with the between-factors *Group* (patients vs. healthy controls) and *Adaptation Direction* (750R/3000L vs. 750L/3000R) and the within-factors *Frequency* (750Hz vs. 3000Hz) and *Frequency of prior adaptation trail* (same vs. different) on mean constant errors in UA auditory trials revealed a highly significant *Frequency * Adaptation Direction* interaction, $F(1, 54) = 33.76$, $p < .001$, $\eta^2_p = 0.38$, indicating a bias in sound localization in UA trials induced by the respective AV spatial mismatch in AV adaptation trials, i.e. reflecting the cumulative VAE. There further was a highly significant *Frequency * Adaptation Direction * Frequency of prior adaptation trial* interaction, $F(1, 54) = 40.26$, $p < .001$, $\eta^2_p = 0.43$, indicating that the similarity in frequency and therefore the direction of the AV spatial mismatch of the previous adaptation trial in the immediate stimulus history had an impact on sound localization responses in UA trials, i.e. reflecting the immediate VAE. There also was a significant *Group * Adaptation Direction * Frequency of prior adaptation trial* interaction, $F(1, 54) = 4.64$, $p = .036$, $\eta^2_p = 0.08$. Two follow-up ANOVAs separately for each group revealed that the *Adaptation Direction * Frequency of prior adaptation trial* interaction did not remain significant for either group (healthy controls $p = .339$; patients $p = .071$). The main effect of *Group* only showed a non-significant trend, $F(1, 54) = 3.51$, $p = .066$, $\eta^2_p = .06$, implying that patients ($M = 2.95$, $SD = 2.95$) and healthy controls ($M = 1.50$, $SD = 2.86$) displayed a non-significant tendency towards localization error bias differences during UA trials. All other effects did not reach statistical significance, all $p > .05$ (see Figure 14 for localization responses in UA trials and Figure 15 for localization responses in UA trials split by same vs. different frequency of prior adaptation trial).

[9] The *Group * Adaptation Direction * Frequency * Frequency on of prior adaptation trial* Bayesian mixed ANOVA unexpectedly reported $BF_{incl} > 1000$ for some factors, which were non-significant in the frequentist analysis, all $p > .05$. We speculate that the current JASP version 0.16.4 is not yet built for a Bayesian ANOVA with two between- and two within-factors, possibly leading to computational errors due to the high amount of resulting candidate models. We assume that $BF_{incl}$ are erroneous and therefore do not report the results of above mentioned Bayesian mixed ANOVA.

***Figure 14***

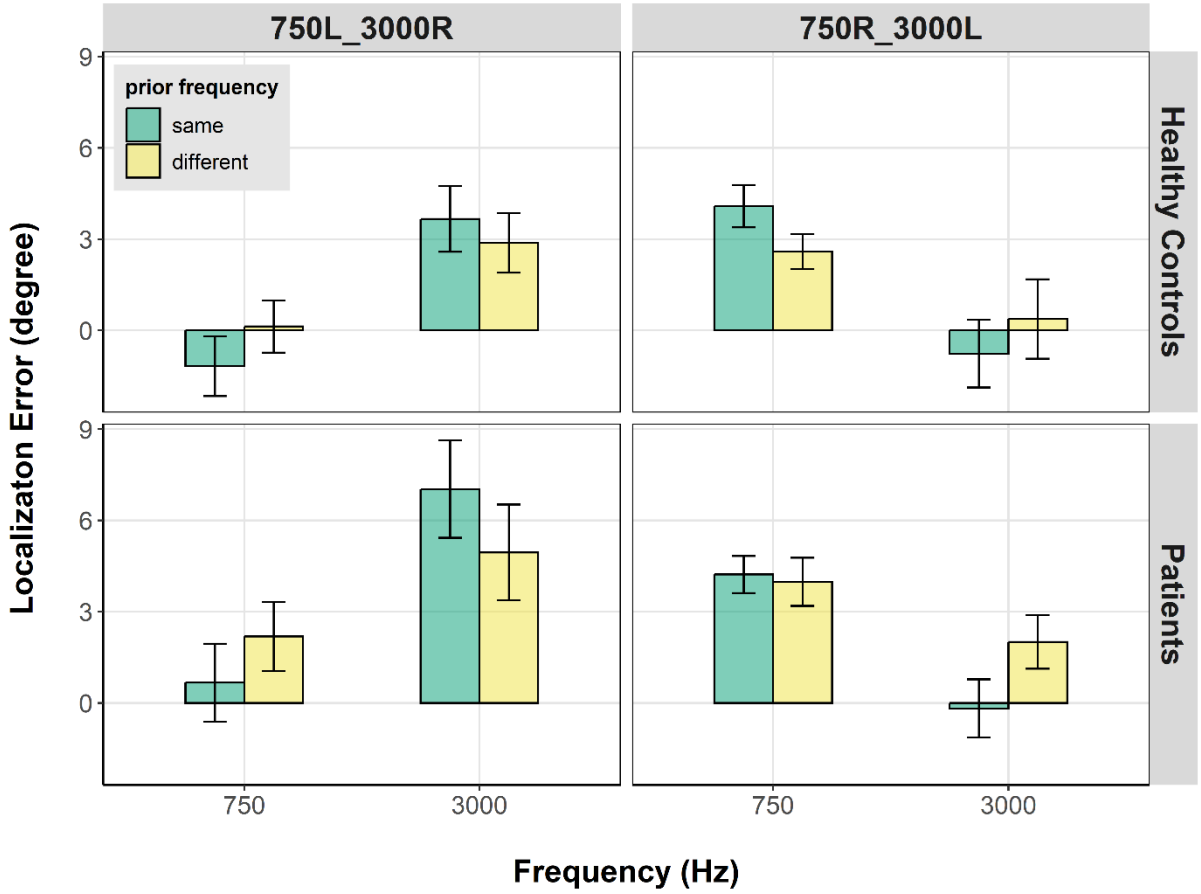
*Localization Error in Unimodal Auditory Trials*




*Notes.* $n_{HC}$ = 31; $n_{PAT}$ = 27. The mean constant error in unimodal auditory trials is depicted per sound frequency (750Hz & 3000 Hz) and group for the non-merged data, i.e. separately for both adaptation directions (750 Hz leftward/3000 Hz rightward & 750 Hz rightward/3000 Hz leftward). Depicted are the mean, error bars denoting the SEM, and the individual data points.

***Figure 15***

*Localization Error in Unimodal Auditory Trials Following an Audiovisual Trial with the Same or Different Frequency*




*Notes.* $n_{HC}$ = 31; $n_{PAT}$ = 27. The mean constant error in unimodal auditory trials following an audiovisual trial of the same or different frequency is depicted per group and per sound frequency (750Hz & 3000 Hz) for the non-merged data, i.e. separately for both adaptation directions (750 Hz leftward/3000 Hz rightward & 750 Hz rightward/3000 Hz leftward). Depicted are the mean and error bars denoting the SEM.

## 3.2 Control Variables and Questionnaire Data

### *3.2.1 Control Tests for Attention, Processing Speed, and Vigilance*

**TMT.** Patients ($M$ = 31.05, $SD$ = 11.06) and healthy controls ($M$ = 26.13, $SD$ = 9.11) showed a non-significant trend towards TMT-A completion time (in seconds) differences, $t(50.52) = 1.83$, $p = .073$, Cohen's $d = 0.49$, $BF_{10} = 1.1$. Groups differed significantly in TMT-B completion time (in seconds), with patients needing more time to complete the test ($M$ = 84.84, $SD$ = 40.51) compared to healthy controls ($M$ = 56.37, $SD$ = 20.44), $t(37.23) = 3.3$, $p$ = .002, Cohen's $d = 0.91$, $BF_{10} = 28.87$.

**CPT-IP.** A mixed ANOVA for CPT-IP *d'* scores revealed a significant main effect of *Group*, $F(1, 112) = 9.03$, $p = .004$, $\eta^2_p = 0.14$, $BF_{incl} = 10.46$, indicating that patients showed overall lower d' scores ($M = 2.32$, $SD = 1.15$) compared to healthy controls ($M = 2.81$, $SD = 1.14$). Further, there was a highly significant main effect of *Digit Load*, $F(2, 112) = 252.38$, $p < .001$, $\eta^2_p = 0.82$, $BF_{incl} > 1000$. Bonferroni-corrected post-hoc t-tests revealed that d' scores were larger in the 2-digit condition ($M = 3.55$, $SD = 0.65$) compared to both the 3-digit condition ($M = 2.74$, $SD = 0.88$), $t(57) = 9.23$, $p < .001$, and the 4-digit condition ($M = 1.46$, $SD = 0.81$), $t(57) = 22.31$, $p < .001$, and larger in the 3-digit condition compared to the 4-digit condition, $t(57) = 13.09$, $p < .001$. There was no significant *Group* * *Digit condition* interaction, $F(2, 112) = 2.06$, $p = .132$, $\eta^2_p = 0.04$, $BF_{incl} = 1.79$.

**Association Between Experimental Data and Control Variables.** VE, cumulative VAE, and immediate VAE data did not show any significant Spearman rank correlations with TMT and CPT-IP data, all Spearman's $\rho \leq .31$ and all $p > .05$ (see Table D.3 in Appendix D for Spearman rank correlation coefficients $\rho$).

### *3.2.2 Questionnaires*

**CAPS.** Groups differed significantly in CAPS scores, with patients reporting significantly higher number of aberrant perceptual experiences ($M = 9.9$, $SD = 6.7$) compared to healthy controls ($M = 2.2$, $SD = 2.4$), $t(31.85) = 5.95$, $p < .001$, Cohen's $d = 1.55$, $BF_{10} > 1000$.

**LSHS.** Groups differed significantly in LSHS-E total scores, with patients reporting significantly higher occurrence of hallucinations ($M = 22.3$, $SD = 11.3$) compared to healthy controls ($M = 4.4$, $SD = 4.4$), $t(32.86) = 7.76$, $p < .001$, Cohen's $d = 2.15$, $BF_{10} > 1000$.

**PCL.** Groups differed significantly in PCL total scores, with patients reporting significantly higher levels of paranoia ($M = 81.7$, $SD = 56.3$) compared to healthy controls ($M = 21.3$, $SD = 18.7$), $t(30.98) = 5.32$, $p < .001$, Cohen's $d = 1.48$, $BF_{10} > 1000$.

**Passivity Experiences.** Groups differed significantly in passivity experiences scores, with patients reporting significantly higher occurrence of passivity experiences ($M = 5.5$, $SD = 4.4$) compared to healthy controls ($M = 0.3$, $SD = 0.5$), $t(26.62) = 6.19$, $p < .001$, Cohen's $d = 1.75$, $BF_{10} > 1000$.

**BCSS.** Groups differed significantly in BCSS negative self scores, with patients reporting significantly more negative beliefs about the self ($M = 6.9$, $SD = 4.9$) compared to

healthy controls ($M$ = 2.0, $SD$ = 2.1), $t$(33.96) = 4.91, $p$ < .001, Cohen's $d$ = 1.35, $BF_{10}$ = 610. Further, groups differed significantly in BCSS negative other scores, with patients reporting significantly more negative beliefs about other people ($M$ = 7.9, $SD$ = 6.1) compared to healthy controls ($M$ = 3.3, $SD$ = 3.9), $t$(42.58) = 3.35, $p$ = .002, Cohen's $d$ = 0.91, $BF_{10}$ = 29.91.

**Correlations Between Questionnaire Scores.** The majority of Spearman rank correlations between CAPS, LSHS-E, PCL, passivity experiences and BCSS negative subscales was highly significant, all Spearman's $\rho$ > .40, all $p$ < .001, all $BF_{10}$ > 15 [10], except for the correlation between CAPS and the BCSS negative other subscale, Spearman's $\rho$ = .39, being significant at $p$ = .03, $BF_{10}$ = 14.71. $p$-values are Bonferroni-corrected (see Table 6 for Spearman rank coefficients $\rho$. Exploratory Spearman rank coefficients $\rho$ split by group can be found in Table D.1 and Table D.2 in Appendix D).

### 3.3 Correlations Between Experimental Data and Questionnaire Scores

There were no significant Spearman rank correlations between VE, cumulative VAE as well as immediate VAE and questionnaire data, all Spearman's $\rho$ < .3, all $p$ > .05, all $BF_{10}$ < 1. $p$-values are Bonferroni-corrected (see Table 6 for Spearman rank coefficients $\rho$. Exploratory Spearman rank coefficients $\rho$ split by group can be found in Table D.1 and Table D.2 in Appendix D).

## 4 Discussion

In this study, we investigated spatial MSI and crossmodal spatial recalibration in psychosis. Only few published studies examined spatial MSI in psychosis with their results suggesting intact spatial MSI (de Gelder et al., 2003; Noel et al., 2020; Stephen et al., 2013; Stone et al., 2011, 2014). Previous studies on spatial recalibration in psychosis exclusively focused on sensorimotor processes (e.g. Bartolomeo et al., 2020; Ferri et al., 2016; Rösler et al., 2015) and to our knowledge no study investigated crossmodal, i.e. sensory-sensory spatial recalibration in psychosis. Thus, we aimed to replicate earlier findings on intact spatial MSI in psychosis (de Gelder et al., 2003; Noel et al., 2020; Stephen et al., 2013; Stone et al., 2011, 2014) and extend findings on spatial recalibration in psychosis to AV stimuli.

[10] The majority of $BF_{10}$ was > 1000, with following exceptions: LSHS ~ BCSS negative other, $BF_{10}$ = 289.61, passivity experiences ~ BCSS negative self, $BF_{10}$ = 207.42, and passivity experiences ~ BCSS negative other, $BF_{10}$ = 17.08.

Patients with the diagnosis of a psychotic disorder and healthy controls completed a VE/VAE paradigm with pseudorandomly alternating UA and AV trials (paradigm adopted from Bruns & Röder, 2015). In each trial, participants localized a sound originating from one of six different speaker positions by means of a hand pointer. In AV trials, the visual stimulus was presented with a fixed spatial disparity either to the left or the right of the sound depending on sound frequency and had to be ignored. Between both groups, we compared the difference in mean deviance of sound localization responses from the veridical sound position between leftward- and rightward-adaptation AV trials, reflecting the VE. Further, we compared the difference in mean deviance of sound localization from the true sound position between leftward- and rightward-adapted UA trials, reflecting the cumulative VAE, and between UA trials immediately preceded by a leftward- vs. rightward-adaptation AV trial, reflecting the immediate VAE.

Our data analysis revealed that patients and controls showed a comparable mean deviance of sound localization responses from the veridical sound position in all three of the above described analyses. This indicates similar sizes of the VE, cumulative as well as immediate VAE in both groups. In the following, these results will be discussed in more detail with regard to our hypotheses and previous evidence on spatial multisensory processing in psychosis.

### 4.1 Hypothesis 1: Psychosis and Spatial Ventriloquism

Our first aim was to replicate previous findings of an intact spatial VE in psychosis (de Gelder et al., 2003). Thus, we expected patients with the diagnosis of a psychotic disorder to show a comparable VE size compared to healthy controls.

In line with our hypothesis, the difference in mean constant error during leftward- vs. rightward-adaptation AV trials did not differ between both groups, reflecting a comparable size of the VE between patients and healthy controls. Further, both groups individually showed a significant VE as indicated by group-specific *t*-tests against zero. Thus, our results suggest that our patient sample showed intact spatial MSI in our VE paradigm, which confirmed our hypothesis.

Our results are consistent with an earlier study by de Gelder et al. (2003), who observed an intact VE using short sequences of AV stimuli. By applying a more typical VE paradigm with single AV stimulus pairs to avoid within-trial recalibration processes (Bruns, 2019), we were able to replicate the findings of de Gelder et al. (2003) in patients with the

diagnosis of a psychotic disorder. Further, our results are also in line with studies reporting intact spatial MSI of visuo-tactile stimuli (Noel et al., 2020) and in AV far-near judgements (Stephen et al., 2013; Stone et al., 2011, 2014) in psychosis. Crucially, unisensory far-near judgement performance was impaired in auditory and visual only trials, suggesting crossmodal compensatory mechanisms for unisensory deficits in patients with the diagnosis of a psychotic disorder (Stephen et al., 2013; Stone et al., 2011, 2014). Alternatively, this could reflect a consequence of greater attentional capture of crossmodal compared to unimodal stimuli, which might improve sensory processing (Tseng et al., 2015).

Importantly, although previous studies on AV far-near judgements in psychosis did not observe multisensory behavioral impairments, these studies reported structural and neurophysiological differences between patients with the diagnosis of a psychotic disorder and healthy controls (Stephen et al., 2013; Stone et al., 2011, 2014). By using a joint MEG and DTI analysis strategy, Stephen et al. (2013) identified neurophysiological and structural alterations in visual processing pathways in patients compared to controls, such as reduced occipital MEG amplitude in response to visual stimulation and lower fractional anisotropy in parietal regions, suggesting visual processing impairments in psychosis. These alterations were associated with overall RT and measures of cognitive functioning, but not however with multisensory performance (Stephen et al., 2013). Stone et al. (2014) reported altered gamma-band oscillations in frontal regions as well as in visual and auditory cortices during unimodal and AV processing in patients compared to controls, which however were not correlated to increased behavioral crossmodal gain in the patient sample. In an earlier study, Stone et al. (2011) observed group differences between patients with the diagnosis of a psychotic disorder and healthy controls in a ERP comparison between AV and unimodal trials over left occipital-temporal regions, reflecting greater neural crossmodal gain in patients compared to controls. In sum, structural and neurophysiological changes during an AV far-near judgement task in psychosis were either not related to behavioral multisensory performance or reflected greater crossmodal gain and thus mirrored behavioral findings of similar multisensory performance between patients and controls (Stephen et al., 2013; Stone et al., 2011, 2014).

Our findings indicate intact spatial MSI in psychosis in a VE paradigm with simple AV stimuli. While this is consistent with previous findings suggesting intact spatial MSI in patients with the diagnosis of a psychotic disorder (de Gelder et al., 2003; Noel et al., 2020; Stephen et al., 2013; Stone et al., 2011, 2014), previous evidence in the temporal domain indicated impaired temporal MSI in psychosis (Zhou et al., 2018). These contrasting findings

in the spatial vs. in temporal domain might offer indirect support for Odegaard's & Shams' (2016) findings in healthy participants, which suggested that spatial and temporal MSI might reflect dissociated processes. In sum, our findings suggest that the integration of spatial crossmodal information might be unimpaired in psychosis, contrary to findings in the linguistic, emotional (Lin et al., 2020; Tseng et al., 2015), temporal (Zhou et al., 2018) or sensorimotor domain (Noel, Cascio, et al., 2017). This in turn might offer support for the proposition of domain-specific rather than global deficits in MSI in psychosis (Tseng et al., 2015).

## 4.2 Hypothesis 2: Psychosis and the Spatial VAE

### *4.2.1 Discussion and Comparison to Previous Evidence*

Our second aim was to provide first experimental evidence for crossmodal, i.e. sensory-sensory spatial recalibration in psychosis. Previous studies on spatial recalibration in psychosis exclusively examined sensorimotor processes and reported impairments in e.g. visuomotor learning processes (Bartolomeo et al., 2020; Rösler et al., 2015). We aimed to extend those findings to the sensory-sensory domain to investigate if impairments of spatial recalibration in psychosis can also be observed in AV stimuli. We examined both cumulative and immediate spatial recalibration in psychosis to obtain a more fine-grained picture of potential alterations in AV spatial recalibration.

The difference in mean constant error during leftward- vs. rightward-adapted UA trials did not differ between patients and controls, reflecting a similar size of the cumulative VAE in both groups. Further, patients and controls did also not differ in the difference in mean constant error during AV trials immediately preceded by a leftward- vs. rightward-adaptation AV trial, reflecting a comparable size of the immediate VAE in both groups. Further, separate *t*-tests against zero confirmed that both groups individually showed a cumulative VAE and immediate VAE. This indicates that our sample of patients with the diagnosis of a psychotic disorder showed intact cumulative and immediate AV spatial recalibration.

Our findings do not support our hypothesis on altered cumulative spatial recalibration in patients with the diagnosis of a psychotic disorder. This indicates that both our patients and control subjects were able to recalibrate their auditory cortical maps to repeated exposure to consistent AV spatial discrepancies (Chen & Vroomen, 2013). In contrast, our results on the immediate VAE confirm our hypothesis on immediate spatial recalibration in psychosis. This indicates that both groups were able to rapidly adapt sound localization to AV spatial

conflicts. In light of previous evidence indicating that spatial MSI and immediate spatial recalibration might be closely related and based on common underlying neural substrates (Bruns et al., 2022; Park & Kayser, 2019; Rohlf et al., 2020), this finding is consistent with our results on an intact VE in our patient sample. These results could suggest that the perceptual and cognitive processes underlying both the VE and the immediate VAE might be relatively spared in patients with the diagnosis of a psychotic disorder. Crucially, further research aiming to replicate our findings are needed to investigate this assumption.

Given the lack of evidence on sensory-sensory spatial recalibration in psychosis, comparing our findings to previous evidence on sensorimotor spatial recalibration might give hints for contrasting sensory-sensory and sensorimotor processes in psychosis and understanding their potential respective role in disorder development.

Bartolomeo et al. (2020) administered a prism adaptation paradigm in psychosis and observed reduced visuomotor recalibration during and after prism application when reaching for a target in patients with the diagnosis of a psychotic disorder compared to healthy controls. Importantly, visuomotor recalibration deficits were not related to motor dysfunction and deficits were not observed during baseline, suggesting a specific impairment in integrating and recalibrating visual and proprioceptive input (Bartolomeo et al., 2020). Further, Rösler et al. (2015) observed impaired visuomotor recalibration of eye movements in patients with the diagnosis of a psychotic disorder in a saccadic remapping paradigm. Crucially, impaired recalibration was larger in patients with higher severity of psychotic symptoms (Rösler et al., 2015), which might indicate a possible link between visuomotor recalibration and psychosis.

Deficits in visuomotor recalibration processes in psychosis such as saccadic remapping impairments might be rooted in deficits in predicting the sensory consequences of own behavior (for a review, see Thakkar & Rolfs, 2019). This hypothesis is based on previous work on corollary discharge, a cognitive process describing how efference copies of motor commands and sensory input of own actions are integrated in the respective sensory brain areas, leading to a sense of agency over own actions. By comparing the efference copy and the sensory input of the action, the brain is able to explain the sensory input and attenuates its neural signal. It has been discussed that this process enables the brain to monitor own behaviour, predict its outcome and by this infer if actions were self-induced (Bansal et al., 2018; Thakkar et al., 2017).

It has been discussed that psychosis might be associated with impaired signaling of motor command efference copies, in turn leading to overly high salience of the sensory consequences of own behaviour, such as seeing the movement of one's hand (Bansal et al., 2018; Poletti et al., 2019). The impaired attenuation of sensory signals of own actions might lead to an impaired sense of agency over own actions and it has been found that impairments in corollary discharge are associated with psychotic symptoms (Bansal et al., 2018). It has been discussed that impaired corollary discharge might impair the distinction between self- and externally caused actions in psychosis, drastically influencing the experience of self-agency or body ownership (Moberget & Ivry, 2019; Poletti et al., 2019).

Impaired corollary discharge as a basis for altered sense of agency and body ownership in psychosis might offer a potential model for psychotic symptoms such as delusions of control or auditory verbal hallucinations (Moberget & Ivry, 2019; Poletti et al., 2019; Thakkar & Rolfs, 2019). While this indicates that model attempts for linking altered multisensory processing and psychosis might be feasible for sensorimotor processes (see **4.1 A Lack of Theoretical Models on Multisensory Processes in Psychosis** in the **General Introduction**), this evidently is specific for the sensorimotor domain and not suitable for multisensory processes in general. Nonetheless, this dysfunction in perceptual prediction can be described within the predictive coding account of psychosis (Sterzer et al., 2018). This account might provide a means to embed altered sensorimotor processing and its potential role for the development or upkeep of psychotic symptoms into a greater framework, which might be suitable to consider various dysfunctions of multisensory processing in psychosis (see **5.1 Causal Inference** in the **General Discussion)**.

In sum, previous research indicated deficits in sensorimotor spatial processing and recalibration in psychosis, which might play a role in disorder development (Bartolomeo et al., 2020; Ferri et al., 2016; Noel, Cascio, et al., 2017; Rösler et al., 2015). Previous evidence offered theoretical links between sensorimotor processing deficits and psychotic symptoms such as passivity delusions and hallucinations (Moberget & Ivry, 2019; Poletti et al., 2019; Thakkar & Rolfs, 2019). In contrast, our results on both an intact cumulative and immediate VAE suggest that sensory-sensory spatial recalibration might be unimpaired in psychosis. One could speculate that spatial recalibration is specifically impaired in processes, which involve proprioceptive information, i.e. stimuli involving the own body and behaviour, but intact in non-proprioceptive and purely external stimuli. This might suggest that sensory-sensory spatial recalibration might not be related to the development of psychotic symptoms

or their maintenance. Evidently, further research on crossmodal spatial recalibration in psychosis is needed to confirm or reject this hypothesis.

### *4.2.2 Associations between the VE, the cumulative, and the immediate VAE*

We observed significant correlations between the VE and both the cumulative and immediate VAE over both groups. This is partly consistent with previous studies on associations between spatial MSI and crossmodal spatial recalibration (see also **1.1 Temporal Multisensory Processing** in **Chapter III**). Both Bruns et al. (2022) and Rohlf et al. (2020) observed correlations between spatial MSI and immediate recalibration, but not with cumulative recalibration. Further, Park & Kayser (2019) reported shared neural substrates in the medial parietal cortex between spatial MSI and immediate recalibration. Thus, our findings of a significant correlation between the VE and immediate VAE match previous evidence, suggesting that spatial MSI and immediate spatial recalibration might be related processes. While previous work suggested that spatial MSI and cumulative spatial recalibration might be less closely related (Bruns et al., 2022; Rohlf et al., 2020), this does not indicate that both processes are independent. It has been proposed that an appropriately integrated multisensory percept might be crucial for the development of recalibration processes (Rohlf et al., 2020), suggesting at least a partial association between spatial MSI and cumulative spatial recalibration. Thus, it could be argued that the correlation between the VE and the cumulative VAE in our data might match the proposition of Rohlf et al. (2020).

Further, we did not observe a correlation between the cumulative and immediate VAE over both groups, indicating that the cumulative VEA size was not related to the size of the immediate VAE. This is consistent with the findings of Bruns & Röder (2015), who also reported no correlation between cumulative and immediate spatial recalibration. Thus, our findings corroborate the proposition of Bruns & Röder (2015) that the cumulative VAE and the immediate VAE might reflect partially different underlying perceptual processes.

### *4.2.3 Accumulation of the Immediate VAE*

We further explored whether groups differed in the accumulation of the immediate VAE as a function of consecutive trials with one sound frequency (and direction of AV spatial disparity), which could give hints if the recalibration rate for immediate recalibration might be different between patients with the diagnosis of a psychotic disorder and healthy controls.

For the accumulation of the immediate VAE as a function of consecutive trials with the same frequency, the immediate VAE descriptively increased in patients but remained

stable in healthy controls. However, we did not observe a significant interaction between the factors group and consecutive trial number of the same frequency. Further, post-hoc tests over both groups indicated that the immediate VAE did not change as a function of consecutive same frequency adaptation trials.

For the accumulation of the immediate VAE as a function of consecutive trials with the different frequency, the immediate VAE decreased with increasing trial number in both groups. Further, the immediate VAE descriptively reached a minimum value in healthy controls after two to three consecutive trials and remained stable, while it further decreased in patients. However, we did not observe an interaction between group and consecutive trial number. Post-hoc tests over both groups confirmed that the immediate VAE decreased with increasing different frequency trials. Thus, it can be assumed that the immediate VAE size reduction as a function of consecutive different frequency adaptation trials was similar in patients and healthy controls.

Overall, our findings on accumulation of spatial recalibration as a function of the recent stimulus history are consistent with the findings of the study by Bruns & Röder (2015), from which we adopted the paradigm. Analogue to our findings, Bruns & Röder (2015) observed that spatial recalibration decreased with consecutive different frequency adaptation trials but was not modulated by the number of consecutive same frequency adaptation trials. Thus, our findings corroborate the authors' proposition that immediate recalibration might be primarily impacted by a mismatch between the current spatial conflict and the orientation of crossmodal spatial disparities in the recent stimulus history rather than a match thereof (Bruns & Röder, 2015).

In sum, our patients and controls did not differ in the accumulation of crossmodal spatial recalibration depending on the recent stimulus history. This suggests that the AV spatial recalibration rate over short time periods might be similar in patients with the diagnosis and healthy controls. Interestingly, this is in contrast to evidence on a reduced adaptation learning rate in visuomotor learning in psychosis (Lencer et al., 2017), further supporting the notion of intact spatial in contrast to impaired sensorimotor multisensory processing in psychosis.

### *4.2.4 Sound Localization Biases in Psychosis*

Our exploratory analysis of the non-merged data (see **2.4 Data Analysis** for a description of the data merging procedure) largely mirrored the findings of the merged data, indicating a comparable VE, cumulative and immediate VAE in both groups. However, both in UA and in AV trials, a non-significant trend for localization bias differences indicated that patients showed a trend to slightly localize sound stimuli more to the right than healthy controls, independent of sound frequency and adaptation.

Only few studies investigated sound localization in psychosis. Nevertheless, localization impairments have been reported, with patients with the diagnosis of a psychotic disorder showing reduced sound localization accuracy especially in the right hemifield (Perrin et al., 2010; Sardari et al., 2022). Thus, the non-significant trend for a slightly larger rightward-bias in our patient sample matches previous findings. However, Sardari et al. (2022) reported that patients with the diagnosis of a psychotic disorder showed a larger leftward-bias than healthy controls, thus reporting a bias in the opposite direction compared to our findings. Our VE/VAE paradigm does not represent a pure sound localization task, since sound stimuli are either jointly presented with a visual stimulus or influenced by a visual stimulus in the recent stimulus history. Thus, the comparability of our findings to pure unimodal sound localization studies (Perrin et al., 2010; Sardari et al., 2022) might be limited. While our findings are consistent with previous studies and suggest that psychosis might be associated with sound localization biases, further studies are needed to investigate the extent, variability, and direction of localization biases in patients with the diagnosis of a psychotic disorder.

Despite only reflecting a non-significant trend, we cannot rule out that a potential rightward-bias in sound localization influenced spatial MSI and AV spatial recalibration in our patient sample. However, the descriptive difference over all trials between patients and controls was rather small, i.e. approximately 1.5 degrees of visual angle. Crucially, our patient sample showed comparable localization as indicated by our pre-analysis check for localization responses relative to speaker location (see ***3.1.1 Localization Responses Relative to Speaker Positions***). Thus, it could be argued that a possible influence of this potential, small rightward-bias on AV spatial processing in our study might be relatively negligible.

## 4.3 Hypothesis 3: Associations between Spatial MSI, Spatial Recalibration and Psychotic Symptoms

Our final study aim was to examine associations between spatial MSI as well as crossmodal spatial recalibration and self-reported psychotic symptoms. In light of the scarcity of studies on spatial MSI and the lack of studies on sensory-sensory spatial recalibration in psychosis, no clear expectations regarding associations or direction of possible correlations between spatial MSI, crossmodal spatial recalibration and psychotic symptoms could be drawn from the literature. We therefore explored potential correlations between our behavioral and questionnaire data.

We did not observe any significant correlation between the size of the VE, the cumulative VAE, or the immediate VAE and self-reports on anomalous perceptual experiences, hallucinations, paranoia, passivity experiences and negative schemata. Thus, our findings indicate that spatial MSI and crossmodal spatial recalibration might not be associated to psychotic symptoms and negative schemata about the self and others.

Previous studies on spatial MSI in psychosis only rarely conducted correlational analyses between measures of spatial MSI and psychotic symptoms. Nevertheless, our findings are consistent with a study by Stone et al. (2011) reporting no correlations between AV far-near judgement performance and negative or positive symptoms. This is further supported by a study by Stephen et al. (2013), who did not observe associations between neurophysiological as well as structural measures during an AV far-near judgement paradigm and positive or negative symptoms. Other studies on spatial MSI in psychosis did not report correlation data between measures for MSI or neurophysiological measurements during spatial MSI and psychotic symptoms (de Gelder et al., 2003; Noel et al., 2020; Stone et al., 2014). Taken together, our findings are in line with findings of Stephen et al. (2013) and Stone et al. (2011) and offer initial evidence suggesting that spatial MSI and psychotic symptoms might not be linked to each other.

In contrast to our findings and previous studies on spatial MSI in psychosis, AV target detection studies reported correlations between AV stimulus detection and negative but not positive symptoms (Williams et al., 2010; Wynn et al., 2014), suggesting a potential link between multisensory processing and negative symptoms of psychotic disorders. It is possible that we failed to observe correlations between our behavioral and questionnaire data, since we exclusively focused on positive rather than negative symptoms. Future studies on spatial

multisensory processing in psychosis using a VE/VAE paradigm should investigate if any association between spatial ventriloquism as well as its aftereffects and negative symptoms can be observed. However, the lack of correlations between AV far-near judgements and negative symptoms in the studies by Stephen et al. (2013) and Stone et al. (2011) rather suggested that there might be no direct association between spatial multisensory processing and negative symptoms of psychosis.

It is possible that recruiting a larger sample size compared to our study might have been fruitful in observing statistical significant correlations between the VE, cumulative VAE, immediate VAE and psychotic symptoms. However, our sample size was either comparable (Stephen et al., 2013) or larger (Stone et al., 2011) than previous studies reporting no association between spatial MSI and psychotic symptoms. Further, since no Spearman rank correlation coefficient $\rho$ approached a potentially meaningful effect size $\rho > .3$ (see **2.4 Data Analysis**), we argue that larger samples might not significantly change our observed correlation pattern. Thus, we conclude that our findings suggest that spatial multisensory processing, i.e. both spatial integration and crossmodal spatial recalibration, might not be linked to psychotic symptoms. Crucially, this conclusion should be drawn cautiously, since evidence on associations between spatial MSI and psychotic symptoms is very scarce and since we are the first to report correlation data between psychotic symptoms and crossmodal spatial recalibration.

### 4.4 Limitations

In addition to potential study limitations mentioned above, i.e. a potential overall rightward-bias in sound localization in the patient group and not having investigated potential associations between AV spatial processing and negative symptoms of psychosis, the following possible limitations of our study should be addressed.

Our patient and controls groups differed significantly in age, with patients having a higher mean age compared to controls. This age difference likely impacted performance in the TMT and CPT-IP, since performance in both tests has been found to decrease with age, indicating a decline in processing speed and vigilance over the lifespan (Mani et al., 2005; Rodewald et al., 2012; Tombaugh, 2004). However, it is unlikely that the age difference influenced behavioral performance in our experimental paradigm, since previous evidence suggested that multisensory processing might be rather stable over the adult lifespan and largely sustained in older adults (Jones & Noppeney, 2021). Further, since we did not observe

group differences in spatial MSI and crossmodal recalibration, we argue that the age difference might have had limited statistical impact on our behavioral findings.

We observed significant group differences in processing speed and vigilance, reflected by performance in the TMT and CPT-IP. This indicated that our patient sample had an overall lower processing speed and vigilance compared to healthy controls, which is consistent with reports on reduced processing speed and vigilance in psychosis (Gebreegziabhere et al., 2022; Mesholam-Gately et al., 2009). Crucially, vigilance did however not decrease faster in patients than controls. This raises the question if lower processing speed and vigilance influenced the behavioral performance of our patient sample in our VE/VAE paradigm. However, we did not observe any association between performance in the TMT as well as the CPT-IP and the VE, cumulative VAE or immediate VAE, suggesting that the performance in our paradigm was not influenced by group differences in processing speed and vigilance as measured by the TMT and CPT-IP.

Due to reasons of restricting experimental session duration, we did not include unimodal auditory or visual localization pretest blocks. Thus, we were not able to obtain estimates of individual localization performance of sound or flash stimuli independent of crossmodal stimulus presentation. It could have been possible that differences in unisensory spatial processing in patients compared to controls, which have previously been reported in psychosis (Hardoy et al., 2004; Perrin et al., 2010; Sardari et al., 2022), impacted spatial multisensory processing and sound localization performance in our VE/VAE paradigm. Further, if unisensory localization differences had been revealed, this would have complicated the interpretation of possible group effects on spatial MSI or crossmodal spatial recalibration. However, we did not observe any group difference in spatial MSI and crossmodal recalibration. Further, our pre-analysis localization performance review procedure (see ***3.1.1 Localization Responses Relative to Speaker Positions***) and the exploratory analysis on the non-merged data provide estimates of comparable sound localization between our patient and control samples. Thus, we argue that our behavioral data might not have been influenced by potential unisensory localization differences between our groups.

## 4.5 Conclusion

To conclude, our study on spatial MSI and crossmodal spatial recalibration in psychosis indicates that our sample of patients with the diagnosis of a psychotic disorder showed both intact spatial ventriloquism and cumulative as well as immediate spatial recalibration. Although we cannot rule out sampling effects, i.e. having recruited a potentially high-functioning outpatient sample (for a comparable argument in our temporal multisensory processing study, see ***4.3.1 Associations between the TBW and Psychotic Symptoms*** in **Chapter III**), we argue that spatial multisensory processing in simple AV stimuli might be unimpaired in psychosis, since previous studies on spatial MSI already observed intact spatial MSI as well (de Gelder et al., 2003; Noel et al., 2020; Stephen et al., 2013; Stone et al., 2011, 2014). Thus, we successfully replicated previous work on spatial MSI in psychosis and extended these findings to AV spatial recalibration. Crucially, our study is the first to report intact cumulative and immediate AV spatial recalibration in psychosis. Therefore, more studies are needed to investigate if our findings can be replicated, if they are generalizable to other stimulus combinations such as visuo-tactile or audio-tactile recalibration and thus to contribute to the overarching research question of globality vs. domain-specificity of deficient multisensory processes in psychosis.

# Chapter V: General Discussion

In this dissertation project, we investigated multisensory processes in psychosis. The aim of the project was to study which exact processes of MSI and crossmodal recalibration are impaired in psychosis. This research aim was based on the initial question if dysfunctions in multisensory processes are at the root of the marked alterations in perception, which have been previously reported in psychosis (Javitt & Freedman, 2015; Uhlhaas & Mishara, 2006). To investigate multisensory processes in psychosis, we conducted three different experimental studies and compared subjects with and without psychotic symptoms. In order to examine various domains of MSI in psychosis, we investigated emotional (study 1), temporal (study 2), and spatial (study 3) MSI. Further, we examined crossmodal temporal (study 2) and spatial (study 3) recalibration in psychosis to provide first evidence for crossmodal learning processes in psychosis.

In our first study, we aimed to investigated if findings on altered emotional MSI previously shown in patients with the diagnosis of a psychotic disorder (Lin et al., 2020; Tseng et al., 2015) could already be observed in psychosis proneness. We administered a paradigm on emotional MSI using emotional face expressions and vocal prosody (adopted from Föcker et al., 2011) in healthy subjects with high vs. low psychosis proneness. Contrary to our expectation, high and low proneness subjects showed comparable performance in categorizing the perceived emotion of facial expressions and vocal prosody. This could be observed in unimodal, bimodal emotionally congruent, and bimodal emotionally incongruent stimulus conditions. This indicates that our sample with high psychosis proneness showed typical emotional MSI.

In our second study, we aimed to replicate previous findings on impaired temporal MSI in patients with the diagnosis of a psychotic disorder (Zhou et al., 2018). Further, we investigated if impairments in temporal multisensory processes in psychosis extend to crossmodal temporal recalibration. Patients with the diagnosis of a psychotic disorder and healthy controls completed a SJ paradigm with simple AV stimuli and interleaved adaptation phases to induce crossmodal recalibration (adapted from Van der Burg et al., 2015). Contrary to our expectation, patients showed a comparable TBW compared to controls. This indicates unimpaired temporal MSI in our patient sample, which contrasts previous evidence. Further, patients and healthy controls showed similar cumulative and immediate temporal recalibration effects. Thus, our findings indicate intact crossmodal temporal recalibration in our patient sample.

In our third study, we aimed to replicate previous findings on an intact spatial VE in psychosis (de Gelder et al., 2003). Further, we examined if previous findings on deficient crossmodal spatial recalibration observed in the sensorimotor domain (e.g. Bartolomeo et al., 2020; Rösler et al., 2015) can also be found in sensory-sensory, i.e. AV stimuli. Patients with the diagnosis of a psychotic disorder and healthy controls completed a paradigm on the VE and VEA with simple AV stimuli (adapted from Bruns & Röder, 2015). Patients and healthy controls showed a similar size of the VE. This indicates intact spatial MSI in our patient sample, replicating the findings of de Gelder et al. (2003). Further, both groups also showed a comparable size of the cumulative and immediate VAE. This indicates intact crossmodal spatial recalibration in our patient sample, providing first findings on typical AV spatial recalibration processes in psychosis.

## 1 Are Multisensory Integration Deficits in Psychosis Global or Domain-Specific?

While deficits in MSI have been observed in various domains in psychosis, it is still not sufficiently clear how widespread these deficits might be. Dysfunctional MSI has been observed in patients with the diagnosis of a psychotic disorder especially in linguistic (Tseng et al., 2015), temporal (Zhou et al., 2018), sensorimotor (Noel, Cascio, et al., 2017) processes. Further, it has been reported that emotional MSI might also be impaired, although previous evidence so far has been inconsistent (Lin et al., 2020; Tseng et al., 2015). In contrast, a handful previous studies suggested intact spatial MSI (Noel et al., 2020; Stone et al., 2014; Tseng et al., 2015) and findings of basic multisensory processes such as crossmodal target detection were inconsistent (Moran et al., 2021; Williams et al., 2010; Wynn et al., 2014).

Overall, the results of our studies do neither offer support for the assumption that MSI in psychosis might be globally impaired nor that it might be selectively impaired in specific domains. In all three of our studies, measures of MSI were comparable between subjects with and without psychosis proneness or the diagnosis of a psychotic disorder, respectively. Our results thus indicate typical emotional MSI in psychosis proneness and intact temporal and spatial MSI in patients with the diagnosis of a psychotic disorder. Crucially, while our findings on emotional and temporal MSI are in contrast to previous evidence in patient samples and might be due to methodological differences, the findings of our third study corroborate prior results suggesting intact spatial MSI in psychosis (Noel et al., 2020; Stone et al., 2014; Tseng et al., 2015).

Thus, embedding our findings into the literature on MSI in psychosis, the existing evidence overall rather tends to suggest that MSI might be selectively impaired in psychosis

and does not reflect a global dysfunction. This however raises the question why some domains of MSI are impaired, why others might remain largely intact, and which factors might influence this selectivity of impairments.

In their review, Tseng et al. (2015) discussed that previous evidence might indicate that MSI is especially impaired in speech and emotional stimuli, suggesting pronounced dysfunctions in processing socially relevant crossmodal information. Crucially, social cognition in terms of unisensory processing difficulties have been reported in psychosis (Healey et al., 2016; Savla et al., 2013) and it has been discussed that these difficulties might play a role in the pathogenesis of psychosis (e.g. D. Martin et al., 2020; Seo et al., 2020; Tripoli et al., 2022). Interestingly, impairments in *both* unisensory and multisensory processing was also observed in temporal and proprioceptive/sensorimotor processes, i.e. in the other domains in which deficient MSI has repeatedly been reported (Noel, Cascio, et al., 2017; Zhou et al., 2018).

These findings can be reconciled with the notion that MSI deficits might primarily manifest in domains, which already show marked impairments beyond crossmodal conditions: in unisensory processes such as facial and vocal emotion processing (Barkl et al., 2014; Gong et al., 2021; Kohler et al., 2010) or unisensory temporal processing (Amadeo et al., 2022) and in proprioceptive processes (Borda & Sass, 2015). Inversely speaking, this could indicate that domains, in which unisensory processes are not or only little impaired, might still largely benefit from the advantages of MSI, in the sense of intact crossmodal gain ameliorating potential unisensory performance deficits.

Further, it is also possible that crossmodal gain remains largely intact in psychosis in less complex tasks, such as spatial processing or target detection with simple stimuli. Initial evidence for this might be found in studies on AV far-near judgements, which observed impaired unisensory but similar multisensory performance in patients compared to controls (Stephen et al., 2013; Stone et al., 2011, 2014). This is further supported by a recent visuo-tactile target detection study by Moran et al. (2021) in patients with the diagnosis of a psychotic disorder. In this study, worse target detection performance was observed in unimodal conditions in comparison to healthy controls, but performance was similar in crossmodal trials. Interestingly, earlier ERPs over occipital and frontal regions, which were discussed to reflect neural substrates of MSI, did not differ between patients and controls. Moran et al. (2021) concluded that compensatory mechanisms of multisensory processes both on the behavioural and neural level, due to e.g. the higher attentional capture of crossmodal

stimuli (e.g. Talsma et al., 2009), resulted in comparable target detection performance in crossmodal trials in both groups (Moran et al., 2021).

In sum, it is possible that in domains with less gravely impaired unisensory performance or lower task complexity, such as spatial processing or target detection tasks with simple stimuli, benefits of multisensory processing are still largely preserved in psychosis. In contrast, MSI might no longer maintain performance benefits in psychosis in domains with marked unisensory processing deficits and higher task demands, possibly even potentiating sensory processing impairments. This encourages a more nuanced perspective of earlier assumptions that unisensory processing deficits in psychosis might result in pronounced multisensory deficits, based on the non-linear gain nature of MSI (Meredith & Stein, 1996). It might be possible that this assumption specifically refers to above mentioned domains, which show pronounced deficits in psychosis, i.e. linguistic, emotional (Healey et al., 2016; Savla et al., 2013), temporal (Amadeo et al., 2022) and proprioceptive information (Borda & Sass, 2015), but not as much in others such as spatial processing.

We conclude that more research is needed to sufficiently address the question of domain specific impairments of MSI in psychosis. Specifically, more studies with robust sample sizes might be needed in domains, which so far have received less scientific attention than others, such as spatial processing. Moreover, by comparing the level of impairments in both uni- and multisensory processing in psychosis per domain as well as across domains, future studies should investigate if MSI might show pronounced deficits in overall markedly impaired domains, potentially even aggravating unisensory processing dysfunctions.

### 2 Is Crossmodal Learning Deficient in Psychosis?

While considerable work has been done on MSI in psychosis, no published study investigated crossmodal, i.e. sensory-sensory recalibration in psychosis. It is thus unclear if psychosis is associated with dysfunctional adjustment to changes in crossmodal features in addition to aberrant integration of crossmodal signals. A failure to recalibrate to changes in crossmodal characteristics might drastically impede the creation and maintenance of holistic and stable percepts and could thereby contribute to erroneous and fragmented perception, which has repeatedly been reported in patients with the diagnosis of a psychotic disorder (Javitt & Freedman, 2015; Uhlhaas & Mishara, 2006).

The results of our two studies on temporal and spatial recalibration, however, do not offer support for this hypothesis. In our study on crossmodal temporal recalibration, patients

with the diagnosis of a psychotic disorder and healthy controls adjusted their SJ responses in a similar degree both to prolonged exposure to a fixed AV asynchrony and to the modality order on the previous AV trial. In our crossmodal spatial recalibration study, patients and controls showed a similar crossmodal capture effect of sound localization both by prolonged exposure to a consistent AV disparity and by the direction of crossmodal disparity on the preceding AV trial. Thus, our results indicate intact cumulative and immediate recalibration both in the temporal and spatial domain in psychosis. In sum, the findings of our studies on crossmodal recalibration provide first evidence that sensory-sensory recalibration might be unaffected in psychosis – in contrast to MSI.

This contrasts previous findings on crossmodal recalibration in the sensorimotor domain, where it has been observed that patients with the diagnosis of a psychotic disorder showed impairments in e.g. visuomotor recalibration in prism adaptation and saccadic adaptation paradigms (Bartolomeo et al., 2020; Lencer et al., 2017; Rösler et al., 2015). Impairments in sensorimotor recalibration in psychosis have been linked to deficits in corollary discharge, i.e. a failure to predict the sensory consequences of one's own behaviour (Bansal et al., 2018; Thakkar & Rolfs, 2019). It has been suggested that this might constitute a key factor for the pathogenesis psychosis, which might explain central symptoms of psychosis such as hallucinations or delusions of control (Moberget & Ivry, 2019; Poletti et al., 2019; Thakkar & Rolfs, 2019). Thus, dysfunctional corollary discharge mechanisms might offer a potential model for linking impaired sensorimotor processing and the development of psychotic symptoms, an issue which has not yet been sufficiently addressed in other domains of multisensory processing in psychosis.

Crucially, proprioceptive processes have repeatedly been observed to be markedly impaired in psychosis (Borda & Sass, 2015). This encourages the speculation that crossmodal recalibration in psychosis might express in a similar pattern of domain specific impairments as could be assumed in MSI (see **1 Are Multisensory Integration Deficits in Psychosis Global or Domain-Specific?**). Preliminary support for this speculation could be provided by our third study. It is possible that we did not observe impairments in sensory-sensory spatial recalibration, because this study focused on an intact domain, namely multisensory spatial processing. In our second study, however, we did not observe impaired crossmodal temporal recalibration in psychosis, i.e. in a sensory processing domain in which impairments have repeatedly been observed in psychosis (Amadeo et al., 2022; Zhou et al., 2018). Thus, the findings of our second study do not offer support for the speculation of domain-specific

recalibration deficits in psychosis. However, our patient sample in this study showed unimpaired temporal MSI, contrasting previous evidence (Zhou et al., 2018). This could indicate that our patient sample had overall rather spared temporal processing capabilities. Thus, it is possible that our findings on crossmodal temporal recalibration in psychosis might not be generalizable to the population of patients with the diagnosis of a psychotic disorder.

In sum, our studies provided first results on sensory-sensory recalibration in psychosis, suggesting intact spatial and temporal recalibration in patients with the diagnosis of a psychotic disorder. The results of our study on crossmodal spatial recalibration mirror both our findings and previous evidence on typical spatial MSI in psychosis (Noel et al., 2020; Stone et al., 2014; Tseng et al., 2015), suggesting that spatial multisensory processing might overall be unaffected in psychosis. In contrast, the picture is less clear in the temporal domain. We observed both unaffected crossmodal temporal recalibration and temporal MSI in our patient sample. The latter, however, contradicts previous evidence on impaired temporal MSI in psychosis (Zhou et al., 2018). Thus, future studies aiming to replicate our findings are needed to investigate if the temporal domain might be impaired in psychosis in the integration but not in recalibration of crossmodal stimuli. Alternatively, a coupling of impaired temporal MSI and impaired crossmodal temporal recalibration would suggest that temporal multisensory processing might overall be impaired in psychosis. Given that temporal processing might be distinctly dysfunctional in psychosis (Amadeo et al., 2022; Ueda et al., 2018), it could be argued that the latter seems more plausible.

It has been discussed that crossmodal recalibration processes reflect perceptual adjustment processes in response to changes in crossmodal features such as changed spatial and temporal relationships. Thus, crossmodal recalibration might play a vital role in the maintenance of stable and holistic representations of the world (Chen & Vroomen, 2013). This encourages the question if potential impairments in these processes might be associated with altered and fragmented perception as it has been previously observed in psychosis (Javitt & Freedman, 2015; Uhlhaas & Mishara, 2006). To our knowledge, we are the first to provide experimental evidence on intact sensory-sensory recalibration in psychosis, adding to the findings of a recent study in rare EOS (Zhou, Cui, et al., 2022). Thus, we encourage further studies with larger sample sizes and experimental designs covering the whole range of (multi)sensory processing domains, thus aiming to replicate our initial findings on crossmodal recalibration in psychosis and to extend them to other domains. Such approaches will contribute to answering the questions a) if psychosis is associated with intact or impaired

crossmodal recalibration, b) if crossmodal recalibration might be impaired in some domains but intact in others, c) if crossmodal recalibration in psychosis might show a similar pattern of domain specific impairments as could be assumed in MSI and d) if potential deficits in crossmodal recalibration might be associated with the development of psychotic symptoms.

## 3 Are Deficits in Multisensory Processing in Psychosis a Precedent or a Consequence of Disorder Development?

In several published studies, it has been discussed that deficits in multisensory processes might play a role in the *development* of psychotic symptoms (e.g. Amadeo et al., 2022; Postmes et al., 2014; Tseng et al., 2015; Wallace & Stevenson, 2014). Impaired MSI might facilitate the learning of erroneous associations or hinder the appropriate binding of information (Odegaard & Shams, 2017), which in turn might impede the construction and maintenance of holistic representations of the environment (de Jong et al., 2009; Postmes et al., 2014; Uhlhaas & Mishara, 2006). However, so far the majority of studies on MSI in psychosis have been conducted in patients with the diagnosis of a psychotic disorder using cross-sectional designs. In order to evaluate if impaired multisensory processes precede the manifestation of psychosis, studies in samples prior to disorder onset such as healthy subjects with high psychosis proneness are needed.

The literature so far fails to answer this question due to a relative scarcity of published studies in psychosis proneness, which focused only on a few domains of multisensory processing. Thus, previous studies in psychosis proneness did not sufficiently cover all domains of MSI, preventing generalized conclusions on deficient MSI as a potential key factor in the development of psychotic symptoms.

Most studies on MSI in psychosis proneness have been conducted in the temporal and in the sensorimotor domain. In several studies, a positive relationship between the size of the crossmodal TBW and higher levels of psychosis proneness could be observed (Dalal et al., 2021; Di Cosmo et al., 2021; Ferri et al., 2017, 2018; Marsicano et al., 2022). These findings are consistent with previous reports of an enlarged TBW in patients with the diagnosis of a psychotic disorder (Zhou et al., 2018). In the sensorimotor domain, previous evidence suggested a weakened sense of ownership over the own body and a higher occurrence to body representation distortions. This has been indicated by a positive relationship between a heightened susceptibility to the rubber hand illusion and increased psychosis proneness (Germine et al., 2013; Kállai et al., 2015; Torregrossa & Park, 2022). Taken together, both in the temporal and sensorimotor domain, findings in psychosis proneness indicated that MSI

might already be impaired before full-blown psychosis manifests, suggesting that dysfunctional multisensory processes might play a role in disorder development.

Only few studies investigated linguistic MSI in psychosis proneness. On the one hand, Muller et al. (2021) observed a higher occurrence of McGurk illusions in subjects with higher psychosis proneness, suggesting higher susceptibility to the illusion and thus enhanced linguistic MSI in their psychosis proneness sample. This finding contrasted the authors' hypothesis and was inconsistent with previous findings in patient samples (Tseng et al., 2015). On the other hand, Muller et al. (2020) did not observe any association between the McGurk effect and higher levels of psychosis proneness in healthy subjects, suggesting typical linguistic MSI in psychosis proneness. Thus, initial findings on linguistic in psychosis proneness so far are inconclusive, hampering inferences if linguistic MSI might already be dysfunctional prior to psychosis onset.

Apart from MSI studies in the temporal, sensorimotor, and linguistic domain in psychosis proneness, to our knowledge no study investigated spatial or emotional multisensory processing. Particularly, addressing this lack of studies on emotional MSI might be crucial for understanding how altered processing of social information might contribute to psychosis. Previous studies discussed a potential role of deficient emotion processing for psychosis development (D. Martin et al., 2020; Seo et al., 2020; Tripoli et al., 2022; Tseng et al., 2016) and emotional MSI has been observed to be altered in patients with the diagnosis of a psychotic disorder (Lin et al., 2020; Tseng et al., 2015). Thus, emotional MSI was investigated in psychosis proneness in the current dissertation project to address the question if dysfunctions in emotional MSI might precede the manifestation of psychosis.

Contrary to our hypothesis, we observed comparable emotional MSI between subjects with high and low psychosis proneness in our study on AV emotion processing. This contrasts the majority of previous findings in patient samples (Lin et al., 2020; Tseng et al., 2015) and previous findings on impaired unisensory emotion processing in psychosis proneness (D. Martin et al., 2020; Seo et al., 2020, S. 202; Tripoli et al., 2022; van Donkersgoed et al., 2015). While our results do not offer support for the assumption of a role of impaired emotional MSI for psychosis development, the question thus remains if our findings might be replicable or rather reflect a finding rather specific for our potentially high-functioning sample.

In sum, MSI has been rather selectively investigated in psychosis proneness. While dysfunctions in MSI in psychosis proneness have been observed in temporal (Dalal et al., 2021; Di Cosmo et al., 2021; Ferri et al., 2017, 2018; Marsicano et al., 2022) as well as sensorimotor processing (Germine et al., 2013; Kállai et al., 2015; Torregrossa & Park, 2022), findings in linguistic MSI are inconclusive (Muller et al., 2020, 2021) and no published study investigated emotional or spatial MSI in psychosis proneness. In our study, we observed typical emotional MSI in psychosis proneness, which however contrasts findings on unisensory emotion processing (D. Martin et al., 2020; Seo et al., 2020, S. 202; Tripoli et al., 2022; van Donkersgoed et al., 2015). Thus, so far the existing evidence does not support clear conclusions if deficits in MSI generally precede disorder manifestation. However, given patterns of impaired MSI in the temporal and sensorimotor domain observed both in psychosis proneness (Dalal et al., 2021; Di Cosmo et al., 2021; Ferri et al., 2017, 2018; Germine et al., 2013; Kállai et al., 2015; Marsicano et al., 2022; Torregrossa & Park, 2022) and patient samples (Noel, Cascio, et al., 2017; Zhou et al., 2018), it is likely that analogue patterns can be found in other domains of multisensory processing. It could be argued that this might particularly be the case in domains, which have previously been observed to be dysfunctional in psychosis in uni- and multisensory processes, such as those involving linguistic and emotional stimuli (Healey et al., 2016; Lin et al., 2020; Savla et al., 2013; Tseng et al., 2015).

Future studies should look into both linguistic and emotional MSI in psychosis proneness to evaluate if multisensory processes in these domains might be already dysfunctional prior to psychosis onset. In light of evidence suggesting unimpaired spatial MSI in psychosis (Tseng et al., 2015; see also **1 Are Multisensory Integration Deficits in Psychosis Global or Domain-Specific?**), future studies should also investigate if spatial multisensory processing is intact in psychosis proneness and remains stable over course of disorder development.

### 4 Can Anomalous Perceptual Experiences Conceptually Link Multisensory Processing and Phenomenology of Psychosis?

In this dissertation project, we aimed to explore a potential mechanism linking experimental findings on multisensory processes and phenomenology of psychosis. This is important since no convincing link so far has been established, despite published assumptions on an association between deficient MSI and the development psychotic symptoms (e.g. Amadeo et al., 2022; Tseng et al., 2015; Wallace & Stevenson, 2014). To our knowledge,

only one tangible model attempt has previously been published. In this account, Postmes et al. (2014) described how psychotic symptoms might develop as a means to explain an estranged experience of the self in the surrounding world due to deficient integration of proprioceptive perceptual processes and internal representations into a holistic percept. However, it is likely that the mechanism described by Postmes et al. (2014) might rather be specific to sensorimotor processes and their potential link to delusions of control and passivity experiences. Thus, it could be argued that the account of Postmes et al. (2014) might not be feasible to cover both a) the whole range of previous findings on impaired MSI in psychosis and b) psychosis as a whole syndrome.

As outlined earlier (see ***4.2 Anomalous Perceptual Experiences as a Potential Link between Multisensory Processing and Psychosis*** in the **General Introduction**), we argued that the concept of anomalous perceptual experiences (Bell et al., 2006; Wright et al., 2018) might be a promising mediator, which could link multisensory processing and psychosis via the mechanism proposed in the disorder model by Garety et al. (2001). Therefore, associations between our findings on MSI in subjects with psychotic symptoms, subjects' responses in the CAPS (Bell et al., 2006), and psychotic symptoms were investigated in the current project.

In all three of our studies on multisensory processes in subjects with psychotic symptoms, no association could be observed between experimental measures of MSI and crossmodal recalibration to anomalous perceptual experiences. Specifically, no significant correlation was found between the CAPS and 1) emotional MSI in subjects with high vs. low psychosis proneness, 2) temporal MSI and crossmodal temporal recalibration, as well as 3) spatial MSI and crossmodal spatial recalibration in patients with the diagnosis of a psychotic disorder and healthy controls. In contrast, in all three studies, significant strong correlations could be observed between the CAPS and questionnaires covering psychotic symptoms.

Thus, while our results corroborate a link between anomalous perceptual experiences and psychotic symptoms, no such association was found to multisensory processes. This raises several questions regarding potential reasons for this finding beyond those already discussed earlier in the respective study, i.e. potential power issues due to sample sizes and reduced variance in multisensory processing between groups due to similar multisensory performance in patients and healthy controls.

It is possible that the CAPS is conceptually closer to phenomenology of psychosis, i.e. closer to the psychological level of observation of psychotic symptoms, than it is to experimental findings of MSI. As such it might not be feasible to offer a conceptual link in-between both levels of observation. This is supported by the strong correlations between the CAPS and measures for psychotic symptoms in our data, indicating a conceptual proximity between both.

Moreover, the item composition of the CAPS might be a possible reason for the lack of correlations. Although the CAPS inquires perceptual anomalies across all senses, i.e. vision, audition, touch, smell, taste and proprioceptive/vestibular experiences (Bell et al., 2006), the questions are almost exclusively formulated targeting a single sensory modality each. Future approaches might benefit from applying an instrument, which targets *crossmodal* rather than unimodal perceptual anomalies, in order to reveal associations to experimental findings of MSI.

To conclude, we did not observe associations between our experimental measures of MSI and crossmodal recalibration, respectively, and self-reports on anomalous perceptual experiences as measured by the CAPS (Bell et al., 2006). Thus, our findings do not offer support for our hypothesis that dysfunctional multisensory processing might facilitate the experience of conscious perceptual anomalies. This suggests that our proposition on a potential link between MSI and the development of psychotic symptoms based on the mechanism described by Garety et al. (2001) might not be sufficiently suitable. Future studies should explore alternative mechanism, via which deficits in multisensory processes might contribute to the development of psychotic symptoms, as it was previously assumed (Amadeo et al., 2022; Tseng et al., 2015; Wallace & Stevenson, 2014). It might be beneficial to examine mechanisms underlying altered multisensory processing in psychosis and relate them to components addressing altered perception in disorder models of psychosis. A possible target for future research to understand the underlying mechanisms of multisensory processing in psychosis and relate them to etiological processes might be found in altered perceptual inference mechanisms (see **5.1 Causal Inference**), as previously observed in psychosis (Sterzer et al., 2018).

## 5 Mechanisms Underlying Multisensory Processes in Psychosis

While a considerable number of published studies investigated multisensory processes in psychosis, the mechanisms underlying dysfunctional MSI in psychosis are still not sufficiently understood. As proposed above (see **4 Can Anomalous Perceptual Experiences Conceptually Link Multisensory Processing and Phenomenology of Psychosis?**), examining the mechanisms underlying multisensory processes in psychosis might be beneficial for understanding how altered multisensory processing might contribute to psychosis. In the following, I will briefly discuss two lines of research on potential mechanisms of altered multisensory processing in psychosis: inference mechanisms governing multisensory processing and neural substrates of MSI in psychosis.

### 5.1 Causal Inference

As briefly introduced earlier (see **2.1 Multisensory Integration** in the **General Introduction**), a central problem for the brain during MSI lies within the decision, if crossmodal stimuli belong to the same event and therefore should be integrated or not (Shams & Beierholm, 2010). It has been discussed that this problem is solved by the brain by means of Bayesian inferencing, a principle called causal inference (Körding et al., 2007; Sato et al., 2007).

Within this inference framework, the likelihood of crossmodal stimuli being integrated is influenced by two types of variables. For one, the incoming sensory signals derived from the each of the crossmodal cues are formalized as their respective sensory likelihood. Both sensory likelihoods describe estimates of the underlying veridical stimuli, weighed by their reliability. Second, a binding prior, called prior of common cause or p' common, describes the likelihood of each crossmodal combination in the absence of sensory input. This prior entails the a priori estimate for crossmodal stimuli originating from the same cause as opposed to separate causes. The brain infers if it is more likely that crossmodal stimuli originated from the same or separate sources based on both types of variables, i.e. their sensory likelihood and the prior of common cause. More specifically, the causal inference model computes the probability of causal structure for the origin of crossmodal cues separately for the assumption of a common and separate causes. The more probable causal structure dominates multisensory perception, influencing the degree of crossmodal influence and MSI (Körding et al., 2007; Sato et al., 2007; Shams & Beierholm, 2010).

Alterations in Bayesian inference processes have repeatedly observed in psychosis (Heinz et al., 2019; Humpston et al., 2019; Sterzer et al., 2018). It has been discussed that

impairments in weighing prior knowledge and sensory input might underlie altered perception and cognition in psychosis, potentially forming the basis for the development and maintenance of psychotic symptoms (for a review, see Sterzer et al., 2018). This predictive coding account of psychosis (Sterzer et al., 2018), is based on earlier work on neural information processing, which assumed that the brain processes information based on hierarchical Bayesian inference. Within this framework, the brain weighs incoming sensory input and prior knowledge of the world by their respective precision. The resulting prediction error, i.e. the mismatch between input and prior, is then propagated upwards the hierarchy, used to update prior knowledge and thus to continuously learn and adjust the inner representations of the self and the environment (Friston, 2005; Mathys, 2011).

Psychosis has been observed to be associated with marked alterations within this predictive coding framework, i.e. aberrant precision weighing of both sensory input and prior knowledge and thus maladaptive signaling of prediction errors (Heinz et al., 2019; Humpston et al., 2019; Sterzer et al., 2018). Thus, alterations in predictive processing might be able to describe how impaired sensory processing and altered perception contribute to the development of psychotic symptoms via maladaptive perceptual learning mechanisms. To date, it is however not sufficiently clear how the alterations in inference mechanisms manifest in psychosis. Previous studies observed both deficient or excessive weighing of prior knowledge, sensory input or prediction errors (e.g. Davies et al., 2018; Limongi et al., 2018; Powers et al., 2017; Schmack et al., 2013, 2015, 2017; Teufel et al., 2015; Weilnhammer et al., 2020). It has been discussed that these inconsistencies might be explicable by the hierarchical structure of the predictive coding framework, describing deficient weighing of prior knowledge in lower levels but excessive weighing in higher levels of the hierarchy (Sterzer et al., 2018). While it is conclusively understood how alterations in predictive information processing in psychosis might shape altered perception, models of altered predictive processing might be nonetheless helpful to understand associations between altered perception and psychosis. Moreover, investigating aberrant inference mechanisms in multisensory perception such as causal inference might benefit the understanding of the potential role of impaired multisensory processing for psychosis.

While alterations in predictive coding have been repeatedly overserved in psychosis (Sterzer et al., 2018), only very few studies investigated causal inference in psychosis in order to elucidate dysfunctional mechanisms underlying multisensory processes in psychosis. To

our knowledge only two published studies investigated if psychosis might be associated to altered causal inference.

Noel et al. (2018) investigated causal inference during temporal MSI in patients with the diagnosis of a psychotic disorder, patients with the diagnosis of an ASD, and two respective healthy control groups. In a SJ paradigm with AV speech stimuli, the authors observed impaired temporal MSI, i.e. an enlarged TBW, in both patient groups. Further, both groups showed altered causal inference parameters compared to controls, with both patient groups, however, showing different alterations. While patients with the diagnosis of an ASD mainly differed in their prior of ascribing a common cause, i.e. p' common, patients with the diagnosis of a psychotic disorder showed (trends for) alterations in both p' common and sensory likelihood. This suggested that although ASD and psychosis are associated with an enlarged TBW, the underlying mechanisms might be different. Specifically, deficient temporal multisensory processes in psychosis might most likely be rooted in a combination of altered causal priors and unreliable sensory representations (Noel et al., 2018).

Odegaard & Shams (2017) investigated multisensory spatial and temporal processing in healthy subjects with varying levels of psychosis proneness. Basing their expectations on previous work on altered weighing of prior knowledge in psychosis (for a review, see Sterzer et al., 2018), Odegaard & Shams (2017) investigated if higher levels of psychosis proneness are associated with a lower binding tendency, i.e. prior of common cause, during spatial and temporal AV integration. Results indicated that individuals with high psychosis proneness showed a decreased binding tendency in the spatial task, but not in the temporal task. This suggested that psychosis proneness might be associated with a reduced ascription of a common cause across spatial dimensions. In contrast, the null findings in the temporal task might have reflected an averaging issue, with some of the subjects with the highest proneness ratings showing a reduced, others an increased binding tendency (Odegaard & Shams, 2017).

In sum, previous work on causal inference in psychosis and psychosis proneness indicated that alterations in the parameters governing causal inference might underlie deficient multisensory processing in psychosis (Noel et al., 2018; Odegaard & Shams, 2017). These alterations in causal inference, reflected by altered priors and sensory representations, are consistent with dysfunctional inference mechanisms in psychosis described by the predictive coding account (Sterzer et al., 2018). This suggests a global, underlying dysfunction in Bayesian inferencing in psychosis at the root of both uni- and multisensory information processing and higher cognition.

However, the findings of Noel et al. (2018) and Odegaard & Shams (2017) should be interpreted with caution. In both studies, several of the reported results reflected non-significant trends or did not survive multiple comparison correction. Further, findings are not entirely consistent. Noel et al. (2018) reporting a tendency for altered priors during temporal MSI in patients, while Odegaard & Shams (2017) did not observe an association between an altered binding tendency and psychosis proneness in a multisensory temporal task. Thus, the literature so far only insufficiently addressed causal inference mechanism during multisensory processing in psychosis. Several open questions remain to be answered:

If altered causal inference processes might reflect a global underlying impairment in psychosis, how could this explain the heterogeneous findings on MSI in psychosis? How can a global underlying deficit in p' common and sensory likelihood explain reduced integration in e.g. linguistic tasks, but excessive integration in temporal tasks? Does the processing of complex compared to simple crossmodal information in psychosis interact with a global underlying dysfunction in causal inference, leading to pronounced dysfunctions in multisensory processing of speech or emotional cues? How do the findings of Odegaard & Shams (2017) of a reduced binding tendency in psychosis proneness during a spatial task fit both previous evidence (Noel et al., 2020; Stone et al., 2014; Tseng et al., 2015) and our findings suggesting intact spatial MSI in psychosis?

In sum, casual inference in psychosis is still poorly understood. Future studies should address open questions to contribute to a better understanding how causal inference dysfunctions might underlie the heterogeneity of multisensory processes in psychosis. For example, future studies should compare parameters of causal inference in psychosis in a variety of crossmodal tasks, such as VE paradigms, SJ judgement tasks, and tasks with speech and emotional stimuli. Alterations of p' common and sensory likelihood should be compared across these processing domains and between patients, subjects with high psychosis proneness, and healthy controls. This approach will aid in answering if alterations in causal inference might reflect a global deficit in psychosis, or if they might manifest differently depending on task complexity and domain. Further, embedding evidence on causal inference in psychosis in the broader framework of the predictive coding account of psychosis (Sterzer et al., 2018) might build the base for understanding if altered multisensory processes might contribute to the development of psychotic symptoms via maladaptive perceptual inference processes.

## 5.2 Neural Substrates of Multisensory Processing in Psychosis

In their paper, Odegaard & Shams (2017) highlighted the importance of neurophysiological and neuroimaging investigations to further understand the mechanism underlying altered multisensory processes in psychosis. Odegaard & Shams (2017) briefly discussed a potential reason for differences in causal inference during MSI in psychosis based on the target domain. The authors argued that white-matter disconnectivity reported in psychosis (Crossley et al., 2016; Friston et al., 2016) might underlie an altered prior of common cause. It could be argued that differences in recruitment of cortical areas during multisensory processing depending on target domain might partially explain why the prior of common cause potentially varies across tasks. Odegaard & Shams (2017) proposed that the respective regions underpinning various target domains in multisensory tasks, such as the superior temporal sulcus in speech (e.g. Nath & Beauchamp, 2012) or the intraparietal sulcus in a spatial task (e.g. Rohe & Noppeney, 2015), are affected to a different degree by structural white-matter disconnectivity in psychosis. As such, a deficit in the binding tendency, i.e. prior of common cause, might manifest differently depending on the underlying neural substrates (Odegaard & Shams, 2017).

Earlier neurophysiological and neuroimaging studies have revealed overlaps in neural substrates of multisensory processes and symptoms of psychosis. For example, an earlier study by Kim et al. (2003) reported that activation in the superior temporal sulcus and the inferior frontal gyrus was involved in both temporal multisensory processes and the emergence of hallucinations, suggesting a neural link between MSI and psychosis.

A recent systematic review by Gröhn et al. (2022) aimed to investigate previous evidence on the neural underpinnings of MSI in patients with the diagnosis of a psychotic disorder. By reviewing EEG, MEG and fMRI studies during MSI in psychosis, Gröhn et al. (2022) were able to draw preliminary conclusions about aberrant neural processes at the root of multisensory processing in psychosis. In EEG and MEG studies, patients with the diagnosis of a psychotic disorder showed altered or reduced cortical responses to crossmodal stimuli or inverse neural response patterns when contrasting congruent and incongruent stimuli compared to healthy controls (e.g. Stekelenburg et al., 2013; Szycik et al., 2013). Reduced oscillatory activity across several bands was observed in psychosis during processing of crossmodal stimuli, suggesting crossmodal information processing deficits (Balz et al., 2016; Roa Romero, Keil, Balz, Gallinat, et al., 2016; Roa Romero, Keil, Balz, Niedeggen, et al., 2016). These impairments were observed predominantly over parietal,

occipital and fronto-temporal areas (e.g. Balz et al., 2016; Liu et al., 2016; Sanfratello et al., 2018) and aberrant neural activity was observed especially in regions covering the superior temporal sulcus, superior temporal gyrus and the inferior temporal gyrus (e.g. Stekelenburg et al., 2013; Straube et al., 2014; Szycik et al., 2013). However, conflicting findings were also observed, reporting no group differences in cortical responses to crossmodal stimuli or even increased neural signaling in patients compared to controls, which suggested intact or enhanced MSI (Müller et al., 2014; Stone et al., 2011). Further, fMRI studies revealed reduced functional connectivity predominantly between the superior temporal sulcus and frontal regions as well as between the inferior frontal gyrus and various cortical areas (Straube et al., 2014; Szycik et al., 2013).

Overall, the results of the review by Gröhn et al. (2022) suggested reduced or at least altered neural responses to crossmodal stimuli and functional disconnectivity between networks involved in processing of crossmodal information. This offers support for the assumption that altered neural mechanisms such as deficient neural signaling or impaired cortical connectivity might underlie dysfunctional MSI in psychosis (Gröhn et al., 2022; Odegaard & Shams, 2017).

However, it is important to note that only studies on AV information processing in psychosis could be included in the review, revealing a lack of evidence on neural substrates in other crossmodal stimulus combinations. Further, the result pattern was not entirely consistent, with some studies reporting no group differences or even enhanced neural multisensory facilitation effects in psychosis (Müller et al., 2014; Stone et al., 2011). Gröhn et al. (2022) argued that previous studies differed to a significant extent in task complexity and experimental design, which might be a possible reason for the observed heterogeneous findings. Crucially, the response pattern was judged as not sufficiently consistent to hypothesize about concrete influencing factors (Gröhn et al., 2022). Lastly, alterations in neural activity in psychosis were repeatedly not mirrored by behavioural multisensory impairments (Gröhn et al., 2022).

It is possible that altered neural activity during multisensory processing in psychosis might not necessarily translate into impaired behavioural performance. It would be interesting to investigate if altered neural responses in ERP, oscillatory activity or functional disconnectivity could have been observed in our psychosis samples, which did not show behavioural differences in emotional, temporal, and spatial multisensory processing compared to control subjects. It is possible that multisensory behavioural performance might be partially

able to compensate for altered neural substrates in psychosis, leading to typical behavioural MSI compared to healthy controls. Preliminary support for this hypothesis can be found in two earlier AV far-near judgment studies. Stephen et al. (2013) observed reduced parietal MEG responses to AV stimuli in patients with the diagnosis of a psychotic disorder. Stone et al. (2014) showed altered gamma-band oscillations frontal regions as well as in visual and auditory cortices in response to AV stimulation in psychosis. In both studies, neural findings were not associated with multisensory behavioural performance (Stephen et al., 2013; Stone et al., 2014).

It could be argued that the to date published evidence on neural alterations underpinning multisensory processes in psychosis partly mirrors the evidence on behavioural findings. In various studies, dysfunctions or at least alterations of multisensory processes could be observed, but results seemed to be inconsistent and sometimes contradictory (Gröhn et al., 2022; Tseng et al., 2015). Although Gröhn et al. (2022) observed no apparent difference in result patterns between social and non-social designs, it might be possible that alterations in neural substrates of MSI in psychosis might potentially be influenced by similar factors as behavioural studies such as symptom severity (Lin et al., 2020), stimulus complexity (Darke et al., 2021; Tseng et al., 2015) or the target domain of multisensory processing (e.g. Tseng et al., 2015; Zhou et al., 2018).

We conclude that future studies on multisensory processing in psychosis should include neurophysiological and neuroimaging techniques to elucidate the potential impact of these proposed factors. Studies in patients with the diagnosis of a psychotic disorder should combine behavioural measurements and neural markers of multisensory processing to investigate if behavioural performance might partially compensate for alterations in neural substrates. Further, studies should investigate neural substrates of MSI in psychosis proneness to evaluate if neural processes underpinning MSI might be already altered prior to disorder onset. This would aid in understanding the developmental trajectory of multisensory processes in psychosis and shed further light on the insufficiently answered question if aberrant MSI might play a role in the development of psychosis (see **3 Are Deficits in Multisensory Processing in Psychosis a Precedent or a Consequence of Disorder Development?**).

## 6 Limitations and Outlook

In this dissertation project, we recruited stable outpatients as patient samples for the studies on temporal and spatial multisensory processing. It is possible that our patient samples differed in in-/outpatient status compared to previous studies and as such might be dissimilar in impairments beyond those tested in our studies, such as functioning level. Further, it could be argued that the patient samples of previous studies might differ in hitherto unaddressed factors, which could have an impact on multisensory processing, such as motivation, concentration or neuro-cognitive capacities. These differences could be due to variations in in-/outpatient status in previous studies. As such, these factors might partially account for performance differences observed in previous studies.

Thus, we argue that it might be important to review, which factors might be partial confounds in explaining the difference in multisensory processes between patients and healthy controls. For example, inpatient status might be associated with a higher level of general impairment than outpatient status. By not accounting for such potential group differences, the specific role of psychosis on group differences in multisensory processing might be overestimated.

To compare the patient status in previous studies to our patient samples, we reviewed published studies on temporal and spatial multisensory processes in psychosis. This revealed that the in-/outpatient status considerably varied between studies, with either outpatient (de Boer-Schellekens et al., 2014; de Gelder et al., 2003; B. Martin et al., 2013; Wynn et al., 2014; Zopf et al., 2021), inpatient (Zhou, Lai, et al., 2022; Zvyagintsev et al., 2017) or mixed samples (Haß et al., 2017) reported. Further, several studies did not sufficiently specify if in- or outpatient samples were recruited (Foucher et al., 2007; Noel et al., 2018, 2020; Stephen et al., 2013; Stevenson, Park, et al., 2017; Stone et al., 2011, 2014; Vanes et al., 2016; Williams et al., 2010). Thus, it might be difficult to assess to what extent in-/outpatient status could have influenced group differences in multisensory processing in previous studies in comparison to our findings.

Next, we briefly reviewed reports on symptom severity in previous studies as a potential estimate, if the level of impairments might have differed between studies. Most of previous studies on temporal and spatial multisensory processes in psychosis reported scores for the PANSS positive and negative subscale, an instrument measuring positive and negative symptoms of psychotic disorders (Kay et al., 1987). Overall, previous studies reported mean scores of approximately 14-16 points per subscale (Noel et al., 2018, 2020; Tseng et al., 2015;

Zhou et al., 2018; Zhou, Lai, et al., 2022; Zopf et al., 2021), which might correspond to a status of no acute illness (Leucht et al., 2005). Since previous studies suggested that samples might not have differed greatly in impairment level, we think it is unlikely that variations in symptom severity or general impairment account for group effects on multisensory processing.

Crucially, our short review of the literature for differences in patient status and symptom level was done post-hoc. Future studies on multisensory processes in psychosis should recruit two samples of in- and outpatients, respectively, and compare them to healthy controls. These approaches should investigate concrete factors, which differ between the two patient groups, such as functioning level, symptom severity, concentration or motivation, and test if these factors partially account for dysfunctions of multisensory processing. This approach would be helpful in order to rule out confounding factors, avoid overestimating the specific impact of psychosis on multisensory processing, and contribute to a clearer understanding of the association between psychosis and multisensory processes.

It is important to note that we relied on AV processes as an operationalization of multisensory processes in psychosis. While this mirrors the majority of designs in previous studies on MSI in psychosis (Lin et al., 2020; Tseng et al., 2015; Zhou et al., 2018), this limits the generalizability of our findings to other crossmodal stimulus combinations. Some previous findings suggested that impairments in MSI might be affected by the targeted sensory modality, for example in the emotional domain (Lin et al., 2020; Tseng et al., 2015). However, studies beyond the AV stimulus combination in psychosis are rare. This highlights the importance for future studies to investigate other crossmodal stimulus combinations in order to obtain a comprehensive picture of multisensory processes in psychosis.

As outlined above, our findings only partially corroborate previous findings on multisensory processes in psychosis. While we were able to replicate findings on intact spatial MSI in psychosis, we did not observe expected impairments in emotional MSI in psychosis proneness and temporal MSI in patients with the diagnosis of a psychotic disorder. While we contributed first findings suggesting typical emotional MSI in psychosis proneness and intact crossmodal recalibration in the spatial and temporal domain in psychosis, more studies are needed in order to confirm these results. Overall, our findings and their discussion with regard to previous evidence suggest a complex pattern of multisensory processes in psychosis, which likely does not reflect a single, globally impaired process but rather an interplay of a

multitude of factors influencing if some multisensory processes might be negatively affected or relatively spared in the disorder.

In our findings, some hints for such possible factors were observable, which should be addressed in future research. As discussed throughout this dissertation, the most apparent factor might lie in the target domain of multisensory processes, meaning if spatial, temporal, emotional, linguistic or sensorimotor processes are involved. Previous evidence suggested impairments in psychosis especially in domains with more complex information, i.e. mainly linguistic, temporal, sensorimotor and (albeit inconsistently) emotional stimuli (Lin et al., 2020; Noel, Cascio, et al., 2017; Tseng et al., 2015; Zhou et al., 2018). However, the assumption of impaired emotional and temporal MSI was not supported by our findings, indicating that the notion of domain-specificity might not yet be comprehensively understood. As we discussed earlier (see **1 Are Multisensory Integration Deficits in Psychosis Global or Domain-Specific?**), it is possible that impairments in multisensory processes might particularly manifest in domains, which already show pronounced deficits in unisensory processing. Other potential factors determining how emotional multisensory processing impairments manifest might be found in the task-relevance of the attended modality (Thaler et al., 2013), stimulus complexity such as static or dynamic stimuli (Darke et al., 2021), emotion category or valence (e.g. Bonfils et al., 2020; Tripoli et al., 2022), symptom severity (Lin et al., 2020) or the respective association with positive vs. negative symptoms of psychotic disorders (e.g. Williams et al., 2010; Wynn et al., 2014). This suggests that a more fine-grained approach in investigating multisensory processes in psychosis might be needed to elucidate the potential impact and interaction of above proposed factors.

A crucial line of research for future study will be to elucidate exact mechanisms, by which altered multisensory processes might contribute to psychosis. In this dissertation project, we provided first evidence that anomalous perceptual experiences (Bell et al., 2006; Garety et al., 2001) likely is not the central mechanism, suggesting that other factors might be more relevant. This calls for more studies on the following three aspects. First, future studies should examine multisensory processes in psychosis proneness or first degree relatives – especially in underrepresented domains such as emotional, linguistic and spatial processing. This will aid in understanding the developmental trajectory of multisensory processes in psychosis and if impairments thereof might constitute a risk factor for the disorder. Second, investigating the underlying neural substrates (Gröhn et al., 2022) and causal inference processes (Noel et al., 2018; Odegaard & Shams, 2017) of multisensory processing in

psychosis will contribute to understand potential dysfunctions in neural and computational mechanisms at the root of multisensory processes in psychosis. Finally, by integrating evidence from different observational layers, i.e. behavioral, neural, and computational findings, future studies should aim to provide models describing the contribution of altered multisensory processes for psychosis. A potential, promising candidate framework can be found in computational approaches such as the predictive coding account of psychosis (Sterzer et al., 2018). However, more research especially on causal inference in psychosis is needed to understand how dysfunctions in multisensory processes in psychosis might be explainable in a predictive coding perspective.

## 7 Concluding Remarks

This dissertation project aimed to contribute experimental evidence on multisensory processes in psychosis. This research aim was built on the hypothesis that altered integration and recalibration of crossmodal information is associated with maladaptive learning and thus contributes to psychosis. The available literature suggested that deficient integration of crossmodal signals might facilitate fragmented perception (Postmes et al., 2014; Uhlhaas & Mishara, 2006) and lead to false inferences on the cause of the sensory information (Odegaard & Shams, 2017). Further, excessive integration might facilitate erroneous binding of information originating from separate sources and thus lead to confusing percepts and the maladaptive learning of perceptual associations (Amadeo et al., 2022; Odegaard & Shams, 2017; Zhou et al., 2018).

The findings of this dissertation project do not support the notion that deficits in multisensory processing in psychosis reflect a global dysfunction. Given the lack of group differences in all three of our experiments, our findings however also fail to offer support for the assumption of a domain-specificity of multisensory processing deficits. Our results might suggest that altered perception in psychosis is not generally driven by dysfunctions in multisensory processing. While the studies in this project did not confirm impairments in emotional MSI in psychosis proneness and deficient temporal MSI in patients with the diagnosis of a psychotic disorder, we observed typical spatial MSI in psychosis, replicating previous findings (de Gelder et al., 2003; Noel et al., 2020; Tseng et al., 2015). More studies are however needed to disentangle factors, which could explain why our subjects did not show dysfunctions in emotional and temporal MSI, as previously reported in psychosis (Lin et al., 2020; Tseng et al., 2015; Zhou et al., 2018).

Further, to our knowledge the studies in project are the first to provide evidence on crossmodal spatial and temporal recalibration in patients with the diagnosis of a psychotic disorder. Our results indicate that psychosis might be associated with intact cumulative and immediate recalibration processes in both domains. Thus, the findings of this project suggest that aberrant perception in psychosis might not be associated with dysfunctional perceptual adjustments to changes in crossmodal characteristics.

The findings of this project did not corroborate that deficits in emotional MSI precede psychosis onset. This in turn does not offer support for the notion that impaired MSI overall plays a role in disorder development. Future lines of research should target concrete mechanisms how altered multisensory processes might contribute to psychosis. In this project, a proposed mechanism describing that dysfunctional multisensory processes might facilitate anomalous perceptual experiences (Bell et al., 2006; Garety et al., 2001) was not supported by our findings.

A road for future research is to relate findings of multisensory processing in psychosis to neural substrates as well as computational mechanisms of perception to elaborate on mechanisms underlying multisensory perception in psychosis. Further, approaches to formulate models how multisensory processing might be linked to psychosis should incorporate relevant factors and empirically test how these add to the explanatory value of a disorder model. It is likely that an interplay of factors such as task and stimulus complexity, involved sensory modalities, alterations in neural substrates and dysfunctions in underlying computational processes governs how multisensory processing manifests in psychosis.

**References**

Addington, J., Case, N., Saleem, M. M., Auther, A. M., Cornblatt, B. A., & Cadenhead, K. S. (2014). Substance use in clinical high risk for psychosis: A review of the literature: Substance use in clinical high risk for psychosis. *Early Intervention in Psychiatry*, *8*(2), 104–112. https://doi.org/10.1111/eip.12100

Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current Biology*, *14*(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Amadeo, M. B., Esposito, D., Escelsior, A., Campus, C., Inuggi, A., Pereira Da Silva, B., Serafini, G., Amore, M., & Gori, M. (2022). Time in schizophrenia: A link between psychopathology, psychophysics and technology. *Translational Psychiatry*, *12*(1), 331. https://doi.org/10.1038/s41398-022-02101-x

American Psychiatric Association. (2022). *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5-TR). American Psychiatric Association Publishing. https://doi.org/10.1176/appi.books.9780890425787

Andreasen, N. C., Nopoulos, P., O'Leary, D. S., Miller, D. D., Wassink, T., & Flaum, M. (1999). Defining the phenotype of schizophrenia: Cognitive dysmetria and its neural mechanisms. *Biological Psychiatry*, *46*(7), 908–920. https://doi.org/10.1016/S0006-3223(99)00152-3

Balz, J., Roa Romero, Y., Keil, J., Krebber, M., Niedeggen, M., Gallinat, J., & Senkowski, D. (2016). Beta/gamma oscillations and event-related potentials indicate aberrant multisensory processing in schizophrenia. *Frontiers in Psychology*, *7*, 1896. https://doi.org/10.3389/fpsyg.2016.01896

Bansal, S., Ford, J. M., & Spering, M. (2018). The function and failure of sensory predictions. *Annals of the New York Academy of Sciences*, *1426*(1), 199–220. https://doi.org/10.1111/nyas.13686

Bansal, S., Murthy, K. G., Fitzgerald, J., Schwartz, B. L., & Joiner, W. M. (2019). Reduced transfer of visuomotor adaptation is associated with aberrant sense of agency in schizophrenia. *Neuroscience*, *413*, 108–122. https://doi.org/10.1016/j.neuroscience.2019.06.012

Barkl, S. J., Lah, S., Harris, A. W. F., & Williams, L. M. (2014). Facial emotion identification in early-onset and first-episode psychosis: A systematic review with meta-analysis. *Schizophrenia Research*, *159*(1), 62–69. https://doi.org/10.1016/j.schres.2014.07.049

Bartolomeo, L. A., Shin, Y.-W., Block, H. J., Bolbecker, A. R., Breier, A. F., O'Donnell, B., & Hetrick, W. P. (2020). Prism adaptation deficits in schizophrenia. *Schizophrenia Bulletin*, *46*(5), 1202–1209. https://doi.org/10.1093/schbul/sbaa019

Beesdo-Baum, K., Zaudig, M., & Wittchen, H.-U. (2019). *Strukturiertes Klinisches Interview für DSM-5-Störungen – Klinische Version* (1. Aufl.). Hogrefe.

Bell, V., Halligan, P. W., & Ellis, H. D. (2006). The Cardiff Anomalous Perceptions Scale (CAPS): A new validated measure of anomalous perceptual experience. *Schizophrenia Bulletin*, *32*(2), 366–377. https://doi.org/10.1093/schbul/sbj014

Bob, P., Pec, O., Mishara, A. L., Touskova, T., & Lysaker, P. H. (2016). Conscious brain, metacognition and schizophrenia. *International Journal of Psychophysiology*, *105*, 1–8. https://doi.org/10.1016/j.ijpsycho.2016.05.003

Bob, P., Vlcek, P., & Raboch, J. (2014). Sensory disturbances, inhibitory deficits, and the P50 wave in schizophrenia. *Neuropsychiatric Disease and Treatment*, *10*, 1309. https://doi.org/10.2147/NDT.S64219

Bonfils, K. A., Haas, G. L., & Salyers, M. P. (2020). Emotion-specific performance across empathy tasks in schizophrenia: Influence of metacognitive capacity. *Schizophrenia Research: Cognition*, *19*, 100139. https://doi.org/10.1016/j.scog.2019.100139

Bonfils, K. A., Ventura, J., Subotnik, K. L., & Nuechterlein, K. H. (2019). Affective prosody and facial emotion recognition in first-episode schizophrenia: Associations with functioning & symptoms. *Schizophrenia Research: Cognition*, *18*, 100153. https://doi.org/10.1016/j.scog.2019.100153

Borda, J. P., & Sass, L. A. (2015). Phenomenology and neurobiology of self disorder in schizophrenia: Primary factors. *Schizophrenia Research*, *169*(1–3), 464–473. https://doi.org/10.1016/j.schres.2015.09.024

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature*, *391*, 756–756. https://doi.org/10.1038/35784

Brittain, P., ffytche, D. H., McKendrick, A., & Surguladze, S. (2010). Visual processing, social cognition and functional outcome in schizophrenia. *Psychiatry Research*, *178*(2), 270–275. https://doi.org/10.1016/j.psychres.2009.09.013

Bruns, P. (2019). The ventriloquist illusion as a tool to study multisensory processing: An update. *Frontiers in Integrative Neuroscience*, *13*, 51. https://doi.org/10.3389/fnint.2019.00051

Bruns, P., Li, L., Guerreiro, M. J. S., Shareef, I., Rajendran, S. S., Pitchaimuthu, K., Kekunnaya, R., & Röder, B. (2022). Audiovisual spatial recalibration but not

integration is shaped by early sensory experience. *IScience*, *25*(6), 104439. https://doi.org/10.1016/j.isci.2022.104439

Bruns, P., Liebnau, R., & Röder, B. (2011). Cross-modal training induces changes in spatial representations early in the auditory processing pathway. *Psychological Science*, *22*(9), 1120–1126. https://doi.org/10.1177/0956797611416254

Bruns, P., & Röder, B. (2010). Tactile capture of auditory localization is modulated by hand posture. *Experimental Psychology*, *57*(4), 267–274. https://doi.org/10.1027/1618-3169/a000032

Bruns, P., & Röder, B. (2015). Sensory recalibration integrates information from the immediate and the cumulative past. *Scientific Reports*, *5*(1), 12739. https://doi.org/10.1038/srep12739

Bruns, P., Spence, C., & Röder, B. (2011). Tactile recalibration of auditory spatial representations. *Experimental Brain Research*, *209*(3), 333–344. https://doi.org/10.1007/s00221-011-2543-0

Caclin, A., Soto-Faraco, S., Kingstone, A., & Spence, C. (2002). Tactile "capture" of audition. *Perception & Psychophysics*, *64*(4), 616–630. https://doi.org/10.3758/BF03194730

Capa, R. L., Duval, C. Z., Blaison, D., & Giersch, A. (2014). Patients with schizophrenia selectively impaired in temporal order judgments. *Schizophrenia Research*, *156*(1), 51–55. https://doi.org/10.1016/j.schres.2014.04.001

Carroll, L. S., & Owen, M. J. (2009). Genetic overlap between autism, schizophrenia and bipolar disorder. *Genome Medicine*, *1*(10), 102. https://doi.org/10.1186/gm102

Carter, O., Bennett, D., Nash, T., Arnold, S., Brown, L., Cai, R. Y., Allan, Z., Dluzniak, A., McAnally, K., Burr, D., & Sundram, S. (2017). Sensory integration deficits support a dimensional view of psychosis and are not limited to schizophrenia. *Translational Psychiatry*, *7*(5), e1118–e1118. https://doi.org/10.1038/tp.2017.69

Castagna, F., Montemagni, C., Maria Milani, A., Rocca, G., Rocca, P., Casacchia, M., & Bogetto, F. (2013). Prosody recognition and audiovisual emotion matching in schizophrenia: The contribution of cognition and psychopathology. *Psychiatry Research*, *205*(3), 192–198. https://doi.org/10.1016/j.psychres.2012.08.038

Chapellier, V., Pavlidou, A., Maderthaner, L., von Känel, S., & Walther, S. (2022). The impact of poor nonverbal social perception on functional capacity in schizophrenia. *Frontiers in Psychology*, *13*, 804093. https://doi.org/10.3389/fpsyg.2022.804093

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: A tutorial review. *Attention, Perception, & Psychophysics*, *75*(5), 790–811. https://doi.org/10.3758/s13414-013-0475-4

Cheng, R.-K., Ali, Y. M., & Meck, W. H. (2007). Ketamine "unlocks" the reduced clock-speed effects of cocaine following extended training: Evidence for dopamine–glutamate interactions in timing and time perception. *Neurobiology of Learning and Memory*, *88*(2), 149–159. https://doi.org/10.1016/j.nlm.2007.04.005

Ciullo, V., Spalletta, G., Caltagirone, C., Jorge, R. E., & Piras, F. (2016). Explicit Time Deficit in Schizophrenia: Systematic Review and Meta-Analysis Indicate It Is Primary and Not Domain Specific. *Schizophrenia Bulletin*, *42*(2), 505–518. https://doi.org/10.1093/schbul/sbv104

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, *1242*, 126–135. https://doi.org/10.1016/j.brainres.2008.04.023

Corcoran, C. M., Keilp, J. G., Kayser, J., Klim, C., Butler, P. D., Bruder, G. E., Gur, R. C., & Javitt, D. C. (2015). Emotion recognition deficits as predictors of transition in individuals at clinical high risk for schizophrenia: A neurodevelopmental perspective. *Psychological Medicine*, *45*(14), 2959–2973. https://doi.org/10.1017/S0033291715000902

Cornblatt, B. A., Risch, N. J., Faris, G., Friedman, D., & Erlenmeyer-Kimling, L. (1988). The Continuous Performance Test, Identical Pairs version (CPT-IP): I. New findings about sustained attention in normal families. *Psychiatry Research*, *26*(2), 223–238. https://doi.org/10.1016/0165-1781(88)90076-5

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. *Schizophrenia Bulletin*, *42*(2), 434–442. https://doi.org/10.1093/schbul/sbv146

Dalal, T. C., Muller, A.-M., & Stevenson, R. (2021). The relationship between multisensory temporal processing and schizotypal traits. *Multisensory Research*, *34*(5), 511–529. https://doi.org/10.1163/22134808-bja10044

Darke, H., Sundram, S., Cropper, S. J., & Carter, O. (2021). Dynamic face processing impairments are associated with cognitive and positive psychotic symptoms across psychiatric disorders. *Npj Schizophrenia*, *7*(1), 36. https://doi.org/10.1038/s41537-021-00166-z

Davies, D. J., Teufel, C., & Fletcher, P. C. (2018). Anomalous perceptions and beliefs are associated with shifts toward different types of prior knowledge in perceptual inference. *Schizophrenia Bulletin*, *44*(6), 1245–1253. https://doi.org/10.1093/schbul/sbx177

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 Psychologica*, *147*, 136–142. https://doi.org/10.1016/j.actpsy.2013.06.013

de Gelder, B., Vroomen, J., Annen, L., Masthof, E., & Hodiamont, P. (2003). Audio-visual integration in schizophrenia. *Schizophrenia Research*, *59*(2–3), 211–218. 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., & Hodiamont, P. (2005). Multisensory integration of emotional faces and voices in schizophrenics. *Schizophrenia Research*, *72*(2–3), 195–203. https://doi.org/10.1016/j.schres.2004.02.013

de Jong, J. J., Hodiamont, P. P. G., & de Gelder, B. (2010). Modality-specific attention and multisensory integration of emotions in schizophrenia: Reduced regulatory effects. *Schizophrenia Research*, *122*(1–3), 136–143. https://doi.org/10.1016/j.schres.2010.04.010

de Jong, J. J., Hodiamont, P. P. G., Van den Stock, J., & de Gelder, B. (2009). Audiovisual emotion recognition in schizophrenia: Reduced integration of facial and vocal affect. *Schizophrenia Research*, *107*(2–3), 286–293. https://doi.org/10.1016/j.schres.2008.10.001

Denervaud, S., Gentaz, E., Matusz, P. J., & Murray, M. M. (2020). Multisensory gains in simple detection predict global cognition in schoolchildren. *Scientific Reports*, *10*(1), 1394. https://doi.org/10.1038/s41598-020-58329-4

Di Cosmo, G., Costantini, M., Ambrosini, E., Salone, A., Martinotti, G., Corbo, M., Di Giannantonio, M., & Ferri, F. (2021). Body-environment integration: Temporal processing of tactile and auditory inputs along the schizophrenia continuum. *Journal of Psychiatric Research*, *134*, 208–214. https://doi.org/10.1016/j.jpsychires.2020.12.034

Di Luca, M., Machulla, T. K., & Ernst, M. O. (2009). Recalibration of multisensory simultaneity: Cross-modal transfer coincides with a change in perceptual latency. *Journal of Vision*, *9*(12), 7–7. https://doi.org/10.1167/9.12.7

Dickson, H., Hedges, E. P., Ma, S. Y., Cullen, A. E., MacCabe, J. H., Kempton, M. J., Downs, J., & Laurens, K. R. (2020). Academic achievement and schizophrenia: A systematic meta-analysis. *Psychological Medicine*, *50*(12), 1949–1965. https://doi.org/10.1017/S0033291720002354

Doehrmann, O., & Naumer, M. J. (2008). Semantics and the multisensory brain: How meaning modulates processes of audio-visual integration. *Brain Research*, *1242*, 136–150. https://doi.org/10.1016/j.brainres.2008.03.071

Dondaine, T., Robert, G., PÃ©ron, J., Grandjean, D., VÃ©rin, M., Drapier, D., & Millet, B. (2014). Biases in facial and vocal emotion recognition in chronic schizophrenia. *Frontiers in Psychology*, *5*, 900. https://doi.org/10.3389/fpsyg.2014.00900

Driver, D. I., Gogtay, N., & Rapoport, J. L. (2013). Childhood onset schizophrenia and early onset schizophrenia spectrum disorders. *Child and Adolescent Psychiatric Clinics of North America*, *22*(4), 539–555. https://doi.org/10.1016/j.chc.2013.04.001

Edwards, J., Jackson, H. J., & Pattison, P. E. (2002). Emotion recognition via facial expression and affective prosody in schizophrenia: A methodological review. *Clinical Psychology Review*. https://doi.org/10.1016/S0272-7358(02)00130-7

Edwards, J., Pattison, P. E., Jackson, H. J., & Wales, R. J. (2001). Facial affect and affective prosody recognition in first-episode schizophrenia. *Schizophrenia Research*, *48*(2–3), 235–253. https://doi.org/10.1016/S0920-9964(00)00099-2

Fakra, E., Jouve, E., Guillaume, F., Azorin, J.-M., & Blin, O. (2015). Relation between facial affect recognition and configural face processing in antipsychotic-free schizophrenia. *Neuropsychology*, *29*(2), 197–204. https://doi.org/10.1037/neu0000136

Fengler, I., Nava, E., Villwock, A. K., Büchner, A., Lenarz, T., & Röder, B. (2017). Multisensory emotion perception in congenitally, early, and late deaf CI users. *PLOS ONE*, *12*(10), e0185821. https://doi.org/10.1371/journal.pone.0185821

Ferri, F., Ambrosini, E., & Costantini, M. (2016). Spatiotemporal processing of somatosensory stimuli in schizotypy. *Scientific Reports*, *6*(1), 38735. https://doi.org/10.1038/srep38735

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. *Schizophrenia Research*, *152*(1), 51–57. https://doi.org/10.1016/j.schres.2013.06.026

Ferri, F., Nikolova, Y. S., Perrucci, M. G., Costantini, M., Ferretti, A., Gatta, V., Huang, Z., Edden, R. A. E., Yue, Q., D'Aurora, M., Sibille, E., Stuppia, L., Romani, G. L., &

Northoff, G. (2017). A neural “tuning curve” for multisensory experience and cognitive-perceptual schizotypy. *Schizophrenia Bulletin*, *43*(4), 801–813. https://doi.org/10.1093/schbul/sbw174

Ferri, F., Venskus, A., Fotia, F., Cooke, J., & Romei, V. (2018). Higher proneness to multisensory illusions is driven by reduced temporal sensitivity in people with high schizotypal traits. *Consciousness and Cognition*, *65*, 263–270. https://doi.org/10.1016/j.concog.2018.09.006

Fiszdon, J. M., & Bell, M. D. (2009). Effects of presentation modality and valence on affect recognition performance in schizophrenia and healthy controls. *Psychiatry Research*, *170*(2–3), 114–118. https://doi.org/10.1016/j.psychres.2008.11.014

Föcker, J., Gondan, M., & Röder, B. (2011). Preattentive processing of audio-visual emotional signals. *Acta Psychologica*, *137*(1), 36–47. https://doi.org/10.1016/j.actpsy.2011.02.004

Föcker, J., & Röder, B. (2019). Event-related potentials reveal evidence for late integration of emotional prosody and facial expression in dynamic stimuli: An ERP study. *Multisensory Research*, *32*(6), 473–497. https://doi.org/10.1163/22134808-20191332

Foucher, J. R., Lacambre, M., Pham, B.-T., Giersch, A., & Elliott, M. A. (2007). Low time resolution in schizophrenia. *Schizophrenia Research*, *97*(1–3), 118–127. https://doi.org/10.1016/j.schres.2007.08.013

Fowler, D., Freeman, D., Smith, B., Kuipers, E., Bebbington, P., Bashforth, H., Coker, S., Hodgekins, J., Gracie, A., Dunn, G., & Garety, P. (2006). The Brief Core Schema Scales (BCSS): Psychometric properties and associations with paranoia and grandiosity in non-clinical and psychosis samples. *Psychological Medicine*, *36*(6), 749–759. https://doi.org/10.1017/S0033291706007355

Foxe, J. J., & Molholm, S. (2009). Ten years at the multisensory forum: Musings on the evolution of a field. *Brain Topography*, *21*(3–4), 149–154. https://doi.org/10.1007/s10548-009-0102-9

Freeman, D. (2007). Suspicious minds: The psychology of persecutory delusions. *Clinical Psychology Review*, *27*(4), 425–457. https://doi.org/10.1016/j.cpr.2006.10.004

Freeman, D. (2016). Persecutory delusions: A cognitive perspective on understanding and treatment. *The Lancet Psychiatry*, *3*(7), 685–692. https://doi.org/10.1016/S2215-0366(16)00066-3

Freeman, D., Garety, P. A., Bebbington, P. E., Smith, B., Rollinson, R., Fowler, D., Kuipers, E., Ray, K., & Dunn, G. (2005). Psychological investigation of the structure of

paranoia in a non-clinical population. *British Journal of Psychiatry*, *186*(5), 427–435. https://doi.org/10.1192/bjp.186.5.427

Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., & Bebbington, P. E. (2002). A cognitive model of persecutory delusions. *British Journal of Clinical Psychology*, *41*(4), 331–347. https://doi.org/10.1348/014466502760387461

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K., Brown, H. R., Siemerkus, J., & Stephan, K. E. (2016). The dysconnection hypothesis (2016). *Schizophrenia Research*, *176*(2–3), 83–94. https://doi.org/10.1016/j.schres.2016.07.014

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, *7*(7), 773–778. https://doi.org/10.1038/nn1268

Gaebel, W., & Zielasek, J. (2015). Focus on psychosis. *Dialogues in Clinical Neuroscience*, *17*(1), 9–18. https://doi.org/10.31887/DCNS.2015.17.1/wgaebel

Gao, C., Wedell, D. H., & Shinkareva, S. V. (2021). Evaluating non-affective cross-modal congruence effects on emotion perception. *Cognition and Emotion*, *35*(8), 1634–1651. https://doi.org/10.1080/02699931.2021.1973966

Garcia-Leon, M. A., Fuentes-Claramonte, P., Valiente-Gómez, A., Natividad, C., Salgado-Pineda, P., Gomar, J. J., Guerrero-Pedraza, A., Portillo, F., Ortiz-Gil, J., Alonso-Lana, S., Maristany, T., Raduà, J., Salvador, R., Sarró, S., & Pomarol-Clotet, E. (2021). Altered brain responses to specific negative emotions in schizophrenia. *NeuroImage: Clinical*, *32*, 102894. https://doi.org/10.1016/j.nicl.2021.102894

García-Pérez, M. A., & Alcalá-Quintana, R. (2012). On the discrepant results in synchrony judgment and temporal-order judgment tasks: A quantitative model. *Psychonomic Bulletin & Review*, *19*(5), 820–846. https://doi.org/10.3758/s13423-012-0278-y

Garety, P. A., Kuipers, E., Fowler, D., Freeman, D., & Bebbington, P. E. (2001). A cognitive model of the positive symptoms of psychosis. *Psychological Medicine*, *31*(2), 189–195. https://doi.org/10.1017/S0033291701003312

Gebreegziabhere, Y., Habatmu, K., Mihretu, A., Cella, M., & Alem, A. (2022). Cognitive impairment in people with schizophrenia: An umbrella review. *European Archives of Psychiatry and Clinical Neuroscience*, *272*(7), 1139–1155. https://doi.org/10.1007/s00406-022-01416-6

Germine, L., Benson, T. L., Cohen, F., & Hooker, C. I. (2013). Psychosis-proneness and the rubber hand illusion of body ownership. *Psychiatry Research*, *207*(1–2), 45–52. https://doi.org/10.1016/j.psychres.2012.11.022

Giannitelli, M., Xavier, J., François, A., Bodeau, N., Laurent, C., Cohen, D., & Chaby, L. (2015). Facial, vocal and cross-modal emotion processing in early-onset schizophrenia spectrum disorders. *Schizophrenia Research*, *168*(1–2), 252–259. https://doi.org/10.1016/j.schres.2015.07.039

Giersch, A., Lalanne, L., Corves, C., Seubert, J., Shi, Z., Foucher, J., & Elliott, M. A. (2009). Extended visual simultaneity thresholds in patients with schizophrenia. *Schizophrenia Bulletin*, *35*(4), 816–825. https://doi.org/10.1093/schbul/sbn016

Gold, R., Butler, P., Revheim, N., Leitman, D. I., Hansen, J. A., Gur, R. C., Kantrowitz, J. T., Laukka, P., Juslin, P. N., Silipo, G. S., & Javitt, D. C. (2012). Auditory emotion recognition impairments in schizophrenia: Relationship to acoustic features and cognition. *American Journal of Psychiatry*, *169*(4), 424–432. https://doi.org/10.1176/appi.ajp.2011.11081230

Gong, B., Li, Q., Zhao, Y., & Wu, C. (2021). Auditory emotion recognition deficits in schizophrenia: A systematic review and meta-analysis. *Asian Journal of Psychiatry*, *65*, 102820. https://doi.org/10.1016/j.ajp.2021.102820

Green, M. J., & Phillips, M. L. (2004). Social threat perception and the evolution of paranoia. *Neuroscience & Biobehavioral Reviews*, *28*(3), 333–342. https://doi.org/10.1016/j.neubiorev.2004.03.006

Green, M. J., Williams, L. M., & Hemsley, D. R. (2000). Cognitive theories of delusion formation: The contribution of visual scanpath research. *Cognitive Neuropsychiatry*, *5*(1), 63–74. https://doi.org/10.1080/135468000395835

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112. https://doi.org/10.1007/BF02289823

Grimsen, C., Brand, A., & Fahle, M. (2013). No evidence for prolonged visible persistence in patients with schizophrenia. *PLoS ONE*, *8*(3), e58940. https://doi.org/10.1371/journal.pone.0058940

Gröhn, C., Norgren, E., & Eriksson, L. (2022). A systematic review of the neural correlates of multisensory integration in schizophrenia. *Schizophrenia Research: Cognition*, *27*, 100219. https://doi.org/10.1016/j.scog.2021.100219

Haddock, G., McCarron, J., Tarrier, N., & Faragher, E. B. (1999). Scales to measure dimensions of hallucinations and delusions: The Psychotic Symptom Rating Scales

(PSYRATS). *Psychological Medicine*, *29*(4), 879–889. https://doi.org/10.1017/S0033291799008661

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382–385. https://doi.org/10.1038/nn827

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: Hemodynamic response function abnormalities. *Human Brain Mapping*, *37*(2), 745–755. https://doi.org/10.1002/hbm.23063

Hardoy, M. C., Carta, M. G., Catena, M., Hardoy, M. J., Cadeddu, M., Dell'Osso, L., Hugdahl, K., & Carpiniello, B. (2004). Impairment in visual and spatial perception in schizophrenia and delusional disorder. *Psychiatry Research*, *127*(1–2), 163–166. https://doi.org/10.1016/j.psychres.2004.03.002

Hartcher-O'Brien, J., Brighouse, C., & Levitan, C. A. (2016). A single mechanism account of duration and rate processing via the pacemaker–accumulator and beat frequency models. *Current Opinion in Behavioral Sciences*, *8*, 268–275. https://doi.org/10.1016/j.cobeha.2016.02.026

Haß, K., Sinke, C., Reese, T., Roy, M., Wiswede, D., Dillo, W., Oranje, B., & Szycik, G. R. (2017). Enlarged temporal integration window in schizophrenia indicated by the double-flash illusion. *Cognitive Neuropsychiatry*, *22*(2), 145–158. https://doi.org/10.1080/13546805.2017.1287693

He, Y., Steines, M., Sammer, G., Nagels, A., Kircher, T., & Straube, B. (2021). Modality-specific dysfunctional neural processing of social-abstract and non-social-concrete information in schizophrenia. *NeuroImage: Clinical*, *29*, 102568. https://doi.org/10.1016/j.nicl.2021.102568

Healey, K. M., Bartholomeusz, C. F., & Penn, D. L. (2016). Deficits in social cognition in first episode psychosis: A review of the literature. *Clinical Psychology Review*, *50*, 108–137. https://doi.org/10.1016/j.cpr.2016.10.001

Heinz, A., Murray, G. K., Schlagenhauf, F., Sterzer, P., Grace, A. A., & Waltz, J. A. (2019). Owards a unifying cognitive, neurophysiological, and computational neuroscience account of schizophrenia. *Schizophrenia Bulletin*, *45*(5), 1092–1100. https://doi.org/10.1093/schbul/sby154

Hillmann, T. E., Ascone, L., Kempkensteffen, J., & Lincoln, T. M. (2017). Scanning to conclusions? Visual attention to neutral faces under stress in individuals with and

without subclinical paranoia. *Journal of Behavior Therapy and Experimental Psychiatry*, *56*, 137–143. https://doi.org/10.1016/j.jbtep.2016.08.013

Hillmann, T. E., Kempkensteffen, J., & Lincoln, T. M. (2015). Visual attention to threat-related faces and delusion-proneness: An eye tracking study using dynamic stimuli. *Cognitive Therapy and Research*, *39*(6), 808–815. https://doi.org/10.1007/s10608-015-9699-z

Hillmann, T. E., Riehle, M., & Lincoln, T. M. (2018). Reduced scanning of salient facial features mediates the association between paranoia and emotion recognition. *Psychiatry Research*, *269*, 430–436. https://doi.org/10.1016/j.psychres.2018.08.052

Hornix, B. E., Havekes, R., & Kas, M. J. H. (2019). Multisensory cortical processing and dysfunction across the neuropsychiatric spectrum. *Neuroscience & Biobehavioral Reviews*, *97*, 138–151. https://doi.org/10.1016/j.neubiorev.2018.02.010

Howes, O. D., & Kapur, S. (2009). The dopamine hypothesis of schizophrenia: Version III--The final common pathway. *Schizophrenia Bulletin*, *35*(3), 549–562. https://doi.org/10.1093/schbul/sbp006

Howes, O. D., & Murray, R. M. (2014). Schizophrenia: An integrated sociodevelopmental-cognitive model. *The Lancet*, *383*(9929), 1677–1687. https://doi.org/10.1016/S0140-6736(13)62036-X

Hsu, J. C. (1996). *Multiple comparisons: Theory and methods*. Springer US. https://doi.org/10.1007/978-1-4899-7180-7

Humpston, C. S., Adams, R. A., Benrimoh, D., Broome, M. R., Corlett, P. R., Gerrans, P., Horga, G., Parr, T., Pienkos, E., Powers, A. R., Raballo, A., Rosen, C., & Linden, D. E. J. (2019). From computation to the first-person: Auditory-verbal hallucinations and delusions of thought interference in schizophrenia-spectrum psychoses. *Schizophrenia Bulletin*, *45*(Supplement_1), S56–S66. https://doi.org/10.1093/schbul/sby073

Huq, S. F., Garety, P. A., & Hemsley, D. R. (1988). Probabilistic judgements in deluded and non-deluded subjects. *The Quarterly Journal of Experimental Psychology Section A*, *40*(4), 801–812. https://doi.org/10.1080/14640748808402300

Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, *1*(1), 69–82. https://doi.org/10.3102/10769986001001069

Irani, F., Seligman, S., Kamath, V., Kohler, C., & Gur, R. C. (2012). A meta-analysis of emotion perception and functional outcomes in schizophrenia. *Schizophrenia Research*, *137*(1–3), 203–211. https://doi.org/10.1016/j.schres.2012.01.023

Izen, S. C., & Ciaramitaro, V. M. (2020). A crowd of emotional voices influences the perception of emotional faces: Using adaptation, stimulus salience, and attention to probe audio-visual interactions for emotional stimuli. *Attention, Perception, & Psychophysics*, *82*(8), 3973–3992. https://doi.org/10.3758/s13414-020-02104-0

Javitt, D. C. (2009a). Sensory processing in schizophrenia: Neither simple nor intact. *Schizophrenia Bulletin*, *35*(6), 1059–1064. https://doi.org/10.1093/schbul/sbp110

Javitt, D. C. (2009b). When doors of perception close: Bottom-up models of disrupted cognition in schizophrenia. *Annual Review of Clinical Psychology*, *5*(1), 249–275. https://doi.org/10.1146/annurev.clinpsy.032408.153502

Javitt, D. C., & Freedman, R. (2015). Sensory processing dysfunction in the personal experience and neuronal machinery of schizophrenia. *American Journal of Psychiatry*, *172*(1), 17–31. https://doi.org/10.1176/appi.ajp.2014.13121691

Jeong, J. W., Kim, H. T., Lee, S.-H., & Lee, H. (2021). Effects of an audiovisual emotion perception training for schizophrenia: A preliminary study. *Frontiers in Psychiatry*, *12*, 522094. https://doi.org/10.3389/fpsyt.2021.522094

Johnston, P. J., Enticott, P. G., Mayes, A. K., Hoy, K. E., Herring, S. E., & Fitzgerald, P. B. (2010). Symptom correlates of static and dynamic facial affect processing in schizophrenia: Evidence of a double dissociation. *Schizophrenia Bulletin*, *36*(4), 680–687. https://doi.org/10.1093/schbul/sbn136

Jones, S. A., & Noppeney, U. (2021). Ageing and multisensory integration: A review of the evidence, and a computational perspective. *Cortex*, *138*, 1–23. https://doi.org/10.1016/j.cortex.2021.02.001

Just, M. A. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity. *Brain*, *127*(8), 1811–1821. https://doi.org/10.1093/brain/awh199

Kállai, J., Hegedüs, G., Feldmann, Á., Rózsa, S., Darnai, G., Herold, R., Dorn, K., Kincses, P., Csathó, Á., & Szolcsányi, T. (2015). Temperament and psychopathological syndromes specific susceptibility for rubber hand illusion. *Psychiatry Research*, *229*(1–2), 410–419. https://doi.org/10.1016/j.psychres.2015.05.109

Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, *160*(1), 13–23. https://doi.org/10.1176/appi.ajp.160.1.13

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261–276. https://doi.org/10.1093/schbul/13.2.261

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. *Schizophrenia Research*, *60*(2–3), 173–181. https://doi.org/10.1016/S0920-9964(02)00299-2

King, B. H., & Lord, C. (2011). Is schizophrenia on the autism spectrum? *Brain Research*, *1380*, 34–41. https://doi.org/10.1016/j.brainres.2010.11.031

Klaver, M., & Dijkerman, H. C. (2016). Bodily experience in schizophrenia: Factors underlying a disturbed sense of body ownership. *Frontiers in Human Neuroscience*, *10*, 305. https://doi.org/10.3389/fnhum.2016.00305

Koelewijn, T., Bronkhorst, A., & Theeuwes, J. (2010). Attention and the multiple stages of multisensory integration: A review of audiovisual studies. *Acta Psychologica*, *134*(3), 372–384. https://doi.org/10.1016/j.actpsy.2010.03.010

Kohler, C. G., Turner, T. H., Bilker, W. B., Brensinger, C. M., Siegel, S. J., Kanes, S. J., Gur, R. E., & Gur, R. C. (2003). Facial emotion recognition in schizophrenia: Intensity effects and error pattern. *American Journal of Psychiatry*, *160*(10), 1768–1774. https://doi.org/10.1176/appi.ajp.160.10.1768

Kohler, C. G., Walker, J. B., Martin, E. A., Healey, K. M., & Moberg, P. J. (2010). Facial emotion perception in schizophrenia: A meta-analytic review. *Schizophrenia Bulletin*, *36*(5), 1009–1019. https://doi.org/10.1093/schbul/sbn192

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS ONE*, *2*(9), e943. https://doi.org/10.1371/journal.pone.0000943

Kozhuharova, P., Saviola, F., Ettinger, U., & Allen, P. (2020). Neural correlates of social cognition in populations at risk of psychosis: A systematic review. *Neuroscience & Biobehavioral Reviews*, *108*, 94–111. https://doi.org/10.1016/j.neubiorev.2019.10.010

Krkovic, K., Clamor, A., Schlier, B., & Lincoln, T. M. (2020). Emotions and persecutory ideation in daily life: On the trail of the "chicken and egg" problem. *Journal of Abnormal Psychology*, *129*(2), 215–223. https://doi.org/10.1037/abn0000495

Lakis, N., & Mendrek, A. (2013). Individuals diagnosed with schizophrenia assign emotional importance to neutral stimuli: An fMRI study. *ISRN Psychiatry*, *2013*, 1–7. https://doi.org/10.1155/2013/965428

Lencer, R., Meermeier, A., Silling, K., Gremmler, S., & Lappe, M. (2017). Instability of visual error processing for sensorimotor adaptation in schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience*, *267*(3), 237–244. https://doi.org/10.1007/s00406-016-0716-3

Leucht, S., Kane, J., Kissling, W., Hamann, J., Etschel, E., & Engel, R. (2005). What does the PANSS mean? *Schizophrenia Research*, *79*(2–3), 231–238. https://doi.org/10.1016/j.schres.2005.04.008

Limongi, R., Bohaterewicz, B., Nowicka, M., Plewka, A., & Friston, K. J. (2018). Knowing when to stop: Aberrant precision and evidence accumulation in schizophrenia. *Schizophrenia Research*, *197*, 386–391. https://doi.org/10.1016/j.schres.2017.12.018

Lin, Y., Ding, H., & Zhang, Y. (2018). Emotional prosody processing in schizophrenic patients: A selective review and meta-analysis. *Journal of Clinical Medicine*, *7*(10), 363. https://doi.org/10.3390/jcm7100363

Lin, Y., Ding, H., & Zhang, Y. (2020). Multisensory integration of emotion in schizophrenic patients. *Multisensory Research*, *33*(8), 865–901. https://doi.org/10.1163/22134808-bja10016

Lincoln, T. M., Keller, E., & Rief, W. (2009). Die Erfassung von Wahn und Halluzinationen in der Normalbevölkerung: Deutsche Adaptationen des Peters et al. Delusions Inventory (PDI) und der Launay Slade Hallucination Scale (LSHS-R). *Diagnostica*, *55*(1), 29–40. https://doi.org/10.1026/0012-1924.55.1.29

Liu, T., Pinheiro, A. P., Zhao, Z., Nestor, P. G., McCarley, R. W., & Niznikiewicz, M. (2016). Simultaneous face and voice processing in schizophrenia. *Behavioural Brain Research*, *305*, 76–86. https://doi.org/10.1016/j.bbr.2016.01.039

Lukow, P. B., Kiemes, A., Kempton, M. J., Turkheimer, F. E., McGuire, P., & Modinos, G. (2021). Neural correlates of emotional processing in psychosis risk and onset – A systematic review and meta-analysis of fMRI studies. *Neuroscience & Biobehavioral Reviews*, *128*, 780–788. https://doi.org/10.1016/j.neubiorev.2021.03.010

Mangelinckx, C., Belge, J. B., Maurage, P., & Constant, E. (2017). Impaired facial and vocal emotion decoding in schizophrenia is underpinned by basic perceptivo-motor deficits. *Cognitive Neuropsychiatry*, *22*(6), 461–467. https://doi.org/10.1080/13546805.2017.1382342

Mani, T., Bedwell, J., & Miller, L. (2005). Age-related decrements in performance on a brief continuous performance test. *Archives of Clinical Neuropsychology*, *20*(5), 575–586. https://doi.org/10.1016/j.acn.2004.12.008

Marsicano, G., Cerpelloni, F., Melcher, D., & Ronconi, L. (2022). Lower multisensory temporal acuity in individuals with high schizotypal traits: A web-based study. *Scientific Reports*, *12*(1), 2782. https://doi.org/10.1038/s41598-022-06503-1

Martin, B., Giersch, A., Huron, C., & van Wassenhove, V. (2013). Temporal event structure and timing in schizophrenia: Preserved binding in a longer "now". *Neuropsychologia*, *51*(2), 358–371. https://doi.org/10.1016/j.neuropsychologia.2012.07.002

Martin, D., Croft, J., Pitt, A., Strelchuk, D., Sullivan, S., & Zammit, S. (2020). Systematic review and meta-analysis of the relationship between genetic risk for schizophrenia and facial emotion recognition. *Schizophrenia Research*, *218*, 7–13. https://doi.org/10.1016/j.schres.2019.12.031

Mathys, C. (2011). A Bayesian foundation for individual learning under uncertainty. *Frontiers in Human Neuroscience*, *5*, 39. https://doi.org/10.3389/fnhum.2011.00039

McAusland, L., Buchy, L., Cadenhead, K. S., Cannon, T. D., Cornblatt, B. A., Heinssen, R., McGlashan, T. H., Perkins, D. O., Seidman, L. J., Tsuang, M. T., Walker, E. F., Woods, S. W., Bearden, C. E., Mathalon, D. H., & Addington, J. (2017). Anxiety in youth at clinical high risk for psychosis: Anxiety in clinical high risk. *Early Intervention in Psychiatry*, *11*(6), 480–487. https://doi.org/10.1111/eip.12274

McGhie, A., & Chapman, J. (1961). Disorders of attention and perception in early schizophrenia. *British Journal of Medical Psychology*, *34*(2), 103–116. https://doi.org/10.1111/j.2044-8341.1961.tb00936.x

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. https://doi.org/10.1038/264746a0

Meredith, M. A., Nemitz, J., & Stein, B. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *The Journal of Neuroscience*, *7*(10), 3215–3229. https://doi.org/10.1523/JNEUROSCI.07-10-03215.1987

Meredith, M. A., & Stein, B. E. (1986). Spatial factors determine the activity of muitisensory neurons in cat superior colliculus. *Brain Research*, *365*(2), 350–354. https://doi.org/10.1016/0006-8993(86)91648-3

Mesholam-Gately, R. I., Giuliano, A. J., Goff, K. P., Faraone, S. V., & Seidman, L. J. (2009). Neurocognition in first-episode schizophrenia: A meta-analytic review. *Neuropsychology*, *23*(3), 315–336. https://doi.org/10.1037/a0014708

Mettler, F. A. (1955). Perceptual capacity, functions of the corpus striatum and schizophrenia. *The Psychiatric Quarterly*, *29*(1–4), 89–111. https://doi.org/10.1007/BF01567442

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*(2), 247–279. https://doi.org/10.1016/0010-0285(82)90010-X

Mitrovic, M., Ristic, M., Dimitrijevic, B., & Hadzi Pesic, M. (2020). Facial emotion recognition and persecutory ideation in paranoid schizophrenia. *Psychological Reports*, *123*(4), 1099–1116. https://doi.org/10.1177/0033294119849016

Moberget, T., & Ivry, R. B. (2019). Prediction, psychosis, and the cerebellum. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*(9), 820–831. https://doi.org/10.1016/j.bpsc.2019.06.001

Modinos, G., Kempton, M. J., Tognin, S., Calem, M., Porffy, L., Antoniades, M., Mason, A., Azis, M., Allen, P., Nelson, B., McGorry, P., Pantelis, C., Riecher-Rössler, A., Borgwardt, S., Bressan, R., Barrantes-Vidal, N., Krebs, M.-O., Nordentoft, M., Glenthøj, B., … for the EU-GEI High Risk Study Group. (2020). Association of adverse outcomes with emotion processing and its neural substrate in individuals at clinical high risk for psychosis. *JAMA Psychiatry*, *77*(2), 190. https://doi.org/10.1001/jamapsychiatry.2019.3501

Modinos, G., Tseng, H.-H., Falkenberg, I., Samson, C., McGuire, P., & Allen, P. (2015). Neural correlates of aberrant emotional salience predict psychotic symptoms and global functioning in high-risk and first-episode psychosis. *Social Cognitive and Affective Neuroscience*, *10*(10), 1429–1436. https://doi.org/10.1093/scan/nsv035

Moran, J. K., Keil, J., Masurovsky, A., Gutwinski, S., Montag, C., & Senkowski, D. (2021). Multisensory processing can compensate for top-down attention deficits in schizophrenia. *Cerebral Cortex*, *31*(12), 5536–5548. https://doi.org/10.1093/cercor/bhab177

Moreno-Küstner, B., Martín, C., & Pastor, L. (2018). Prevalence of psychotic disorders and its association with methodological issues. A systematic review and meta-analyses. *PLOS ONE*, *13*(4), e0195687. https://doi.org/10.1371/journal.pone.0195687

Muller, A.-M., Dalal, T. C., & Stevenson, R. A. (2020). Schizotypal traits are not related to multisensory integration or audiovisual speech perception. *Consciousness and Cognition*, *86*, 103030. https://doi.org/10.1016/j.concog.2020.103030

Muller, A.-M., Dalal, T. C., & Stevenson, R. A. (2021). Schizotypal personality traits and multisensory integration: An investigation using the McGurk effect. *Acta Psychologica*, *218*, 103354. https://doi.org/10.1016/j.actpsy.2021.103354

Müller, V. I., Kellermann, T. S., Seligman, S. C., Turetsky, B. I., & Eickhoff, S. B. (2014). Modulation of affective face processing deficits in Schizophrenia by congruent emotional sounds. *Social Cognitive and Affective Neuroscience*, *9*(4), 436–444. https://doi.org/10.1093/scan/nss107

Murray, M. M., Lewkowicz, D. J., Amedi, A., & Wallace, M. T. (2016). Multisensory processes: A balancing act across the lifespan. *Trends in Neurosciences*, *39*(8), 567–579. https://doi.org/10.1016/j.tins.2016.05.003

Nath, A. R., & Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *NeuroImage*, *59*(1), 781–787. https://doi.org/10.1016/j.neuroimage.2011.07.024

Noel, J.-P., Cascio, C. J., Wallace, M. T., & Park, S. (2017). The spatial self in schizophrenia and autism spectrum disorder. *Schizophrenia Research*, *179*, 8–12. https://doi.org/10.1016/j.schres.2016.09.021

Noel, J.-P., De Niear, M. A., Stevenson, R., Alais, D., & Wallace, M. T. (2017). Atypical rapid audio-visual temporal recalibration in autism spectrum disorders: Audiovisual temporal recalibration in ASD. *Autism Research*, *10*(1), 121–129. https://doi.org/10.1002/aur.1633

Noel, J.-P., De Niear, M., Van der Burg, E., & Wallace, M. T. (2016). Audiovisual simultaneity judgment and rapid recalibration throughout the lifespan. *PLOS ONE*, *11*(8), e0161698. https://doi.org/10.1371/journal.pone.0161698

Noel, J.-P., Failla, M. D., Quinde-Zlibut, J. M., Williams, Z. J., Gerdes, M., Tracy, J. M., Zoltowski, A. R., Foss-Feig, J. H., Nichols, H., Armstrong, K., Heckers, S. H., Blake, R. R., Wallace, M. T., Park, S., & Cascio, C. J. (2020). Visual-tactile spatial multisensory interaction in adults with autism and schizophrenia. *Frontiers in Psychiatry*, *11*, 578401. https://doi.org/10.3389/fpsyt.2020.578401

Noel, J.-P., Stevenson, R. A., & Wallace, M. T. (2018). Atypical audiovisual temporal function in autism and schizophrenia: Similar phenotype, different cause. *European Journal of Neuroscience*, *47*(10), 1230–1241. https://doi.org/10.1111/ejn.13911

Nuechterlein, K. H., Barch, D. M., Gold, J. M., Goldberg, T. E., Green, M. F., & Heaton, R. K. (2004). Identification of separable cognitive factors in schizophrenia. *Schizophrenia Research*, *72*(1), 29–39. https://doi.org/10.1016/j.schres.2004.09.007

Nuechterlein, K. H., Green, M. F., Kern, R. S., Baade, L. E., Barch, D. M., Cohen, J. D., Essock, S., Fenton, W. S., Frese, F. J., Gold, J. M., Goldberg, T., Heaton, R. K., Keefe, R. S. E., Kraemer, H., Mesholam-Gately, R., Seidman, L. J., Stover, E.,

Weinberger, D. R., Young, A. S., … Marder, S. R. (2008). The MATRICS Consensus Cognitive Battery, Part 1: Test selection, reliability, and validity. *American Journal of Psychiatry*, *165*(2), 203–213. https://doi.org/10.1176/appi.ajp.2007.07010042

Odegaard, B., & Shams, L. (2016). The brain's tendency to bind audiovisual signals is stable but not general. *Psychological Science*, *27*(4), 583–591. https://doi.org/10.1177/0956797616628860

Odegaard, B., & Shams, L. (2017). The relationship between audiovisual binding tendencies and prodromal features of schizophrenia in the general population. *Clinical Psychological Science*, *5*(4), 733–741. https://doi.org/10.1177/2167702617704014

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Park, H., & Kayser, C. (2019). Shared neural underpinnings of multisensory integration and trial-by-trial perceptual recalibration in humans. *ELife*, *8*, e47001. https://doi.org/10.7554/eLife.47001

Pearl, D., Yodashkin-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. *Comprehensive Psychiatry*, *50*(2), 186–192. https://doi.org/10.1016/j.comppsych.2008.06.004

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, *16*(10), 504–510. https://doi.org/10.1016/j.tics.2012.08.009

Pena-Garijo, J., Lacruz, M., Masanet, M. J., Palop-Grau, A., Plaza, R., Hernandez-Merino, A., Edo-Villamon, S., & Valllina, O. (2023). Specific facial emotion recognition deficits across the course of psychosis: A comparison of individuals with low-risk, high-risk, first-episode psychosis and multi-episode schizophrenia-spectrum disorders. *Psychiatry Research*, *320*, 115029. https://doi.org/10.1016/j.psychres.2022.115029

Perrin, M. A., Butler, P. D., DiCostanzo, J., Forchelli, G., Silipo, G., & Javitt, D. C. (2010). Spatial localization deficits and auditory cortical dysfunction in schizophrenia. *Schizophrenia Research*, *124*(1–3), 161–168. https://doi.org/10.1016/j.schres.2010.06.004

Pinkham, A. E., Brensinger, C., Kohler, C., Gur, R. E., & Gur, R. C. (2011). Actively paranoid patients with schizophrenia over attribute anger to neutral faces.

*Schizophrenia Research*, *125*(2–3), 174–178. https://doi.org/10.1016/j.schres.2010.11.006

Pinkham, A. E., Penn, D. L., Perkins, D. O., & Lieberman, J. (2003). Implications for the neural basis of social cognition for the study of schizophrenia. *American Journal of Psychiatry*, *160*(5), 815–824. https://doi.org/10.1176/appi.ajp.160.5.815

Piwek, L., Pollick, F., & Petrini, K. (2015). Audiovisual integration of emotional signals from others' social interactions. *Frontiers in Psychology*, *9*, 116. https://doi.org/10.3389/fpsyg.2015.00611

Poletti, M., Tortorella, A., & Raballo, A. (2019). Impaired corollary discharge in psychosis and at-risk states: Integrating neurodevelopmental, phenomenological, and clinical perspectives. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*(9), 832–841. https://doi.org/10.1016/j.bpsc.2019.05.008

Portnova, G., Maslennikova, A., Zakharova, N., & Martynova, O. (2021). The deficit of multimodal perception of congruent and non-congruent fearful expressions in patients with schizophrenia: The ERP study. *Brain Sciences*, *11*(1), 96. https://doi.org/10.3390/brainsci11010096

Postmes, L., Sno, H. N., Goedhart, S., van der Stel, J., Heering, H. D., & de Haan, L. (2014). Schizophrenia as a self-disorder due to perceptual incoherence. *Schizophrenia Research*, *152*(1), 41–50. https://doi.org/10.1016/j.schres.2013.07.027

Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. *Science*, *357*(6351), 596–600. https://doi.org/10.1126/science.aan3458

Prikken, M., van der Weiden, A., Baalbergen, H., Hillegers, M. H. J., Kahn, R. S., Aarts, H., & van Haren, N. E. M. (2019). Multisensory integration underlying body-ownership experiences in schizophrenia and offspring of patients: A study using the rubber hand illusion paradigm. *Journal of Psychiatry & Neuroscience*, *44*(3), 177–184. https://doi.org/10.1503/jpn.180049

Radeau, M., & Bertelson, P. (1974). The after-effects of ventriloquism. *Quarterly Journal of Experimental Psychology*, *26*(1), 63–71. https://doi.org/10.1080/14640747408400388

Reininghaus, U., Kempton, M. J., Valmaggia, L., Craig, T. K. J., Garety, P., Onyejiaka, A., Gayer-Anderson, C., So, S. H., Hubbard, K., Beards, S., Dazzan, P., Pariante, C., Mondelli, V., Fisher, H. L., Mills, J. G., Viechtbauer, W., McGuire, P., van Os, J., Murray, R. M., … Morgan, C. (2016). Stress sensitivity, aberrant salience, and threat

anticipation in early psychosis: An experience sampling study. *Schizophrenia Bulletin*, *42*(3), 712–722. https://doi.org/10.1093/schbul/sbv190

Reitan, R. M., & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery: Theory and interpretation*. Neuropsychology Press.

Riemer, M. (2018). Delusions of control in schizophrenia: Resistant to the mind's best trick? *Schizophrenia Research*, *197*, 98–103. https://doi.org/10.1016/j.schres.2017.11.032

Roa Romero, Y., Keil, J., Balz, J., Gallinat, J., & Senkowski, D. (2016). Reduced frontal theta oscillations indicate altered crossmodal prediction error processing in schizophrenia. *Journal of Neurophysiology*, *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. (2016). Alpha-band oscillations reflect altered multisensory processing of the McGurk illusion in schizophrenia. *Frontiers in Human Neuroscience*, *10*, 41. https://doi.org/10.3389/fnhum.2016.00041

Rodewald, K., Bartolovic, M., Debelak, R., Aschenbrenner, S., Weisbrod, M., & Roesch-Ely, D. (2012). Eine Normierungsstudie eines modifizierten Trail Making Tests im deutschsprachigen Raum. *Zeitschrift für Neuropsychologie*, *23*(1), 37–48. https://doi.org/10.1024/1016-264X/a000060

Rohe, T., & Noppeney, U. (2015). Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLOS Biology*, *13*(2), e1002073. https://doi.org/10.1371/journal.pbio.1002073

Rohlf, S., Li, L., Bruns, P., & Röder, B. (2020). Multisensory integration develops prior to crossmodal recalibration. *Current Biology*, *30*(9), 1726-1732.e7. https://doi.org/10.1016/j.cub.2020.02.048

Rösler, L., Rolfs, M., van der Stigchel, S., Neggers, S. F. W., Cahn, W., Kahn, R. S., & Thakkar, K. N. (2015). Failure to use corollary discharge to remap visual target locations is associated with psychotic symptom severity in schizophrenia. *Journal of Neurophysiology*, *114*(2), 1129–1136. https://doi.org/10.1152/jn.00155.2015

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. *Schizophrenia Research*, *97*(1–3), 173–183. https://doi.org/10.1016/j.schres.2007.08.008

Samad, M., & Shams, L. (2016). Visual–somatotopic interactions in spatial perception. *NeuroReport*, *27*(3), 180–185. https://doi.org/10.1097/WNR.0000000000000521

Samad, M., & Shams, L. (2018). Recalibrating the body: Visuotactile ventriloquism aftereffect. *PeerJ*, *6*, e4504. https://doi.org/10.7717/peerj.4504

Sanfratello, L., Aine, C., & Stephen, J. (2018). Neuroimaging investigations of dorsal stream processing and effects of stimulus synchrony in schizophrenia. *Psychiatry Research: Neuroimaging*, *278*, 56–64. https://doi.org/10.1016/j.pscychresns.2018.05.005

Sardari, S., Pourrahimi, A., Fathi, M., Talebi, H., & Mazhari, S. (2022). Auditory processing in schizophrenia: Behavioural evidence of abnormal spatial awareness. *Laterality*, *27*(1), 71–85. https://doi.org/10.1080/1357650X.2021.1955910

Sartorato, F., Przybylowski, L., & Sarko, D. K. (2017). Improving therapeutic outcomes in autism spectrum disorders: Enhancing social communication and sensory processing through the use of interactive robots. *Journal of Psychiatric Research*, *90*, 1–11. https://doi.org/10.1016/j.jpsychires.2017.02.004

Sass, L. A., & Parnas, J. (2003). Schizophrenia, consciousness, and the self. *Schizophrenia Bulletin*, *29*(3), 427–444. https://doi.org/10.1093/oxfordjournals.schbul.a007017

Sato, Y., Toyoizumi, T., & Aihara, K. (2007). Bayesian Inference Explains Perception of Unity and Ventriloquism Aftereffect: Identification of Common Sources of Audiovisual Stimuli. *Neural Computation*, *19*(12), 3335–3355. https://doi.org/10.1162/neco.2007.19.12.3335

Savla, G. N., Vella, L., Armstrong, C. C., Penn, D. L., & Twamley, E. W. (2013). Deficits in domains of social cognition in schizophrenia: A meta-analysis of the empirical evidence. *Schizophrenia Bulletin*, *39*(5), 979–992. https://doi.org/10.1093/schbul/sbs080

Schlier, B., Jaya, E. S., Moritz, S., & Lincoln, T. M. (2015). The Community Assessment of Psychic Experiences measures nine clusters of psychosis-like experiences: A validation of the German version of the CAPE. *Schizophrenia Research*, *169*(1–3), 274–279. https://doi.org/10.1016/j.schres.2015.10.034

Schmack, K., Gòmez-Carrillo de Castro, A., Rothkirch, M., Sekutowicz, M., Rössler, H., Haynes, J.-D., Heinz, A., Petrovic, P., & Sterzer, P. (2013). Delusions and the role of beliefs in perceptual inference. *The Journal of Neuroscience*, *33*(34), 13701–13712. https://doi.org/10.1523/JNEUROSCI.1778-13.2013

Schmack, K., Rothkirch, M., Priller, J., & Sterzer, P. (2017). Enhanced predictive signalling in schizophrenia. *Human Brain Mapping*, *38*(4), 1767–1779. https://doi.org/10.1002/hbm.23480

Schmack, K., Schnack, A., Priller, J., & Sterzer, P. (2015). Perceptual instability in schizophrenia: Probing predictive coding accounts of delusions with ambiguous stimuli. *Schizophrenia Research: Cognition*, *2*(2), 72–77. https://doi.org/10.1016/j.scog.2015.03.005

Schroeder, C. E., & Foxe, J. (2005). Multisensory contributions to low-level, 'unisensory' processing. *Current Opinion in Neurobiology*, *15*(4), 454–458. https://doi.org/10.1016/j.conb.2005.06.008

Schroger, E., & Widmann, A. (1998). Speeded responses to audiovisual signal changes result from bimodal integration. *Psychophysiology*, *35*(6), 755–759. https://doi.org/10.1111/1469-8986.3560755

Senkowski, D., & Moran, J. K. (2022). Early evoked brain activity underlies auditory and audiovisual speech recognition deficits in schizophrenia. *NeuroImage: Clinical*, *33*, 102909. https://doi.org/10.1016/j.nicl.2021.102909

Seo, E., Park, H. Y., Park, K., Koo, S. J., Lee, S. Y., Min, J. E., Lee, E., & An, S. K. (2020). Impaired facial emotion recognition in individuals at ultra-high risk for psychosis and associations with schizotypy and paranoia level. *Frontiers in Psychiatry*, *11*, 577. https://doi.org/10.3389/fpsyt.2020.00577

Sestito, M., Umiltà, M. A., De Paola, G., Fortunati, R., Raballo, A., Leuci, E., Maffei, S., Tonna, M., Amore, M., Maggini, C., & Gallese, V. (2013). Facial reactions in response to dynamic emotional stimuli in different modalities in patients suffering from schizophrenia: A behavioral and EMG study. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00368

Seubert, J. (2010). Multisensory integration of emotionally valenced olfactory–visual information in patients with schizophrenia and healthy controls. *Journal of Psychiatry & Neuroscience*, *35*(3), 185–194. https://doi.org/10.1503/jpn.090094

Shams, L., & Beierholm, U. R. (2010). Causal inference in perception. *Trends in Cognitive Sciences*, *14*(9), 425–432. https://doi.org/10.1016/j.tics.2010.07.001

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*(1), 147–152. https://doi.org/10.1016/S0926-6410(02)00069-1

Siddi, S., Ochoa, S., Laroi, F., Cella, M., Raballo, A., Saldivia, S., Quijada, Y., Laloyaux, J., Rocha, N. B., Lincoln, T. M., Schlier, B., Ntouros, E., Bozikas, V. P., Gawęda, Ł., Machado, S., Nardi, A. E., Rodante, D., Deshpande, S. N., Haro, J. M., & Preti, A. (2019). A cross-national investigation of hallucination-like experiences in 10

countries: The E-CLECTIC study. *Schizophrenia Bulletin*, *45*(Supplement_1), S43–S55. https://doi.org/10.1093/schbul/sby156

Simon, D. M., Nidiffer, A. R., & Wallace, M. T. (2018). Single trial plasticity in evidence accumulation underlies rapid recalibration to asynchronous audiovisual speech. *Scientific Reports*, *8*(1), 12499. https://doi.org/10.1038/s41598-018-30414-9

Simpson, C., Pinkham, A. E., Kelsven, S., & Sasson, N. J. (2013). Emotion recognition abilities across stimulus modalities in schizophrenia and the role of visual attention. *Schizophrenia Research*, *151*(1–3), 102–106. https://doi.org/10.1016/j.schres.2013.09.026

Sinha, P., Kjelgaard, M. M., Gandhi, T. K., Tsourides, K., Cardinaux, A. L., Pantazis, D., Diamond, S. P., & Held, R. M. (2014). Autism as a disorder of prediction. *Proceedings of the National Academy of Sciences*, *111*(42), 15220–15225. https://doi.org/10.1073/pnas.1416797111

Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule: Spatial coincidence and multisensory integration. *Annals of the New York Academy of Sciences*, *1296*(1), 31–49. https://doi.org/10.1111/nyas.12121

Stefanis, N. C., Hanssen, M., Smirnis, N. K., Avramopoulos, D. A., Evdokimidis, I. K., Stefanis, C. N., Verdoux, H., & Van Os, J. (2002). Evidence that three dimensions of psychosis have a distribution in the general population. *Psychological Medicine*, *32*(2), 347–358. https://doi.org/10.1017/S0033291701005141

Stein, B. E., Burr, D., Constantinidis, C., Laurienti, P. J., Meredith, M. A., Perrault, T. J., Ramachandran, R., Röder, B., Rowland, B. A., Sathian, K., Schroeder, C. E., Shams, L., Stanford, T. R., Wallace, M. T., Yu, L., & Lewkowicz, D. J. (2010). Semantic confusion regarding the development of multisensory integration: A practical solution: Seeking consistency in multisensory terminology. *European Journal of Neuroscience*, *31*(10), 1713–1720. https://doi.org/10.1111/j.1460-9568.2010.07206.x

Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. MIT Press.

Stein, B. E., Meredith, M. A., Huneycutt, W. S., & McDade, L. (1989). Behavioral indices of multisensory integration: Orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience*, *1*(1), 12–24. https://doi.org/10.1162/jocn.1989.1.1.12

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.

*Schizophrenia Research*, *147*(2–3), 253–261. https://doi.org/10.1016/j.schres.2013.04.038

Stephen, J. M., Coffman, B. A., Jung, R. E., Bustillo, J. R., Aine, C. J., & Calhoun, V. D. (2013). Using joint ICA to link function and structure using MEG and DTI in schizophrenia. *NeuroImage*, *83*, 418–430. https://doi.org/10.1016/j.neuroimage.2013.06.038

Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., & Corlett, P. R. (2018). The predictive coding account of psychosis. *Biological Psychiatry*, *84*(9), 634–643. https://doi.org/10.1016/j.biopsych.2018.05.015

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. *Schizophrenia Research*, *179*, 97–103. https://doi.org/10.1016/j.schres.2016.09.035

Stevenson, R. A., Toulmin, J. K., Youm, A., Besney, R. M. A., Schulz, S. E., Barense, M. D., & Ferber, S. (2017). Increases in the autistic trait of attention to detail are associated with decreased multisensory temporal adaptation. *Scientific Reports*, *7*(1), 14354. https://doi.org/10.1038/s41598-017-14632-1

Stevenson, R. A., Zemtsov, R. K., & Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1517–1529. https://doi.org/10.1037/a0027339

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. *Frontiers in Human Neuroscience*, *8*, 788. https://doi.org/10.3389/fnhum.2014.00788

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

Straube, B., Green, A., Sass, K., & Kircher, T. (2014). Superior temporal sulcus disconnectivity during processing of metaphoric gestures in schizophrenia. *Schizophrenia Bulletin*, *40*(4), 936–944. https://doi.org/10.1093/schbul/sbt110

Strauss, G. P., Keller, W. R., Koenig, J. I., Sullivan, S. K., Gold, J. M., & Buchanan, R. W. (2015). Endogenous oxytocin levels are associated with the perception of emotion in dynamic body expressions in schizophrenia. *Schizophrenia Research*, *162*(1–3), 52–56. https://doi.org/10.1016/j.schres.2015.01.022

Stuke, H., Weilnhammer, V. A., Sterzer, P., & Schmack, K. (2018). Delusion proneness is linked to a reduced usage of prior beliefs in perceptual decisions. *Schizophrenia Bulletin*, *45*(1). https://doi.org/10.1093/schbul/sbx189

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 Research: Neuroimaging*, *106*(1), 1–14. https://doi.org/10.1016/S0925-4927(00)00081-0

Szycik, 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. *Schizophrenia Research*, *110*(1–3), 111–118. https://doi.org/10.1016/j.schres.2009.03.003

Szycik, 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., & Woldorff, M. G. (2009). Intermodal attention affects the processing of the temporal alignment of audiovisual stimuli. *Experimental Brain Research*, *198*(2–3), 313–328. https://doi.org/10.1007/s00221-009-1858-6

Tatz, J. R., Undorf, M., & Peynircioğlu, Z. F. (2021). Effect of impoverished information on multisensory integration in judgments of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(3), 481–497. https://doi.org/10.1037/xlm0000953

Tenckhoff, A., Tost, H., & Braus, D. F. (2002). Veränderte Wahrnehmung zeitlicher Relationen bei schizophrenen Psychosen. *Der Nervenarzt*, *73*(5), 428–433. https://doi.org/10.1007/s00115-001-1254-3

Teufel, C., Subramaniam, N., Dobler, V., Perez, J., Finnemann, J., Mehta, P. R., Goodyer, I. M., & Fletcher, P. C. (2015). Shift toward prior knowledge confers a perceptual advantage in early psychosis and psychosis-prone healthy individuals. *Proceedings of the National Academy of Sciences*, *112*(43), 13401–13406. https://doi.org/10.1073/pnas.1503916112

Thakkar, K. N., Diwadkar, V. A., & Rolfs, M. (2017). Oculomotor prediction: A window into the psychotic mind. *Trends in Cognitive Sciences*, *21*(5), 344–356. https://doi.org/10.1016/j.tics.2017.02.001

Thakkar, K. N., Nichols, H. S., McIntosh, L. G., & Park, S. (2011). Disturbances in body ownership in schizophrenia: Evidence from the rubber hand illusion and case study of a spontaneous out-of-body experience. *PLoS ONE*, *6*(10), e27089. https://doi.org/10.1371/journal.pone.0027089

Thakkar, K. N., & Rolfs, M. (2019). Disrupted corollary discharge in schizophrenia: Evidence from the oculomotor system. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *4*(9), 773–781. https://doi.org/10.1016/j.bpsc.2019.03.009

Thaler, N. S., Strauss, G. P., Sutton, G. P., Vertinski, M., Ringdahl, E. N., Snyder, J. S., & Allen, D. N. (2013). Emotion perception abnormalities across sensory modalities in bipolar disorder with psychotic features and schizophrenia. *Schizophrenia Research*, *147*(2–3), 287–292. https://doi.org/10.1016/j.schres.2013.04.001

Thoenes, S., & Oberfeld, D. (2017). Meta-analysis of time perception and temporal processing in schizophrenia: Differential effects on precision and accuracy. *Clinical Psychology Review*, *54*, 44–64. https://doi.org/10.1016/j.cpr.2017.03.007

Toh, W. L., Rossell, S. L., & Castle, D. J. (2011). Current visual scanpath research: A review of investigations into the psychotic, anxiety, and mood disorders. *Comprehensive Psychiatry*, *52*(6), 567–579. https://doi.org/10.1016/j.comppsych.2010.12.005

Tombaugh, T. (2004). Trail Making Test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology*, *19*(2), 203–214. https://doi.org/10.1016/S0887-6177(03)00039-8

Torregrossa, L. J., & Park, S. (2022). Body ownership across schizotypy dimensions: A rubber hand illusion experiment. *Psychiatry Research Communications*, *2*(3), 100058. https://doi.org/10.1016/j.psycom.2022.100058

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In *Cognitive theory* (3. Aufl., S. 199–239). Erlbaum.

Tripoli, G., Quattrone, D., Ferraro, L., Gayer-Anderson, C., La Cascia, C., La Barbera, D., Sartorio, C., Seminerio, F., Rodriguez, V., Tarricone, I., Berardi, D., Jamain, S., Arango, C., Tortelli, A., Llorca, P.-M., de Haan, L., Velthorst, E., Bobes, J., Bernardo, M., … Murray, G. K. (2022). Facial emotion recognition in psychosis and associations with polygenic risk for schizophrenia: Findings from the multi-center EU-GEI case–

control study. *Schizophrenia Bulletin*, *54*(1), 1104–1114. https://doi.org/10.1093/schbul/sbac022

Tschacher, W., Giersch, A., & Friston, K. (2017). Embodiment and schizophrenia: A review of implications and applications. *Schizophrenia Bulletin*, *43*(4), 745–753. https://doi.org/10.1093/schbul/sbw220

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. *Neuroscience & Biobehavioral Reviews*, *55*, 444–452. https://doi.org/10.1016/j.neubiorev.2015.04.019

Tseng, H.-H., Chen, S.-H., Liu, C.-M., Howes, O., Huang, Y.-L., Hsieh, M. H., Liu, C.-C., Shan, J.-C., Lin, Y.-T., & Hwu, H.-G. (2013). Facial and prosodic emotion recognition deficits associate with specific clusters of psychotic symptoms in schizophrenia. *PLoS ONE*, *8*(6), e66571. https://doi.org/10.1371/journal.pone.0066571

Tseng, H.-H., Roiser, J. P., Modinos, G., Falkenberg, I., Samson, C., McGuire, P., & Allen, P. (2016). Corticolimbic dysfunction during facial and prosodic emotional recognition in first-episode psychosis patients and individuals at ultra-high risk. *NeuroImage: Clinical*, *12*, 645–654. https://doi.org/10.1016/j.nicl.2016.09.006

Turetsky, B. I., Kohler, C. G., Indersmitten, T., Bhati, M. T., Charbonnier, D., & Gur, R. C. (2007). Facial emotion recognition in schizophrenia: When and why does it go awry? *Schizophrenia Research*, *94*(1–3), 253–263. https://doi.org/10.1016/j.schres.2007.05.001

Turi, M., Karaminis, T., Pellicano, E., & Burr, D. (2016). No rapid audiovisual recalibration in adults on the autism spectrum. *Scientific Reports*, *6*(1), 21756. https://doi.org/10.1038/srep21756

Ueda, N., Maruo, K., & Sumiyoshi, T. (2018). Positive symptoms and time perception in schizophrenia: A meta-analysis. *Schizophrenia Research: Cognition*, *13*, 3–6. https://doi.org/10.1016/j.scog.2018.07.002

Uhlhaas, P. J., & Mishara, A. L. (2006). Perceptual anomalies in schizophrenia: Integrating phenomenology and cognitive neuroscience. *Schizophrenia Bulletin*, *33*(1), 142–156. https://doi.org/10.1093/schbul/sbl047

Underwood, R., Peters, E., & Kumari, V. (2015). Psychobiology of threat appraisal in the context of psychotic experiences: A selective review. *European Psychiatry*, *30*(7), 817–829. https://doi.org/10.1016/j.eurpsy.2015.07.001

Uno, K., & Yokosawa, K. (2022). Cross-modal correspondence between auditory pitch and visual elevation modulates audiovisual temporal recalibration. *Scientific Reports*, *12*(1), 21308. https://doi.org/10.1038/s41598-022-25614-3

Van den Stock, J., de Jong, S. J., Hodiamont, P. P. G., & de Gelder, B. (2011). Perceiving emotions from bodily expressions and multisensory integration of emotion cues in schizophrenia. *Social Neuroscience*, *6*(5–6), 537–547. https://doi.org/10.1080/17470919.2011.568790

Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid recalibration to audiovisual asynchrony. *The Journal of Neuroscience*, *33*(37), 14633–14637. https://doi.org/10.1523/JNEUROSCI.1182-13.2013

Van der Burg, E., Alais, D., & Cass, J. (2015). Audiovisual temporal recalibration occurs independently at two different time scales. *Scientific Reports*, *5*(1), 14526. https://doi.org/10.1038/srep14526

van Dijke, A., van 't Wout, M., Ford, J. D., & Aleman, A. (2016). Deficits in degraded facial affect labeling in schizophrenia and borderline personality disorder. *PLOS ONE*, *11*(6), e0154145. https://doi.org/10.1371/journal.pone.0154145

van Donkersgoed, R. J. M., Wunderink, L., Nieboer, R., Aleman, A., & Pijnenborg, G. H. M. (2015). Social cognition in individuals at ultra-high risk for psychosis: A meta-analysis. *PLOS ONE*, *10*(10), e0141075. https://doi.org/10.1371/journal.pone.0141075

Vanes, L. D., White, T. P., Wigton, R. L., Joyce, D., Collier, T., & Shergill, S. S. (2016). Reduced susceptibility to the sound-induced flash fusion illusion in schizophrenia. *Psychiatry Research*, *245*, 58–65. https://doi.org/10.1016/j.psychres.2016.08.016

Verdoux, H., & van Os, J. (2002). Psychotic symptoms in non-clinical populations and the continuum of psychosis. *Schizophrenia Research*, *54*(1–2), 59–65. https://doi.org/10.1016/S0920-9964(01)00352-8

Verdoux, H., Van Os, J., Maurice-Tison, S., Gay, B., Salamon, R., & Bourgeois, M. L. (1999). Increased occurrence of depression in psychosis-prone subjects: A follow-up study in primary care settings. *Comprehensive Psychiatry*, *40*(6), 462–468. https://doi.org/10.1016/S0010-440X(99)90091-3

Vos, T., Abajobir, A. A., Abate, K. H., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abdulkader, R. S., Abdulle, A. M., Abebo, T. A., Abera, S. F., Aboyans, V., Abu-Raddad, L. J., Ackerman, I. N., Adamu, A. A., Adetokunboh, O., Afarideh, M., Afshin, A., Agarwal, S. K., Aggarwal, R., … Murray, C. J. L. (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328

diseases and injuries for 195 countries, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, *390*(10100), 1211–1259. https://doi.org/10.1016/S0140-6736(17)32154-2

Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: A tutorial review. *Attention, Perception, & Psychophysics*, *72*(4), 871–884. https://doi.org/10.3758/APP.72.4.871

Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cognitive Brain Research*, *22*(1), 32–35. https://doi.org/10.1016/j.cogbrainres.2004.07.003

Vyas, N. S., Patel, N. H., & Puri, B. K. (2011). Neurobiology and phenotypic expression in early onset schizophrenia. *Early Intervention in Psychiatry*, *5*(1), 3–14. https://doi.org/10.1111/j.1751-7893.2010.00253.x

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J., van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., … Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review*, *25*(1), 58–76. https://doi.org/10.3758/s13423-017-1323-7

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. *Annual Review of Psychology*, *71*(1), 193–219. https://doi.org/10.1146/annurev-psych-010419-051112

Weilnhammer, V., Röd, L., Eckert, A.-L., Stuke, H., Heinz, A., & Sterzer, P. (2020). Psychotic experiences in schizophrenia and sensitivity to sensory evidence. *Schizophrenia Bulletin*, *46*(4), 927–936. https://doi.org/10.1093/schbul/sbaa003

Welch, R. B., & Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, *88*(3), 638–667. https://doi.org/10.1037/0033-2909.88.3.638

Wenke, D., & Haggard, P. (2009). How voluntary actions modulate time perception. *Experimental Brain Research*, *196*(3), 311–318. https://doi.org/10.1007/s00221-009-1848-8

White, T. P., Wigton, R. L., Joyce, D. W., Bobin, T., Ferragamo, C., Wasim, N., Lisk, S., & Shergill, S. S. (2014). Eluding the illusion? Schizophrenia, dopamine and the McGurk effect. *Frontiers in Human Neuroscience*, *8*, 565. https://doi.org/10.3389/fnhum.2014.00565

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

Won, S., Lee, W. K., Kim, S.-W., Kim, J. J., Lee, B. J., Yu, J.-C., Lee, K. Y., Lee, S.-H., Kim, S.-H., Kang, S. H., Kim, E., & Chung, Y.-C. (2019). Distinct differences in emotional recognition according to severity of psychotic symptoms in early-stage schizophrenia. *Frontiers in Psychiatry*, *10*, 564. https://doi.org/10.3389/fpsyt.2019.00564

Wozny, D. R., & Shams, L. (2011). Recalibration of auditory space following milliseconds of cross-modal discrepancy. *The Journal of Neuroscience*, *31*(12), 4607–4612. https://doi.org/10.1523/JNEUROSCI.6079-10.2011

Wright, A. C., Fowler, D., & Greenwood, K. E. (2018). Developing a dynamic model of anomalous experiences and function in young people with or without psychosis: A cross-sectional and longitudinal study protocol. *BMJ Open*, *8*(11), e022546. https://doi.org/10.1136/bmjopen-2018-022546

Wroblewski, A., He, Y., & Straube, B. (2020). Dynamic Causal Modelling suggests impaired effective connectivity in patients with schizophrenia spectrum disorders during gesture-speech integration. *Schizophrenia Research*, *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 behavioural and event-related potential study. *Cognitive Neuropsychiatry*, *19*(4), 319–336. https://doi.org/10.1080/13546805.2013.866892

Zhou, H., Cai, X., 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. *Neuroscience & Biobehavioral Reviews*, *86*, 66–76. https://doi.org/10.1016/j.neubiorev.2017.12.013

Zhou, H., Cui, X., Yang, B., Shi, L., Luo, X., Cheung, E. F. C., Lui, S. S. Y., & Chan, R. C. K. (2022). Audiovisual temporal processing in children and adolescents with

schizophrenia and children and adolescents with autism: Evidence from simultaneity-judgment tasks and eye-tracking data. *Clinical Psychological Science*, *10*(3), 482–498. https://doi.org/10.1177/21677026211031543

Zhou, H., Lai, I. Y. S., Hung, K. S. Y., Chan, M. K. M., Ho, Z. T. Y., Lam, J. P. H., Lui, S. S. Y., & Chan, R. C. K. (2022). Audiovisual temporal processing in adult patients with first-episode schizophrenia and high-functioning autism. *Schizophrenia*, *8*(1), 75. https://doi.org/10.1038/s41537-022-00284-2

Zhou, H., Shi, L., Yang, H., Cheung, E. F. C., & Chan, R. C. K. (2020). Audiovisual temporal integration and rapid temporal recalibration in adolescents and adults: Age-related changes and its correlation with autistic traits. *Autism Research*, *13*(4), 615–626. https://doi.org/10.1002/aur.2249

Zhou, H., Wang, Y., Zhang, R., Cheung, E. F. C., Pantelis, C., & Chan, R. C. K. (2021). Neural correlates of audiovisual temporal binding window in individuals with schizotypal and autistic traits: Evidence from resting-state functional connectivity. *Autism Research*, *14*(4), 668–680. https://doi.org/10.1002/aur.2456

Zopf, R., Boulton, K., Langdon, R., & Rich, A. N. (2021). Perception of visual-tactile asynchrony, bodily perceptual aberrations, and bodily illusions in schizophrenia. *Schizophrenia Research*, *228*, 534–540. https://doi.org/10.1016/j.schres.2020.11.038

Zvyagintsev, M., Parisi, C., Chechko, N., Nikolaev, A. R., & Mathiak, K. (2013). Attention and multisensory integration of emotions in schizophrenia. *Frontiers in Human Neuroscience*, *7*, 674. https://doi.org/10.3389/fnhum.2013.00674

Zvyagintsev, M., Parisi, C., & Mathiak, K. (2017). Temporal processing deficit leads to impaired multisensory binding in schizophrenia. *Cognitive Neuropsychiatry*, *22*(5), 361–372. https://doi.org/10.1080/13546805.2017.1331160

## Contents Appendix

# Appendix A:
# General Supplementary Information

## 1. Glossary of Applied Instruments and Questionnaires

***Table A.1***

*Glossary of Applied Instruments and Questionnaires*

| Abbreviation | Measured Construct | Test Score | Possible (Range of) Values |
|---|---|---|---|
| BCSS | negative and positive beliefs about the self and others | sum score per subscale<br>· negative self,<br>· positive self,<br>· negative other,<br>· positive other | per subscale: 0-24 |
| CAPE | positive, negative, and depressive symptoms of psychosis | sum score per scale<br>· positive<br>· negative<br>· depressive | positive: 0-54<br>negative: 0-42<br>depressive: 0-24 |
| CAPS | anomalous perceptual experiences | number of experiences &<br>per experience:<br>· distress<br>· intrusiveness<br>· frequency | number: 0-32<br>distress, intrusiveness, & frequency: 0-160 |
| CPT-IP | vigilance | d prime | |
| EHI | handedness | | left, right, ambidexter |
| LSHS-E | hallucinations | total sum score | 0-64 |
| Passivity experiences | passivity experiences | total sum score | 0-21 |
| PCL | Paranoia | total sum score & sum for<br>· frequency<br>· conviction<br>· distress | total sum: 0-216<br>per subscale: 0-72 |
| PSYRATS | auditory verbal hallucinations & delusions | score per individual item; items measure various symptom characteristics | per item: 0-4 |
| SCID-V | diagnosis criteria of psychiatric disorders | | |
| TMT-A/-B | processing speed, attention, executive functioning | completion time (in seconds)<br>number of errors | |

*Notes.* BCSS = Brief Core Schema Scales; CAPE = Community Assessment of Psychic Experiences; CAPS = Cardiff Anomalous Perception Scale; CPT-IP = Continuous Performance Test, Identical Pairs; EHI = Edinburgh Handedness Inventory; LSHS-E; Launey-Slade Hallucination Scale, Extended; Passivity Experiences = Factor "Bizzare Experiences" from CAPE; PCL = Paranoia Checklist; PSYRATS = Psychotic Symptom Rating Scales; SCID-V Structural Clinical Interview for DSM-V Diagnoses; TMT-A/-B = Trail Making Test, Version A & B.

# Appendix B: Supplementary Results for Chapter II (Study 1)

# 1 Supplementary Analyses

## 1.1 Reaction Time Analysis

### *1.1.1 Attend Face*

A 2 (*Group*) x 3 (*Stimulus Condition*) mixed ANOVA for IE scores in attend face blocks revealed a significant main effect of *Stimulus Condition*, $F(2, 140) = 4.15$, $p = .018$, $\eta^2_p = 0.06$, indicating lower reaction times in the congruent ($M = 1567.89$, $SD = 276.6$) compared to the incongruent condition ($M = 1607.80$, $SD = 256.37$), $t(71) = -2.9$, $p = .005$, but no differences between the unimodal ($M = 1581.71$, $SD = 279.88$) and both the congruent, $t(71) = -1.04$, $p = .3$, and incongruent condition, $t(71) = -1.74$, $p = .86$. There was no main effect of *Group*, $F(1, 70) = 1.83$, $p = .18$, $\eta^2_p = 0.03$, and no *Group* * *Stimulus Condition* interaction, $F(2, 140) = 0.45$, $p = .636$, $\eta^2_p = 0.006$ (see Figure 3 second row for mean perceived emotional intensity per group and stimulus condition).

### *1.1.2 Attend Voice*

A 2 (*Group*) x 3 (*Stimulus Condition*) mixed ANOVA for IE scores in attend voice blocks revealed a significant main effect of *Stimulus Condition*, $F(2, 140) = 47.65$, $p < .001$, $\eta^2_p = .41$, indicating lower reactions times in the congruent ($M = 1622.50$, $SD = 276.99$) compared to the incongruent condition ($M = 1753.25$, $SD = 260.7$), $t(71) = -8.78$, $p < .001$, and lower reactions times in the unimodal ($M = 1633.26$, $SD = 246.7$) compared to the incongruent condition, $t(71) = -8.28$, $p < .001$, but no differences between the congruent and the unimodal condition, $t(71) = -0.69$, $p = .49$. There was no main effect of *Group*, $F(1, 70) = 0.52$, $p = .475$, $\eta^2_p < 0.007$, and no *Group* * *Stimulus Condition* interaction, $F(2, 140) = 1.95$, $p = .146$, $\eta^2_p = 0.03$ (see Figure 3 second row for mean perceived emotional intensity per group and stimulus condition).

## 1.2 Accuracy Analysis

### *1.2.1 Attend Face*

A 2 (*Group*) x 3 (*Stimulus Condition*) mixed ANOVA for accuracy attend face blocks was not interpretable due to a lack of homogeneity of covariances (Box test $p = .035$) and lack of homogeneity of error variances (Levene test $p < .05$). Therefore, we will report *t*-Tests for the factors *Group* and *Stimulus Condition*, as recommended by Hsu (1996). There was no difference in accuracy between the high proneness ($M = 84.09$, $SD = 6.56$) and the low proneness group ($M = 83.42$, $SD = 5.54$), $t(70) = -0.46$, $p = .645$. Accuracy scores were higher in the congruent ($M = 88.56$, $SD = 6.56$) compared to both the unimodal ($M = 84.00$, $SD = 6.45$), $t(71) = 5.28$, $p < .001$, and the incongruent condition ($M = 78.7$, $SD = 10.43$), $t(71) =$

8.05, $p < .001$, as well as higher in the unimodal compared to the incongruent condition, $t(71) = 4.71$, $p < .001$ (see Figure 3 third row for mean perceived emotional intensity per group and stimulus condition).

### *1.2.2 Attend Voice*

A 2 (*Group*) x 3 (*Stimulus Condition*) mixed ANOVA for accuracy in attend voice blocks was not interpretable due to a lack of homogeneity of covariances (Box test $p = .036$) and to a lack of homogeneity of error variances (Levene test $p < .05$). Therefore, we will report *t*-Tests for the factors *Group* and *Stimulus Condition*, as recommended by Hsu (1996). There was no difference in accuracy between the high proneness ($M = 81.16$, $SD = 7.22$) and the low proneness group ($M = 83.04$, $SD = 5.37$), $t(70) = 1.26$, $p = .212$. Accuracy scores were higher in the congruent ($M = 90.77$, $SD = 6.16$) compared to both the unimodal ($M = 82.74$, $SD = 6.15$), $t(71) = 11.58$, $p < .001$, and the incongruent condition ($M = 72.78$, $SD = 11.21$), $t(71) = 14.20$, $p < .001$, as well as higher in the unimodal compared to the incongruent condition, $t(71) = 9.08$, $p < .001$ (see Figure 3 third row for mean perceived emotional intensity per group and stimulus condition).

## 2 Supplementary Tables

***Table B.1***

*Pearson Correlation Coefficients r Between Inverse Efficiency Scores and Questionnaire Scores over all Participants.*

| | | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. | 9. | 10. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Attend Face | 1. IE uni | | | | | | | | | | |
| | 2. IE con | **.85***** | | | | | | | | | |
| | 3. IE inc | **.73***** | **.63***** | | | | | | | | |
| Attend Voice | 4. IE uni | **.82***** | **.71***** | **.61***** | | | | | | | |
| | 5. IE con | **.79***** | **.77***** | **.56***** | **.87***** | | | | | | |
| | 6. IE inc | **.62***** | **.48***** | **.57***** | **.64***** | **.54***** | | | | | |
| | 7. CAPS | -.16 | **-.33** | -.12 | -.05 | -.12 | -.01 | | | | |
| | 8. LSHS-E | -.11 | -.24 | -.1 | -.05 | -.14 | -.14 | **.7***** | | | |
| | 9. PCL [a] | .09 | -.03 | .02 | .03 | .03 | .17 | **.42***** | **.52***** | | |
| | 10.BCSS n.S. | -.05 | -.12 | -.14 | -.04 | -.02 | -.04 | .14 | .27 | **.63***** | |
| | 11.BCSS n.O. | .06 | .08 | .07 | .14 | .04 | .23 | .28 | **.40*** | **.52***** | .29 |

*Notes.* N = 72. Pearson correlation coefficients $r > .3$ are marked in bold. *p*-values are Bonferroni-corrected. IE uni = Inverse Efficiency in unimodal conditions; IE con = Inverse Efficiency in emotionally congruent conditions; IE inc = Inverse Efficiency in emotionally incongruent conditions; CAPS = Cardiff Anomalous Perception Scale, total score; LSHS-E = Launey-Slade Hallucination Scale-Extended, total score; PCL = Paranoia Checklist, total score; BCSS n.S. = Brief Core Schema Scales, negative self subscale; BCSS n.O. = Brief Core Schema Scales, negative other subscale.

[a] One subject in the low and one in the high proneness group had ≥50% missings in the PCL. Their PCL scores were corrected by the respective group mean.

* $p < .05$, ** $p < .01$, *** $p < .001$

***Table B.2***

*Inferential Statistics of Mixed ANOVAs per Target Emotion and Dependent Variable*

| | **RT** | | | **Accuracy** | | | **Intensity** | | |
|---|---|---|---|---|---|---|---|---|---|
| | $F$ (df) | $p$ | $\eta^2_p$ | $F$ (df) | $p$ | $\eta^2_p$ | $F$ (df) | $p$ | $\eta^2_p$ |
| **Happy** | | | | | | | | | |
| ***Face*** | | | | | | | | | |
| Group | 0.93 (1,70) | .338 | 0.01 | 2.06 (1,70) | .156 | 0.03 | 1.04 (1,70) | .312 | 0.02 |
| Cond. | 1.79 (2,140) | .171 | 0.03 | 2.75 (1.74, 121.74) | ***.075*** | 0.04 | 5.46 (2,140) | **.005** | 0.07 |
| G * C | 2.97 (2,140) | ***.055*** | 0.04 | 0.99 (1.74, 121.74) | .366 | 0.01 | 0.6 (2,140) | .549 | 0.009 |
| ***Voice*** | | | | | | | | | |
| Group | 0.51 (1,68) | .479 | 0.007 | 0.21 (1,70) | .645 | 0.003 | 0.15 (1,70) | .703 | 0.002 |
| Cond. | 42.31 (2,136) | **<.001** | 0.38 | 153.04 (1.63, 114.25) | **<.001** | 0.69 | 54.95 (2,140) | **<.001** | 0.44 |
| G * C | 3 (2,136) | ***.053*** | 0.04 | 2.54 (1.63, 114.25) | ***.094*** | 0.04 | 0.2 (2,140) | .819 | 0.003 |
| **Angry** | | | | | | | | | |
| ***Face*** | | | | | | | | | |
| Group | 1.7 (1,70) | .196 | 0.02 | 1.26 (1,70) | .265 | 0.02 | 0.64 (1,70) | .425 | 0.009 |
| Cond. | 2.30 (2,140) | .104 | 0.03 | 14.52 (1.58, 110.73) | **<.001** | 0.17 | 14.47 (1.53, 107.28) | **<.001** | 0.17 |
| G * C | 0.77 (2,140) | .464 | 0.01 | 0.95[a] (1.58, 110.73) | .372[a] | 0.01[a] | 0.002 (1.53, 107.28) | .994 | <0.001 |
| ***Voice*** | | | | | | | | | |
| Group | 0.09 (1,70) | .769 | 0.001 | [b] | [b] | [b] | 0.86 (1,70) | .356 | 0.01 |
| Cond. | 20.76 (2,140) | **<.001** | 0.23 | [b] | [b] | [b] | 39.32 (2,140) | **<.001** | 0.36 |
| G * C | 3.24 (2,140) | **.042** | 0.04 | [b] | [b] | [b] | 0.39 (2,140) | .68 | 0.005 |

| | RT | | | Accuracy | | | Intensity | | |
|---|---|---|---|---|---|---|---|---|---|
| | *F* (df) | *p* | $\eta^2_p$ | *F* (df) | *p* | $\eta^2_p$ | *F* (df) | *p* | $\eta^2_p$ |
| **Sad** | | | | | | | | | |
| ***Face*** | | | | | | | | | |
| Group | 1.70 (1,70) | .197 | 0.02 | [b] | [b] | [b] | 0.01 (1,70) | .93 | 0.001 |
| Cond. | 1.77 (2,140) | .175 | 0.03 | [b] | [b] | [b] | 21.03 (1.78, 124.37) | **<.001** | 0.23 |
| G * C | 1.12 (2,140) | .329 | 0.02 | [b] | [b] | [b] | 1.61 (1.78, 124.37) | .206 | 0.02 |
| ***Voice*** | | | | | | | | | |
| Group | 0.21 (1,70) | .652 | 0.003 | [b] | [b] | [b] | 0.14 (1,70) | .705 | 0.002 |
| Cond. | 17.16 (2,140) | **<.001** | 0.20 | [b] | [b] | [b] | 40.37 (1.83, 128.3) | **<.001** | 0.37 |
| G * C | 0.42[a] (2,140) | .959[a] | 0.001[a] | [b] | [b] | [b] | 0.15 (1.83, 128.3) | .845 | 0.002 |
| **Neutral** | | | | | | | | | |
| ***Face*** | | | | | | | | | |
| Group | 3.15 (1,68) | ***.081*** | 0.04 | 2.59 (1,70) | .112 | 0.04 | 6.65 (1,70) | **.012** | 0.09 |
| Cond. | 8.97 (2,140) | **<.001** | 0.12 | 8.64 (1.79, 125.2) | **<.001** | 0.11 | 4.48 (1.84, 128.78) | **.016** | 0.06 |
| G * C | 1.79 (2,140) | .17 | 0.03 | 0.59 (1.79, 125.2) | .539 | 0.008 | 0.65 (1.84, 128.78) | .512 | 0.009 |
| ***Voice*** | | | | | | | | | |
| Group | 3.5 (1,70) | ***.066*** | 0.05 | 0.07 (1,70) | .797 | 0.001 | [b] | [b] | [b] |
| Cond. | 11.84 (2,140) | **<.001** | 0.15 | 11.86 (2,140) | **<.001** | 0.15 | [b] | [b] | [b] |
| G * C | 5.70 (2,140) | **.004** | 0.08 | 2.83 (2,140) | ***.062*** | 0.04 | [b] | [b] | [b] |

*Notes.* $N = 72$. *p*-values < .05 are marked in bold, *p*-values <.1 in bold and italics. Cond. = factor *Stimulus Condition*; G * C = *Group* * *Stimulus Condition* interaction.

[a] interaction not interpretable due to a violation of equivalence of covariance matrices (Box-test $p < .05$); [b] mixed ANOVA not applicable due to a violation of error variance homogeneity (Levene-test $p < .05$), post-hoc comparisons for *Group* and *Stimulus Condition* are reported in the text (as recommended by Hsu, 1996).

# Appendix C: Supplementary Results for Chapter III (Study 2)

# 1 Supplementary Tables

***Table C.1***

*Spearman Rank Correlation Coefficients ρ for Associations Between Experimental Data and Questionnaire Scores in the Patient Group*

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. |
|---|---|---|---|---|---|---|---|---|
| 1. TBW | | | | | | | | |
| 2. Cumulative Recalibration | -.09 | | | | | | | |
| 3. Immediate Recalibration | .14 | -.14 | | | | | | |
| 4. CAPS | -.20 | .17 | **-.30** | | | | | |
| 5. LSHS-E | **-.41** | -.08 | **-.41** | **.66*** ** | | | | |
| 6. PCL | -.15 | -.25 | .02 | **.47** | .22 | | | |
| 7. Pass. Exp | -.09 | -.27 | **-.38** | **.32** | **.41** | **.44** | | |
| 8. BCSS n.S. | .01 | -.05 | -.11 | .24 | .28 | .03 | .12 | |
| 9. BCSS n.O. | -.11 | -.14 | .27 | .24 | .21 | **.57*** | .18 | **.50** |

*Notes.* $n = 25$. Spearman rank correlation coefficients $\rho > .3$ are marked in bold. $p$-values are Bonferroni-corrected. TBW = temporal binding window, defined as the bandwidth of the Gaussian fit over all test trials; Cumulative Recalibration = cumulative recalibration effect, defined as the difference in mean of the Gaussian fits, i.e. PSS, over the first 50 visual lead vs. auditory lead adapted test trials; Immediate Recalibration = immediate recalibration effect, defined as the difference in mean of the Gaussian fits, i.e. PSS, over test trials preceded by a visual first vs. auditory first test trial; CAPS = Cardiff Anomalous Perception Scale, total score; LSHS-E = Launey-Slade Hallucination Scale-Extended, total score; PCL = Paranoia Checklist, total score; Pass. Exp. = total score in passivity experiences items; BCSS n.S. = Brief Core Schema Scales, negative self subscale; BCSS n.O. = Brief Core Schema Scales, negative other subscale.

* $p < .05$, ** $p < .01$, *** $p < .001$

***Table C.2***

*Spearman Rank Correlation Coefficients ρ for Associations Between Experimental Data and Questionnaire Scores in the Healthy Control Group*

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. |
|---|---|---|---|---|---|---|---|---|
| 1. TBW | | | | | | | | |
| 2. Cumulative Recalibration | .15 | | | | | | | |
| 3. Immediate Recalibration | .11 | .15 | | | | | | |
| 4. CAPS | -.13 | .17 | .23 | | | | | |
| 5. LSHS-E | .00 | .28 | .05 | **.57*** | | | | |
| 6. PCL | .16 | **.34** | .26 | **.49** | **.50** | | | |
| 7. Pass. Exp | .06 | -.10 | .01 | .15 | .09 | .14 | | |
| 8. BCSS n.S. | **-.39** | .01 | .29 | **.34** | **.47** | .24 | -.10 | |
| 9. BCSS n.O. | **-.37** | -.12 | .03 | **.40** | **.42** | .07 | -.02 | **.63**** |

*Notes.* $n = 27$. Spearman rank correlation coefficients $\rho > .3$ are marked in bold. $p$-values are Bonferroni-corrected. TBW = temporal binding window, defined as the bandwidth of the Gaussian fit over all test trials; Cumulative Recalibration = cumulative recalibration effect, defined as the difference in mean of the Gaussian fits, i.e. PSS, over the first 50 visual lead vs. auditory lead adapted test trials; Immediate Recalibration = immediate recalibration effect, defined as the difference in mean of the Gaussian fits, i.e. PSS, over test trials preceded by a visual first vs. auditory first test trial; CAPS = Cardiff Anomalous Perception Scale, total score; LSHS-E = Launey-Slade Hallucination Scale-Extended, total score; PCL = Paranoia Checklist, total score; Pass. Exp. = total score in passivity experiences items; BCSS n.S. = Brief Core Schema Scales, negative self subscale; BCSS n.O. = Brief Core Schema Scales, negative other subscale.

* $p < .05$, ** $p < .01$, *** $p < .001$

***Table C.3***

*Spearman Rank Correlation Coefficients ρ for Associations Between Control Variables and Experimental Data over both Groups.*

| | TBW | Cumulative Recalibration | Immediate Recalibration |
|---|---|---|---|
| TMT-A time | .31 | -.02 | .12 |
| TMT-B time | .40* | -.16 | .03 |
| CPT-IP *d'* [a] | -.52*** | -.03 | .09 |

*Notes.* $n = 52$. $p$-values are Bonferroni-corrected. TBW = temporal binding window, defined as the bandwidth of a Gaussian fit over all test trials; Cumulative Recalibration = cumulative recalibration effect, defined as the difference in mean of the Gaussian fits, i.e. PPS, between the first 50 visual lead vs. auditory lead adapted test trials; Immediate Recalibration = immediate recalibration effect, defined as the difference in mean of the Gaussian fits, i.e. PSS, between test trials preceded by a visual first vs. auditory first test trial; TMT-A time = completion time in Trail-Making-Test, Version A; TMT-B time = completion time in Trail-Making-Test, Version B; CPT-IP *d'* = *d'* score in Continuous Performance Test, Identical Pairs.

[a] One patient did not complete the CPT-IP due to fatigue at the end of both sessions. Their CPT *d'* scores were replaced by the group mean to not lose the subject for the analysis.

* $p < .05$, ** $p < .01$, *** $p < .001$

# Appendix D: Supplementary Results for Chapter IV (Study 3)

# 1 Supplementary Tables

***Table D.1***

*Spearman Rank Correlation Coefficients ρ for Associations Between VE, Cumulative VAE, immediate VAE and Questionnaire Scores in the Patient Group*

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. | 9. |
|---|---|---|---|---|---|---|---|---|---|
| 1. VE | | | | | | | | | |
| 2. VAEc | **.52*** | | | | | | | | |
| 3. VAEi | .06 | .17 | | | | | | | |
| 4. CAPS | .14 | .07 | -.08 | | | | | | |
| 5. LSHS-E | .25 | .14 | -.12 | **.80***** | | | | | |
| 6. PCL | .30 | .18 | .07 | **.60*** | **.45** | | | | |
| 7. Pass. Exp. | .24 | -.06 | .14 | **.38** | **.37** | .23 | | | |
| 8. BCSS n.S. | -.20 | -.21 | -.06 | **.45** | **.35** | **.43** | .19 | | |
| 9. BCSS n.O. | .17 | .02 | -.34 | **.35** | .29 | **.62*** | .17 | **.58*** | |

*Notes.* n = 27. Spearman rank correlation coefficients $\rho > .3$ are marked in bold. *p*-values are Bonferroni-corrected. VE = ventriloquist effect size; VAEc = cumulative ventriloquist aftereffect size; VAEi = immediate ventriloquist aftereffect size; CAPS = Cardiff Anomalous Perception Scale, total score; LSHS-E = Launey-Slade Hallucination Scale-Extended, total score; PCL = Paranoia Checklist, total score; Pass. Exp. = total score in passivity experiences items; BCSS n.S. = Brief Core Schema Scales, negative self subscale; BCSS n.O. = Brief Core Schema Scales, negative other subscale.

* $p < .05$, ** $p < .01$, *** $p < .001$

***Table D.2***

*Spearman Rank Correlation Coefficients ρ for Associations Between VE, Cumulative VAE, immediate VAE and Questionnaire Scores in the Healthy Control Group*

| | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. | 9. |
|---|---|---|---|---|---|---|---|---|---|
| 1. VE | | | | | | | | | |
| 2. VAEc | **.36** | | | | | | | | |
| 3. VAEi | **.57**** | .12 | | | | | | | |
| 4. CAPS | .15 | **.38*** | .18 | | | | | | |
| 5. LSHS-E | .03 | .17 | .05 | **.45*** | | | | | |
| 6. PCL | -.05 | .10 | -.16 | .23 | .27 | | | | |
| 7. Pass. Exp. | -.08 | -.21 | -.10 | .16 | .22 | .10 | | | |
| 8. BCSS n.S. | .26 | .24 | .07 | .22 | .30 | .23 | .20 | | |
| 9. BCSS n.O. | .15 | .07 | -.15 | .06 | .27 | .23 | **.34** | .24 | |

*Notes.* n = 31. Spearman rank correlation coefficients $\rho > .3$ are marked in bold. *p*-values are Bonferroni-corrected. VE = ventriloquist effect size; VAEc = cumulative ventriloquist aftereffect size; VAEi = immediate ventriloquist aftereffect size; CAPS = Cardiff Anomalous Perception Scale, total score; LSHS-E = Launey-Slade Hallucination Scale-Extended, total score; PCL = Paranoia Checklist, total score; Pass. Exp. = total score in passivity experiences items; BCSS n.S. = Brief Core Schema Scales, negative self subscale; BCSS n.O. = Brief Core Schema Scales, negative other subscale.

* $p < .05$, ** $p < .01$, *** $p < .001$

***Table D.3***

*Spearman Rank Correlation Coefficients ρ for Associations Between Control Tests and Experimental Data over both Groups.*

| | VE | VAEc | VAEi |
|---|---|---|---|
| TMT-A time | *.11* | *.17* | *.25* |
| TMT-B time | *.10* | *-.02* | *.31* |
| CPT-IP *d'* | *-.23* | *-.08* | *-.3* |

*Notes.* n = 58. *p*-values are Bonferroni-corrected. VE = ventriloquist effect size; VAEc = cumulative ventriloquist aftereffect size; VAEi = immediate ventriloquist aftereffect size; TMT-A time = completion time in Trail-Making-Test, Version A; TMT-B time = completion time in Trail-Making-Test, Version B; CPT-IP *d' = d'* score in Continuous Performance Test, Identical Pairs.

* $p < .05$, ** $p < .01$, *** $p < .001$

**Acknowledgements**

Die Forschung i.R. dieses Dissertationsprojekt wurde aus bereichseigenen Mitteln der Arbeitsbereiche Biologische Psychologie & Neuropsychologie und Klinischer Psychologie & Psychotherapie gefördert.

Meine Arbeit an diesem Dissertationsprojekt ist durch die Unterstützung vieler Menschen ermöglicht worden, denen ich an dieser Stelle meine Dankbarkeit ausdrücken möchte.

Zuerst möchte ich meinen Betreuerinnen Brigitte Röder und Tania Lincoln danken, die mir die Möglichkeit geschaffen haben, an diesem Projekt zu arbeiten, und mich i.R. dieses Kooperationsprojekts leitend unterstützt haben. Ich danke den Mitgliedern meines Komitees Anja Riesel, Philipp Sterzer und Frank Rösler. Vielen Dank an Patrick Bruns, der im Laufe der Jahre immer ein offenes Ohr für mich hatte und mir bei inhaltlichen und methodischen Fragen sehr weitergeholfen hat.

Ich bin sehr dankbar für alle Kolleg*innen der AGs Biologischen Psychologie und Neuropsychologie und Klinischen Psychologie und Psychotherapie für all den inhaltlichen, motivationalen und sozialen Support. Ganz besonderer Dank gilt Liesa Stange und Daniela Schönberger für die wertvolle Freundschaft, die sich aus unserem Trio infernale entwickelt hat und mein Leben auch in Zukunft sehr bereichern wird. Danke an meine ILMA-Mitstreiter*innen Alexander Kramer, Carolin Heitmann, Cora Kubetschek, Laura Kuhne, Madita Linke, Rashi Pant, Sina Storm, Valentina Marcazzan und Veronika Zweckerl. Danke an Angelika Quade, Dagmar Tödter, Esther Wolfram und Nicola Kaczmarek für ihre Unterstützung bei zahlreichen kleinen und größeren Organisations- und Verwaltungsfragen. Vielen Dank an Ivo Denneberg, Julia Mönnicke und Theresa Vordermark, welche mich i.R. von Abschlussarbeiten unermüdlich unterstützt haben, sowie meine studentischen Hilfskräfte Erva Dilekoglu und Julia Wessels für ihre tatkräftige Hilfe bei der Datenerhebung.

Vielen lieben Dank an meine Familie und Freund*innen, ohne die dieses Projekt sehr viel schwerer zu bewältigen gewesen wäre. Danke an meine großartigen Eltern Silvia und Robert, meine Schwestern Lisa und Moni, deren Partner Matthias und Lukas und die kleine Bande Lea, Leon, Jari, Maja und Leni. Last but most certainly not least: Viel Liebe an all meine Freund*innen, durch die Hamburg mein Zuhause geworden ist. Ich kann euch hier nicht alle nennen - aber jede*r ist ausdrücklich gemeint!

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR PSYCHOLOGIE UND
BEWEGUNGSWISSENSCHAFT
Institut für Bewegungswissenschaft
Institut für Psychologie

**Erklärung gemäß *(bitte Zutreffendes ankreuzen)***

☐ **§ 4 (1c) der Promotionsordnung des Instituts für Bewegungswissenschaft der Universität Hamburg vom 18.08.2010**

☒ **§ 5 (4d) der Promotionsordnung des Instituts für Psychologie der Universität Hamburg vom 20.08.2003**

Hiermit erkläre ich,

Andreas Weiß (Vorname, Nachname),

dass ich mich an einer anderen Universität oder Fakultät noch keiner Doktorprüfung unterzogen oder mich um Zulassung zu einer Doktorprüfung bemüht habe.

Hamburg, 29.01.24
Ort, Datum

A. Weiß
Unterschrift

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR PSYCHOLOGIE UND
BEWEGUNGSWISSENSCHAFT
Institut für Bewegungswissenschaft
Institut für Psychologie

**Eidesstattliche Erklärung nach *(bitte Zutreffendes ankreuzen)***

☐ **§ 7 (4) der Promotionsordnung des Instituts für Bewegungswissenschaft der Universität Hamburg vom 18.08.2010**

☒ **§ 9 (1c und 1d) der Promotionsordnung des Instituts für Psychologie der Universität Hamburg vom 20.08.2003**

Hiermit erkläre ich an Eides statt,

1. dass die von mir vorgelegte Dissertation nicht Gegenstand eines anderen Prüfungsverfahrens gewesen oder in einem solchen Verfahren als ungenügend beurteilt worden ist.
2. dass ich die von mir vorgelegte Dissertation selbst verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und keine kommerzielle Promotionsberatung in Anspruch genommen habe. Die wörtlich oder inhaltlich übernommenen Stellen habe ich als solche kenntlich gemacht.

Hamburg, 29.01.24
Ort, Datum

A. Wey
Unterschrift