

Review

Not peer-reviewed version

---

# Cognitive Functioning and Symptomatic Correlations during Audiovisual Multisensory Integration in Schizophrenia: A Systematic Review and Meta-analysis

---

[Enze Tang](#) , Nuo Chen , Jian Li , [Hongwei Ding](#) <sup>\*</sup> , [Hui Chen](#) <sup>\*</sup>

Posted Date: 20 August 2024

doi: 10.20944/preprints202408.1256.v1

Keywords: audiovisual integration, schizophrenia, systematic review, meta-analysis, symptomatic correlation, multisensory



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Review*

# Cognitive Functioning and Symptomatic Correlations during Audiovisual Multisensory Integration in Schizophrenia: A Systematic Review and Meta-Analysis

Enze Tang <sup>1,2,3</sup>, Nuo Chen <sup>3,4</sup>, Jian Li <sup>2</sup>, Hongwei Ding <sup>\*,3,4</sup> and Hui Chen <sup>\*,1,2</sup>

<sup>1</sup> The First Affiliated Hospital, Zhejiang University School of Medicine, China; [angenze@zju.edu.cn](mailto:angenze@zju.edu.cn)

<sup>2</sup> Department of Psychology and Behavioral Sciences, Zhejiang University, China; [allen607@zju.edu.cn](mailto:allen607@zju.edu.cn)

<sup>3</sup> Speech-Language-Hearing Center, School of Foreign Languages, Shanghai Jiao Tong University, China; [chennuo\\_1224@sjtu.edu.cn](mailto:chennuo_1224@sjtu.edu.cn)

<sup>4</sup> National Research Centre for Language and Well-being, Shanghai, China

\* Correspondence: [hwding@sjtu.edu.cn](mailto:hwding@sjtu.edu.cn); [chenhui@zju.edu.cn](mailto:chenhui@zju.edu.cn)

**Abstract: Background:** Multisensory integration (MSI) avoids cognitive overload, and the traditional all-or-nothing perspective for dysfunctional MSI in schizophrenia has been increasingly challenged. This study aimed to systematically review and meta-analyze extant evidence on audiovisual MSI in schizophrenia. **Methods:** We searched Web of Science, PubMed, PsycINFO and PsycARTICLES databases from inception until April 16<sup>th</sup>, 2024. Studies using appropriate audiovisual integration tasks to compare either behavioral or neural responses between schizophrenia individuals and healthy controls were included. Primary outcome measures included between-group differences (Hedge's  $g$ ) and symptomatic correlations (Fisher's  $z$ ) in three components of audiovisual integration tasks, including audiovisual processing, adaptation and integration. The multilevel random-effects models were used to address statistical dependency. **Results:** In total, 209 between-group and 116 correlational effect sizes were synthesized from 46 studies, representing 1,108 non-overlapped schizophrenia participants and 1,084 healthy controls. The overall dysfunctional audiovisual MSI in schizophrenia ( $g = -0.50$ , [95% CI, -0.64 to -0.37]) was primarily ascribed to the impaired audiovisual processing ( $g = -0.53$ , [95% CI, -0.71 to -0.35]) and adaptation ( $g = -0.71$ , [95% CI, -0.99 to -0.42]), rather than integration itself ( $g = -0.23$ , [95% CI, -0.56 to 0.11]). Pooled correlational effect size was extremely small ( $z = 0.16$ , [95% CI, 0.03 to 0.30]). **Conclusions:** Individuals with schizophrenia show dysfunctional multisensory processing and adaptation, but preserve a less impaired multisensory integration ability. Our findings provide implications for clinical interventions, as the relatively preserved audiovisual integration ability may serve as a possible avenue for cognitive training programs.

**Keywords:** audiovisual integration; schizophrenia; systematic review; meta-analysis; symptomatic correlation; multisensory

## Introduction

Facing the overwhelming information flood in real-life situations, we normally have to appropriately unify the incoming inputs from segregated sensory events. Multisensory integration (MSI) thus prevents us from cognitive overload and scaffolds our cognitive responses to life events. Schizophrenia is characterized by dysfunctional MSI capability, such as the widened temporal binding window (TBW) [1], mitigated temporal acuity [2] and reduced illusory perception. [3] These changes in multisensory processing usually influence cognitive and social functioning of people with schizophrenia in real-life scenarios, [4] such as attention, [5,6] memory, [7] speech-related communication, [8,9] and emotion interpretation. [10–12] In contrast, schizophrenia individuals have also been found to preserve the ability to benefit from multisensory presentations in certain MSI tasks. [13,14] Considering that MSI is a complex concept encompassing several components beyond the

integration of information from different sensory modalities, [15] is it possible that individuals with schizophrenia show only partial abnormality in certain subordinate procedures during MSI processing?

Of various sensory modalities (e.g., visual, auditory, tactile and olfactory), the present study focused on audiovisual integration, one particular type of MSI, for a few reasons. Audiovisual integration is crucial for face-to-face communications and the exchange of emotional messages, and difficulties in audiovisual integration may lead to socio-communicative problems, especially for people with mental disorders, thus negatively influencing their quality of life. [16] Specifically for schizophrenia patients, perceptual dysfunction in visual and auditory modalities has been regarded as a hallmark symptom, [17,18] and the integration of these two modalities has been substantially investigated in this population, [1,19–21] which seemed to be correlated with schizophrenic symptoms like hallucination. [22] Therefore, since the widely attended audiovisual MSI could deepen our understanding of schizophrenia and potentially improve patients' quality of life, the present study has primarily synthesized MSI findings from the audiovisual domain.

Drawing on existing review articles concerning the MSI ability in populations with neuropsychiatric disorders, this ability could be investigated from three essential components. [23–25] First, the multisensory gain [26] denotes the benefits or performance enhancement obtained from multisensory stimuli compared to unisensory stimuli, which is also known as the redundant signals effect [27] or congruent facilitation effect. [28] A few paradigms, including simultaneity judgement [2] and temporal order judgement [29] have been commonly used to explore the multisensory gain. A second MSI component is simply the direct processing of multisensory inputs. For example, in the McGurk [30] and sound-induced flash illusion (SIFI) [3] paradigms, participants need to fuse information concurrently from different sensory modalities into one percept, which reflects the outcome of MSI processing. Third, the processing discrepancy between different multisensory conditions, as a reflection of the MSI adaptation ability, [31] has also been attended as part of the MSI capability. For example, in a typical emotion identification task, the identification of emotions in one modality can be influenced differently by different emotions in the other modality. [11] Recent meta-analyses suggested that neural correlates of audiovisual integration in the human brain are highly context-dependent, with both commonly and uniquely engaged brain regions as a function of analytical contrast, stimulus complexity and attention. [15,32] In support of the rationale to classify audiovisual MSI into different components, unique activation patterns of brain regions were identified if respectively analyzing the bimodal-unimodal contrasts (e.g., bilateral temporal cortices), audiovisual processing (e.g., right insula, right middle temporal gyrus) and bimodal-bimodal contrasts (e.g., right inferior frontal gyrus, right superior temporal gyrus, left occipital gyrus). [15] Motivated by such findings, it shall be necessary to analyze the behavioral and neurological changes of schizophrenia patients across different audiovisual MSI components, as unique patterns may be identified when distinct neural correlates are involved.

Considering distinct neural mechanisms underlying these subordinate components of audiovisual MSI, extant results on different MSI components would also be respectively aggregated in the current study (for more detailed illustrations, see eIntroduction in supplementary File S1; Supplementary Tables S1 and S2). By definition, the multisensory-unisensory (MU) contrast denotes the performance enhancement or processing benefit of audiovisual information, compared to unisensory stimuli. The audiovisual processing (AVP) component refers to the direct processing of audiovisual stimuli, without contrasting other unisensory or multisensory stimuli. The multisensory-multisensory (MM) contrast means the differences in processing different types of multisensory stimuli involved in particular MSI tasks, which reflects the MSI adaptation ability between various multisensory conditions.

Another interesting question is, to which degree the clinical symptoms of schizophrenia individuals could predict their audiovisual MSI performance? Most prior studies reported non-significant correlations between MSI performance and schizophrenia symptom severity, [5,29,33–37] but some still identified significant clinical relevance. [21,38–41] These inconsistencies encourage us to reconsider the clinical implications of such correlations. For example, are such correlations

indicative of the state or trait characteristics of schizophrenia? [42] Can neurological and behavioral changes during MSI tasks serve as potential endophenotypes of schizophrenia? [8,19]

To our knowledge, none to date has derived meta-analytical estimates for the subordinate components of audiovisual MSI processing or the clinical correlations between MSI and symptomatic measures in schizophrenia. Most extant review articles only qualitatively focused on specific cognitive domains [25,43,44] or neural connectivity [24] during MSI processing in schizophrenia, and very few have conducted meta-analyses to examine certain measurement metrics. [45,46] Therefore, this systematic review and meta-analysis had two primary objectives. First, we aimed to quantitatively synthesize the behavioral and neurological statistics about the differences between schizophrenia patients and healthy controls in audiovisual MSI tasks, and also examine the effects of various moderators through subgroup analyses and meta-regression. Second, we also aggregated correlational coefficients between symptom severity and audiovisual MSI performance in the schizophrenia population.

## Methods

### *Search Strategy*

The present systematic review and meta-analysis follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (PRISMA checklists: Supplementary Tables S3 and S4). [47] Four term blocks were designed and used in literature search, which respectively attended to the participant, modality, task and audiovisual integration aspects. We used designed retrieval queries (Supplementary Table S5) for literature search in the Web of Science, PubMed, PsycINFO and PsycARTICLES databases from inception to June 12th, 2023 initially, and updated the search on April 16th, 2024. With the combination of these retrieval queries, studies that investigated MSI capability in schizophrenia patients, especially across the auditory and visual modalities, shall be maximally and exclusively retrieved. To capture potentially unattended publications, forward and backward citation searches were conducted for six highly relevant review articles [4,24,25,43,44,48] and two meta-analytical studies. [45,46]

### *Inclusion and Exclusion Criteria*

In accordance to the PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) [42] PICO criteria, [49] the inclusion criteria of this systematic review and meta-analysis were formulated in four aspects, including population, intervention, control and outcome. First, studies had to include a schizophrenia clinical group and a comparable healthy control (HC) group (i.e., no history of any mental disorders and not at genetic risk for schizophrenia). Second, studies should use appropriate audiovisual integration tasks that could reflect the MSI capability, such as the McGurk, SIFI, simultaneity judgement (SJ) and temporal order judgement (TOJ) paradigms. Here, studies that observed the integration of other modalities (e.g., tactile and olfactory senses), [50] measured parent assessments or subjective feelings (e.g., Audio-Visual Abnormalities Questionnaire), [51] or required the participants to only respond to catch trials irrelevant to the MSI analyses [52–54] were excluded. Third, studies had to include at least one behavioral or neurological measure for the MSI tasks, and report necessary statistics (e.g.,  $t$ ,  $F$ , mean, SD,  $N$ , SE,  $r$ ) for calculating the between-group or correlational effect sizes. Last, studies were published in a peer-reviewed journal in English, and non-empirical publications (e.g., systematic reviews, editorials, commentaries) were excluded. We were also encouraged to conduct a supplementary fMRI meta-analysis. For this additional meta-analysis, we further removed studies that did not report the raw peak coordinates or inferential statistics (e.g.,  $t$  or  $z$ ) of clusters in standard stereotaxic coordinate space (e.g., Talairach or Montreal Neurological Institute, MNI).

In a multi-step screening procedure, two researchers (E.T. and N.C.) independently performed the literature search and screening. With all collected literature after duplication removal, they first inspected titles and abstracts according to the inclusion criteria, and removed the obviously inappropriate articles. Before the independent screening, both researchers agreed on the inclusion criteria and independently screened the titles and abstracts of 20 articles. At this step, disagreements



were addressed and consistent screening criteria were settled by discussion. Then, the two researchers independently screened all non-duplicated articles ( $k = 4,762$ ) and reached agreements on the articles that were not excluded by the title and abstract screening step through discussion. Afterwards, both researchers reviewed the full text of saved articles ( $k = 130$ ), and independently selected the eligible studies for our systematic review and meta-analysis. Disagreements and uncertainties were resolved by discussion with experienced professors (H.D. and H.C.).

### *Data Extraction*

Essential statistics for the calculation of between-group and correlational effect sizes were extracted, and demographic and methodological factors were coded for moderator analysis (see eMethods in supplementary File S1). Based on the descriptions and statistical meaning of these statistics in original articles, they were further labeled by three components of MSI (e.g., AVP, MM, MU). Two researchers (E.T. and N.C.) independently extracted raw statistics and coded moderators, and they reached an agreement on the extracted statistics by discussion. All extracted data were double screened for accuracy by two other researchers (J.L. and H.D.), and disagreements were resolved by discussion.

### *Quality Assessment*

To assess the quality and risk of bias of individual studies, we modified the Downs and Black's Checklist for Measuring Quality [55] and Newcastle-Ottawa Scale for Assessing the Quality of Nonrandomized Studies in Meta-Analysis. [56] The complete checklist for quality assessment used in this study is provided (see eMethods in supplementary File S1). Each included study was evaluated independently by two researchers (E.T. and N.C.). With a substantial inter-rater agreement (Cohen's kappa = 0.76), the remaining disagreements were addressed by discussion.

### *Statistical Analysis*

#### *Effect Size Aggregation*

Two primary outcomes were calculated and reported. First, we calculated the pooled effect size (Hedge's  $g$ ) that indicated between-group differences in MSI capability. A few subgroup analyses were conducted to decrease the between-study and within-study heterogeneity. We calculated the aggregated effect sizes of behavioral and neurological measures both together and separately, and further examined the pooled effect sizes of three MSI components, respectively. Second, we calculated the correlational effect sizes (Fisher's  $z$ ) that reflected the correlational relationships between MSI capability and schizophrenic symptoms.

Random-effects meta-analyses were performed using the *Metafor* [57] and *clubSandwich* [58] packages in R, version 4.3.1 (R Core Team). The Hedge's  $g$  was reported for the between-group analyses, which accounted for the bias caused by small sample sizes. According to the commonly acknowledged benchmarks, a  $g$  value of 0.2, 0.5 and 0.8 indicates a small, medium and large effect size, respectively. For all analyses, we combined the effect sizes from overlapping samples or the same sample into larger clusters by fitting the correlated and hierarchical effects (CHE) model [59] with robust variance estimation, [60,61] which addressed the unit-of-analysis error (i.e., multivariate and multilevel dependency). Detailed descriptions for our fitted models and principles for sample cluster specification are provided in eMethods in supplementary File S1. Between-study heterogeneity was primarily evaluated using the  $\tau^2$ ,  $I^2$ , and prediction intervals.

The supplem [entary fMRI meta-analysis was performed using the Effect-size Seed-based  $d$  Mapping software (ES-SDM) [62] which accounted for both increases and decreases of the outcomes with random-effects modeling. All raw  $z$ -scores were converted to  $t$ -values, and all coordinates were transformed to the MNI space. The default kernel size and thresholds of ES-SDM were maintained (full width at half maximum, FWHM = 20 mm, cluster extent = 10 voxels, voxel  $P = 0.005$ ). [63]

Publication bias was diagnosed by inspections of contour-enhanced funnel plots, [64] Egger's test [65] and the trim-and-fill analyses. [66] For the sensitivity evaluation, we conducted leave-one-out analyses by excluding one single unique effect size at a time to assess whether our findings were

strongly influenced by any single cases or outliers. In addition, the fail-safe  $N$  [67] was also performed to evaluate the robustness of our results.

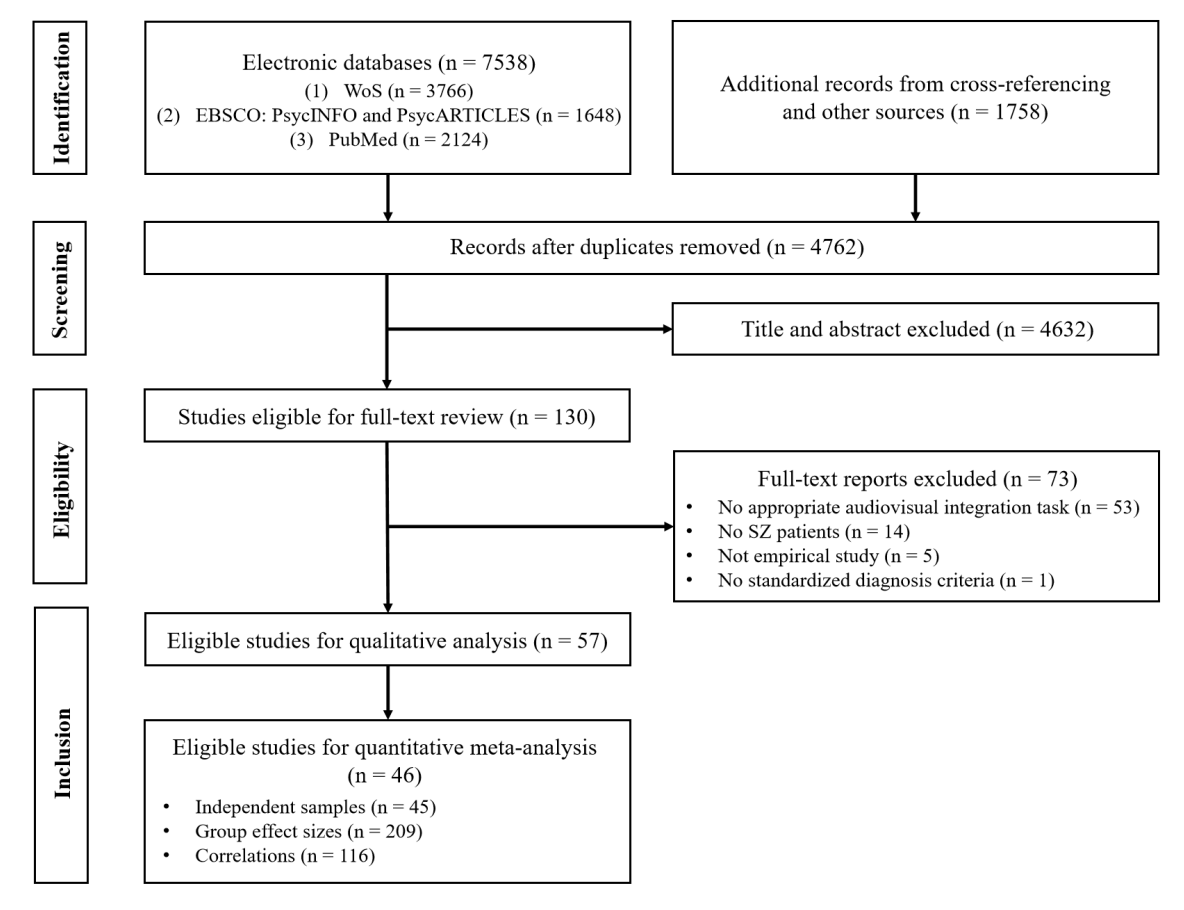
#### *Moderator Analyses*

Meta-regression models with random effects were constructed to demonstrate the association of participant-related and study-related moderators with the pooled group difference and correlation effects, respectively. Participant-related moderators encompassed demographic (e.g., age, sex distribution, education level and IQ) and clinical information, including illness duration, defined daily dose (DDD), and symptom measures (e.g., PANSS, SANS, SAPS). Study-related moderators included sample size, experimental tasks (e.g., McGurk, SIFI, SJ, emotion identification), audiovisual metrics (e.g., hit rates, illusion rates, TBW and neural responses), stimuli types (e.g., visual face/flash paired with auditory beep/syllable/sentence), and stimuli complexity (i.e., social and linguistic complexity).

## **Results**

#### *Literature Search*

Figure 1 presents the PRISMA flowchart, reflecting the identification, screening and selection of our literature search. In total, 4,762 non-repetitive articles were collected from the original electronic search and cross-referencing. After screening the titles and abstracts based on our inclusion and exclusion criteria, 130 articles were identified as eligible for full-text assessment, among which 57 articles were qualified for the qualitative review. Of the 73 excluded articles here, 53 studies did not adopt eligible audiovisual integration tasks to measure the MSI capability, 14 studies did not involve a schizophrenia patient group, five studies were not empirical experiments, and one study did not explicitly refer to any standardized diagnosis criteria. Afterwards, due to the lack of sufficient data for effect size calculation, 11 articles were further excluded from our quantitative analyses, thus leaving 46 eligible articles in our final sample for the meta-analytical review (the complete reference list is provided in supplementary File S2).



**Figure 1.** PRISMA flow diagram for literature search and study selection.

Published between 1998 and 2024, the 46 included articles contributed to a total of 209 between-group effect sizes and 116 correlation coefficients, which aggregated statistics from 2,192 unique individual participants (1,108 non-overlapped schizophrenia patients and 1,084 non-overlapped HCs). Most of them were originated from the United States ( $k = 14$ , 30.43%), followed by Germany ( $k = 13$ , 28.26%), China ( $k = 6$ , 13.04%), the United Kingdom ( $k = 4$ , 8.70%), the Netherlands ( $k = 3$ , 6.52%), France ( $k = 2$ , 4.35%), Belgium ( $k = 1$ , 2.17%), Greece ( $k = 1$ , 2.17%), Israel ( $k = 1$ , 2.17%), and Switzerland ( $k = 1$ , 2.17%). Sample sizes ranged from 6 to 53 for the patient group, from 7 to 57 for the HC group, and from 13 to 109 for the total number of participants. For all the patient participants, the mean age ranged from 14.84 to 48.43 years, the proportion of male participants ranged from 32.26% to 100%, and the education level ranged from 8.4 to 14.8 years. Specific descriptive characteristics for literature search and qualitative synthesis are demonstrated in Table 1 (group differences) and Supplementary Table S6 (correlational coefficients). Qualitative syntheses of existing evidence are demonstrated in eResults in supplementary File S1.

**Table 1.** Characteristics of Studies Included in Quantitative Analyses for Group Effects.

Author/Study	Male		Female		Experiment Methods	
	n	%	n	%	k(s)	stimuli
Lin et al. (1998)	10	100%	10	13%	1	h-beep
Lin et al. (2016)	10	10%	10	10%	1	h-beep
Gelder et al. (2002)	10	13%	10	17%	Gurk	2-syllable
Gelder et al. (2005)	10	12%	10	11%	1	sentence/prosody
Richer et al. (2007)	10	10%	10	17%	1	h-beep

ud et al. (2010)	0%	0	36%	0		rd-word
o et al. (2023)	38%	30	72%	0	✓	ture-sound
meirad et al. (2023a)	0%	36	33%	0	✓	ture-sound
meirad et al. (2023b)	0%	0	4%	0	WI	word
lon et al. (2016)	2%	9	35%	0		h-beep
3 et al. (2017)	7%	0	7%	0	I	h-beep
t al. (2021)	0%	3	0%	35		h-beep
et al. (2016)	2%	0	4%	0		burst
& Pan (2020)	0%	9	5%	38		syllable
ngelinckx et al. (2017)	3%	31	✓	11		word/prosody
rtin et al. (2013)	38%	0	38%	0	Gurk, SJ	syllable
yer et al. (2015)	38%	0	38%	35		nber-number
ller et al. (2014)	3%	0	0%	30		burst
rl et al. (2009)	7%	0	0%	0	Gurk	syllable
i Romero et al. (2016a)	9%	4	7%	0		syllable
i Romero et al. (2016b)	3%	7	3%	9	Gurk	syllable
s et al. (2007)	4%	0	1%	0	WI	word
fratello et al. (2018)	2%	0	4%	0		er-beep
s et al. (2014)	4%	0	1%	0		d-word
kowski & Moran (2022)	0%	2	4%	6	SI	syllable
						word
pson et al. (2013)	1%	1	0%	37		sentence/prosod
						y
ucny et al. (2013)	7%	0	2%	0		nber-noise
ne et al. (2014)	8%	0	8%	0		er-beep
ube et al. (2013)	0%	0	00%	0		/gesture-sentence
guladze et al. (2001)	3%	0	3%	0		word
guladze et al. (2002)	0%	0	5%	0		d-word
cik et al. (2013)	7%	0	7%	0		word
maian et al. (2024)	9%	9	5%	2		h-beep
hacher & Bergomi (2011)	1%	38	7%	4		h-beep
i den Stock et al. (2011)	0%	0	0%	0		ly action-human
						burst
ies et al. (2016)	0%	0	0%	0	I	h-beep
zel et al. (2016)	0%	2	0%	9		word/prosody
llace Cluster						
el et al. (2018)	0%	0	0%	0		syllable
venson et al. (2017)	0%	0	0%	0		h-beep; face-
						syllable
ite et al. (2014)	0%	3	0%	7	Gurk	syllable
od et al. (2006)	✓	3	✓	1		h-beep
et al. (2017)	5%	0	5%	4		phrase
nn et al. (2014)	0%	0	0%	0		er-beep
						h-beep;
u et al. (2021)	6%	34	7%	0	AD	syllable; face-
						sentence
u et al. (2022)	9%	3	0%	2		h-beep; face-
						syllable
ragintsev et al. (2013)	0%	0	0%	0	EI	h-beep; face-
						pseudowords

Note. ES = Effect sizes, SZ = Schizophrenia, HC = Healthy control, %male = Percentage of male participants, Mage = Mean age in years, TI = Target identification, SIFI = Sound induced flash



illusion, McGurk = McGurk effect, N/A = Not available, EI = Emotion identification, SJ = Simultaneity judgement, LD = Lexical decision, LTM = long-term memory, SN = Speech-in-noise, WI = Word identification, MJ = Match judgement, SI = Syllable identification, DJ = Distance judgement, CJ = Causality judgement, AD = Audiovisual asynchrony detection task. <sup>a</sup>The number of between-group (g) and correlational (z) effect sizes extracted from each study.

Primary Outcomes

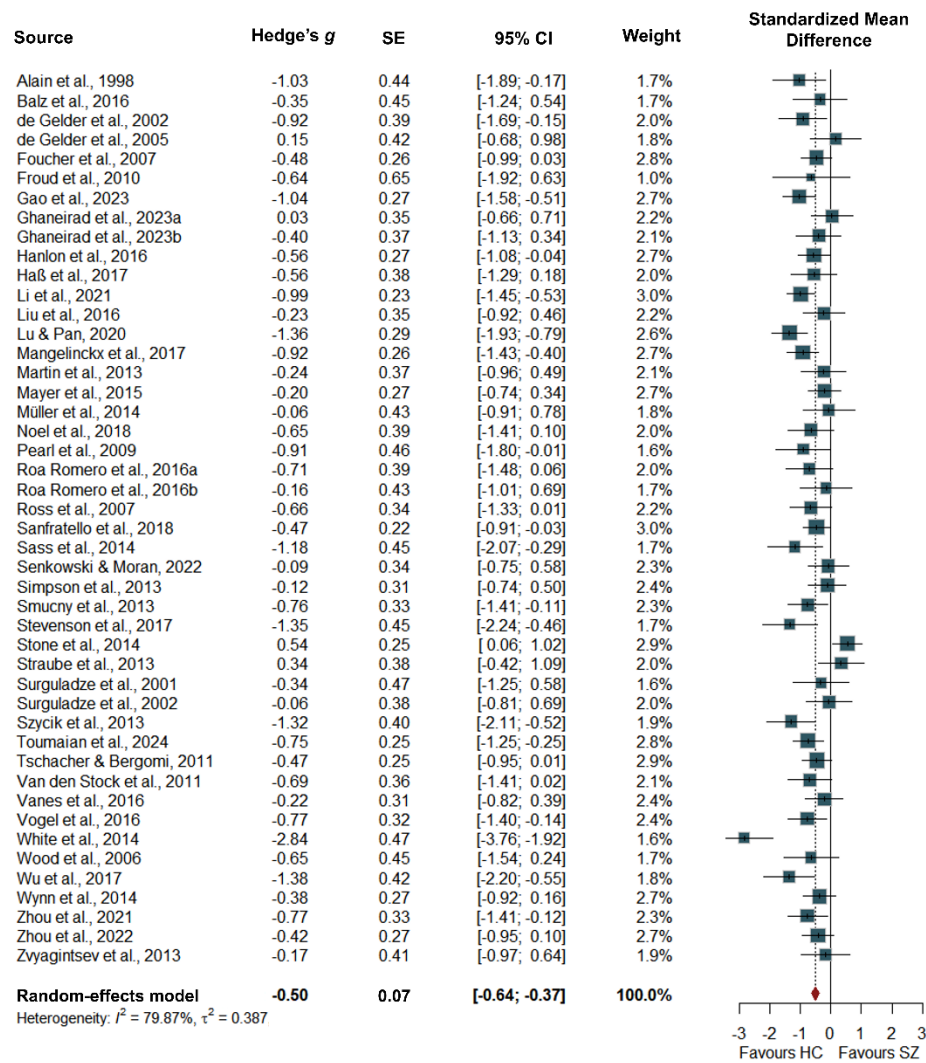
Group Differences

Primary outcomes and important subgroup analyses are displayed in Table 2. Aggregated statistics suggested that schizophrenia patients showed overall impairments at a moderate magnitude in audiovisual MSI tasks ( $g = -0.50$ ,  $SE = 0.07$ ; Figure 2). There existed high heterogeneity among these studies at both the between-cluster ( $\tau^2_{between-cluster} = 0.053$ ,  $I^2 = 10.87\%$ ) and within-cluster ( $\tau^2_{within-cluster} = 0.334$ ,  $I^2 = 69.00\%$ ) levels. However, the impairment of audiovisual MSI in schizophrenia seemed to be ascribed to the AVP ( $g = -0.53$ ,  $SE = 0.09$ ,  $p < .001$ ; Supplementary Figure 1) and MM ( $g = -0.71$ ,  $SE = 0.15$ ,  $p < .001$ ; Supplementary Figure 2) components, rather than the MU ( $g = -0.23$ ,  $SE = 0.17$ ,  $p = .225$ ; Supplementary Figure 3) component. Subgroup Analyses further suggested that both behavioral ( $g = -0.52$ ,  $SE = 0.07$ ,  $p < .001$ ; Supplementary Figure 4) and neural ( $g = -0.51$ ,  $SE = 0.12$ ,  $p < .001$ ; Supplementary Figure 5) measures of the group differences in audiovisual MSI obtained similar results.

Table 2. Primary Outcomes and Subgroup Analyses for Between-group and Clinical Correlational Effects.

Outcome	(95% CI)				%
Group difference					
Overall	0 (-0.64 to -0.37)	3 to 0.72	.01	.37	.37
Components					
MM	1 (-0.99 to -0.42)	6 to 0.45	.01	.24	.31
AVP	3 (-0.71 to -0.35)	3 to 0.67	.01	.39	.38
MU	3 (-0.56 to 0.11)	3 to 0.78	.5	.34	.4
Measures					
Behavioral	2 (-0.66 to -0.37)	5 to 0.62	.01	.30	.6
Neural	1 (-0.74 to -0.27)	7 to 0.86	.01	.39	.8
Correlation					
Overall	.5 (0.03 to 0.30)	8 to 0.70	.0	.71	.74
Measures					
Behavioral	.5 (-0.10 to 0.29)	7 to 0.76	.3	.37	.32
Neural	.5 (0.04 to 0.46)	5 to 0.75	.7	.33	.35

Abbreviations: PI, prediction interval; MSI, multisensory integration; MM, multisensory-multisensory contrast; AVP, audiovisual processing; MU, multisensory-unisensory contrast.

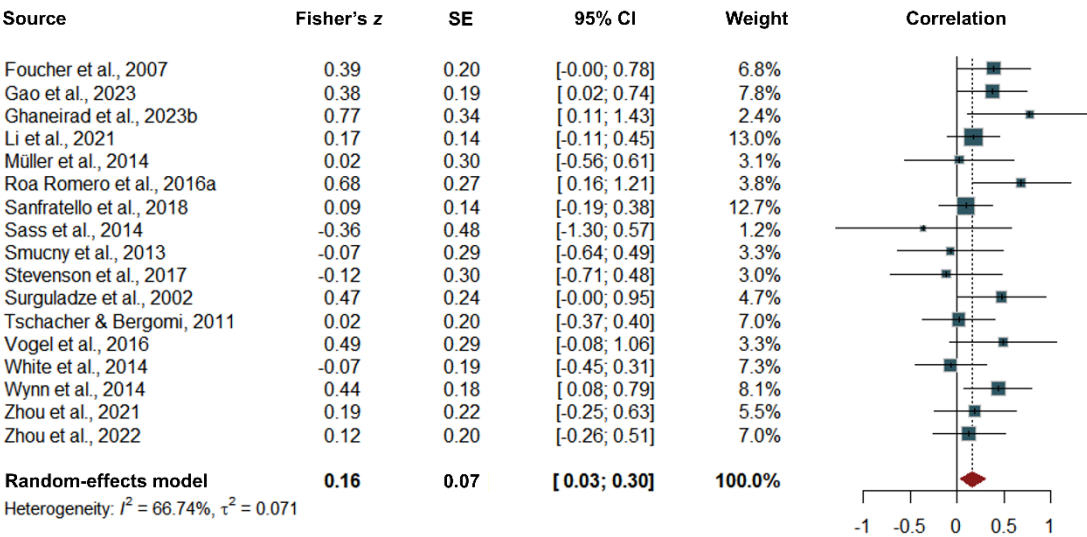


**Figure 2.** Forest Plot of Individual Effect Sizes of Group Differences in Audiovisual Multisensory Integration between Schizophrenia and Healthy Controls.

Our fMRI meta-analysis was conducted based on 35 available peak coordinates and *t*-values of clusters extracted from five studies. [19,31,68–70] There was no significant residual heterogeneity between studies ( $\tau = 0.00$ ,  $Q = 6.41$ ,  $p = .17$ ). The aggregated fMRI data reflected decreased activation in the left supramarginal gyrus (Brodmann area 48, cluster size = 161 voxels, MNI coordinates = [19,31,68–70],  $z = -2.07$ ,  $p_{uncorrected} = .02$ ,  $I^2 = 10.69\%$ ) among schizophrenia patients during audiovisual MSI tasks. However, the significance disappeared after FWER-correction, so this aggregated evidence should be considered weak.

Correlations

Existing evidence only suggested weak correlations between the symptomatic measures and audiovisual MSI performance in schizophrenia ( $z = 0.16$ ,  $SE = 0.07$ ,  $p = .030$ ; Figure 3). Moderate heterogeneity was identified at both the between-cluster ( $\tau^2_{between-cluster} = 0.030$ ,  $I^2 = 28.47\%$ ) and within-cluster ( $\tau^2_{within-cluster} = 0.041$ ,  $I^2 = 38.27\%$ ) levels. Neither the behavioral ( $z = 0.10$ ,  $SE = 0.10$ ,  $p = .348$ ; Supplementary Figure 6) nor neural ( $z = 0.25$ ,  $SE = 0.11$ ,  $p = .077$ ; Supplementary Figure 7) measures in MSI tasks were significantly associated with patients' symptoms.



**Figure 3.** Forest Plot of Individual Effect Sizes of Correlations Between Audiovisual Multisensory Integration and Symptom Measures in Schizophrenia.

*Moderator Analyses*

Detailed statistics of all meta-regression models are displayed in Supplementary Tables S7 (continuous moderators for group difference), S8 (categorical moderators for group difference) and S9 (all moderators for correlational coefficients). Significant moderators associated with the group difference and correlation effects are presented in Supplementary Table S10. Detailed descriptions for moderator analyses are illustrated in eResults in supplementary File S1.

For MSI in general, schizophrenia patients with higher scores of the Brief Psychiatric Rating Scale (BPRS) showed increased audiovisual MSI capability ( $b = 0.038$ ,  $p = .035$ ). In addition, various measuring metrics were predictive of different group effects ( $F(8, 200) = 2.358$ ,  $p = .019$ ), with priming effects (PE;  $g = 0.55$ , 95 % CI [0.11, 1.00]) eliciting the best MSI performance in schizophrenia and d'-value ( $g = -1.04$ , 95 % CI [-1.23, -0.86]) resulting in the largest between-group difference. Neither social ( $F(1, 207) = 0.066$ ,  $p = .798$ ) nor linguistic ( $F(1, 207) = 0.181$ ,  $p = .671$ ) complexity was predictive of group effect sizes. Of the three components, AVP was associated with clinical characteristics, while MM and MU were significantly associated with experimental stimuli. For AVP, higher BPRS scores ( $b = 0.033$ ,  $p = .013$ ) were associated with smaller between-group differences, and various measuring metrics ( $F(8, 132) = 3.396$ ,  $p = .001$ ) could also differentiate the discrepancy in MSI performance between patients and HCs. As for MM, higher proportion of male participation ( $b = 0.027$ ,  $p = .011$ ) was predictive of better MSI adaptation within the schizophrenia group. As for the measuring indices for MM, the neural response ( $g = -1.04$ , 95 % CI [-1.66, -0.41]) could sensitively detect the largest negative effect sizes. Last, the between-group differences in integration ability, as reflected by the MU component, were only predicted by different stimuli ( $F(7, 23) = 2.685$ ,  $p = .035$ ), with soccer-beep stimuli ( $g = 0.64$ , 95 % CI [-0.29, 1.56]) predicting the most integration benefits for schizophrenia patients and number-noise stimuli the least ( $g = -1.01$ , 95 % CI [-1.86, -0.15]).

Interestingly, the more socially ( $F(1, 114) = 6.470$ ,  $p = .012$ ) or linguistically ( $F(1, 114) = 13.131$ ,  $p = .0004$ ) complex the stimuli were, the more possible the correlations between severe symptoms and worse MSI capability could be identified in schizophrenia. Of different measures, eye movements ( $z = 0.63$ , 95 % CI [-0.16, 1.42]) predicted the highest positive correlations, while PE ( $z = -0.17$ , 95 % CI [-2.82, 2.48]) and illusion rates ( $z = -0.02$ , 95 % CI [-0.94, 0.89]) were predictive of negative correlations.

*Quality Assessment, Publication Bias and Sensitivity Analyses*

Quality assessment (QA) scores (see eResults in supplementary File S1; Supplementary Table S11) in our meta-analysis ranged from moderate (score = 8) to high (score = 17) for individual studies, with an average of 13.67 (SD = 2.15, median = 14). The QA scores were not significantly associated with either group ( $b = 0.02$ , SE = 0.03, 95% CI [-0.05, 0.09],  $p = .46$ ) or correlation ( $b = 0.07$ , SE = 0.03, 95% CI [-0.08, 0.22],  $p = .21$ ) effects.

By Egger's test, the funnel plots (Supplementary Figure 8) showed symmetrical distributions of both between-group ( $b = -0.814$ , 95% CI [-2.05, -0.42],  $p = .20$ ) and correlation ( $b = 0.182$ , 95% CI [-0.76, 1.13],  $p = .71$ ) effect sizes. The trim-and-fill analyses added 24 studies for the group effects, resulting in a corrected medium pooled effect size ( $k = 233$ ;  $g = -0.57$ , SE = 0.08, 95 % CI [-0.73, -0.42], prediction interval [-1.88, 0.73],  $p < .001$ ), and five studies were added for the correlation effects, remaining a corrected small pooled effect size ( $k = 121$ ;  $z = 0.18$ , SE = 0.07, 95 % CI [0.05, 0.31], prediction interval [-0.37, 0.73],  $p = .018$ ).

Sensitivity analyses by the leave-one-out method identified 54 outliers for the group effects (Supplementary Figure 9), and the pooled effect size without outliers remained moderately negative ( $g = -0.54$ , SE = 0.05, 95 % CI [-0.63, -0.44], prediction interval [-1.04, -0.04],  $p < .001$ ) with decreased overall heterogeneity ( $\tau^2 = 0.063$ ,  $I^2 = 39.78\%$ ). For the correlation effects, 16 outliers were eliminated (Supplementary Figure 10) and the pooled effect size remained small ( $z = 0.25$ , 95 % CI [0.15, 0.34], prediction interval [0.02, 0.47],  $p < .001$ ) with low heterogeneity ( $\tau^2 = 0.011$ ,  $I^2 = 24.81\%$ ). The fail-safe  $N$  analysis suggested an addition of 30,131 and 1,776 unpublished studies could respectively negate the two pooled effect sizes, which corroborated the robustness of current findings.

## Discussion

This systematic review and multilevel meta-analysis synthesized extant evidence on the audiovisual MSI capability of schizophrenia patients, which calculated pooled between-group effect sizes and aggregated the correlational effect sizes between schizophrenia symptoms and patients' MSI performance. For the first time, we analyzed patients' MSI capability respectively in three subordinate MSI components, and inspected the clinical correlations and significant moderators. Our findings challenged the convention that schizophrenia patients showed systematic impairments in MSI processing, and could address a few inconsistent previous findings by providing a new research perspective. We would like to demonstrate two noteworthy results here, and provide supplementary discussions in eDiscussion (supplementary File S1).

First, our findings acknowledged that the moderately impaired audiovisual MSI capability in schizophrenia was primarily ascribed to audiovisual processing (AVP) and adaptation (MM), rather than the integration (MU) itself. Our identified partially deficient audiovisual MSI capability (overall, AVP and MM) in schizophrenia supported previous review articles, which indicated the overall dysfunctional MSI in schizophrenia at the temporal, [45,46] (non-)emotional [25] and neurological [24] domains. Based on our meta-regression results (Supplementary Table S8), the typically used temporal tasks (e.g., SJ) and illusion-based tasks (e.g., McGurk) consistently revealed worse MSI performance of patients in the AVP and MM components. Surprisingly, our aggregated data identified only small between-group differences ( $g = -0.23$ ) concerning the benefit brought by the multisensory integration (MU) itself. This corroborated sporadic evidence claiming that schizophrenia patients showed comparable [14] or even greater [71,72] performance enhancement than HCs while processing audiovisual stimuli, compared to the unisensory information. As displayed in Supplementary Table S8, paradigms involving spatial distance judgement, lexical decision, long-term memory and speech-in-noise recognition distinguished the impaired AVP and MM components from the less impaired or even enhanced MU component. Pertaining to different measuring indices, the TBW and  $d'$ -difference detected the largest deficits of patients in the AVP and MM components, while the RT difference between audiovisual and unisensory conditions was found to best reflect the relatively preserved integration ability of patients. However, since only 31 MU effect sizes were pooled, our meta-regression findings are currently tentative and require further examinations.

In addition to the traditional meta-analysis evaluating the between-group effect sizes, our additional fMRI meta-analysis identified weak evidence for decreased activation in the left supramarginal gyrus in schizophrenia patients during audiovisual MSI tasks, which supplemented one recent systematic review reporting patients' abnormal EEG oscillations and amplitudes in temporal-parietal regions during MSI processing. [24] Empirical studies have also observed reduced activation in the left temporoparietal junction, [69] the general parietal areas (dorsal visual stream), and delayed peak in the intraparietal sulcus [73] in schizophrenia during audiovisual MSI tasks. However, since only five studies reported sufficient statistics for this meta-analysis and 88.57% (31 out of 35) of our extracted raw data reflected the AVP and MM components, our current fMRI meta-analysis only provided preliminary evidence and required further investigations. More qualitative discussions on the neurological findings are provided in eDiscussion (Supplementary file S1).

A few mechanisms have been proposed to interpret the between-group differences. First, impairments in audiovisual MSI performance could be attributed to the deficits in unisensory representations. [3] Since schizophrenia patients showed diminished decoding of the auditory speech signals [14,74] and abnormally disturbed visual perception, [17] the between-group differences in audiovisual MSI might possibly result from the combination of the weakness and instability of unisensory perception in schizophrenia. [2] Moreover, during audiovisual MSI tasks, the mutual influences between the two modalities may either facilitate or prohibit information processing, [12,75] and unisensory deficits can thus sway the MSI outcomes, which would even be amplified when participants were asked to selectively attend to one modality and inhibit distractions from the other modality. [6,76] However, recent research has reported that schizophrenia patients could still display insensitive audiovisual temporal acuity even if their unisensory temporal processing was normal, [77] so patients' altered audiovisual MSI could not be satisfactorily explained by unisensory dysfunction alone. [21] To determine the mechanisms underlying MSI changes in schizophrenia, the Bayesian modelling method has shown its merits in a few atypical populations (e.g., children, older adults and autism populations). [78] Since MSI Bayesian models involve the computational process itself (i.e., the integration process) and the weighted unisensory estimates (i.e., unisensory processing), such models may supplement previous experimental evidence and help reveal the specific procedures resulting in the changes of MSI performance in schizophrenia. In addition to these interpretations, a few studies have also ascribed MSI changes in schizophrenia to impaired top-down modulation [79] or inaccurate cognitive event structures (e.g., temporal, spatial), [30] which might be independent of sensory binding operations. Although an investigation of the mechanisms underlying audiovisual MSI changes in schizophrenia is beyond the current scope, we encourage more sophisticated investigations on the pathological causes for MSI changes in schizophrenia beyond the phenomenal revelation.

Our second important finding was the weak correlation between schizophrenic symptoms and MSI performance, and the predictability of higher social or linguistic complexity for stronger correlations. During data extraction, most included studies did not report the correlation coefficients because of non-significant results. [5,29,33–37] Our analyses synthesized 116 coefficients from a subset of 17 studies, and the small pooled effect size in our meta-analysis in turn corroborated the widely acknowledged weak or null correlations in individual studies. One reason for the absence of clinical correlations with MSI capability was the relatively small sample size in previous studies. [14] Another possibility could be the inclusion of patients with only moderate to low levels of symptoms in most studies. Of all included samples in our study, the score ranges for different symptom measures were similarly restricted (PANSSP: 8.21 to 17.4; PANSSN: 10.13 to 23.9; PANSSG: 18.3 to 38.29; SAPS: 5.1 to 37.9; SANS: 5.6 to 46; BPRS: 15.4 to 46.93). Hence, with the restricted range of symptoms, the performance of symptomatically severe patients has been rarely observed, thus excluding the possibility of obtaining significant overall correlations between patients' cognitive performance (e.g., audiovisual MSI) and clinical symptoms. Although our meta-regression analyses (Supplementary Table S9) suggested that the weak correlation was consistent across different symptom measures (e.g., BPRS, PANSS, SANS and SAPS) and symptom types (e.g., positive and



negative), these findings concerning correlational effects might not reflect the whole schizophrenia spectrum and required extra caution for interpretation.

Besides the two primary findings, our meta-regressive results could also inspire future investigations. Although the conclusiveness of our findings was challenged because of the limitations in original studies, such as the small number of particular types of effect size, the restricted symptom coverage, and insufficient report of raw data, a few feasible directions can still be enlightened. Interestingly, we identified that social and speech complexity of stimuli was predictive of different correlational strengths, but did not seem to influence the overall group differences or any subordinate components. In specific, stimuli with complex social (e.g., face, body, gesture) or linguistic (e.g., word, phrase, sentence) characteristics may help detect severe symptoms when patients performed poorly in audiovisual MSI tasks. One possible explanation was that MSI processing of complex stimuli imposed extra difficulty by introducing higher cognitive demands and flexibility engagement, so they could detect the abnormalities in schizophrenia patients more sensitively without potential ceiling effects.<sup>11,70</sup> However, the number of stimuli with high or low complexity for the correlation estimate was unbalanced, with much more socially complex (69.83%) but linguistically simple (86.21%) stimuli used in previous studies. As such, certain bias might be introduced due to the design of original studies, thus mitigating the accountability and predictability of our findings on stimuli complexity.

As for other significant meta-regressive findings, the unexpected relationship between higher BPRS score and smaller between-group difference requires further exploration, because higher scores in other scales of symptom measures did not consistently predict better MSI performance of patients (Supplementary Table S7). It should be noted that the significantly positive correlations here did not indicate better MSI performance of patients than healthy controls, but instead reflected mitigated impairments (i.e., smaller negative between-group effect sizes) of patients' MSI capability. Here, we tentatively propose that the coexistence of hyposensitivity and hypersensitivity to sensory stimuli in schizophrenia may account for this correlation. [80–82] Specifically, patients tend to omit relevant sensory information and are inactive in seeking sensory input (hyposensitivity), and they, at the same time, perceive all concurrent sensory stimuli as similarly salient without effectively filtering irrelevant information because of the defective inhibitory gating (hypersensitivity). [18,80] Recent evidence found that higher sensory hypersensitivity was correlated with higher schizotypal traits and symptom severity. [82] Therefore, it could be possible that the patient group showed overall deficits in MSI than HCs partially due to sensory hyposensitivity, but patients with more severe or typical schizophrenic symptoms integrated multisensory inputs better than those with fewer symptoms because of their higher hypersensitivity level. In this case, the clinical implications of correlations between clinical characteristics (e.g., symptoms and illness duration) and patients' performance still need clearer illustrations. For example, prior studies proposed that non-significant correlations might suggest the MSI performance as a trait marker of schizophrenia, [1] which was independent of the current status and illness duration. [42] It has also been recently pointed out that neurological changes (e.g., cerebellar activation and N1 amplitude reduction) during sensory integration may serve as a possible endophenotype of schizophrenia. [8,19] Therefore, both significant and non-significant correlational statistics are clinically indispensable for our understanding about schizophrenia, and their exact clinical meanings are still open for more substantial investigations.

Our findings demonstrated certain clinical implications, as the relatively preserved audiovisual integration ability in schizophrenia may promote novel designs of cognitive training programs. Traditional cognitive training for schizophrenia patients was mainly carried out through single sensory modalities from two weeks to 12 months. [83] Although such training improved patients' certain higher-order cognition, it was still challenged for the sustainability of improvements, efficacy on symptom mitigation and the time course of training. [83–85] Our meta-analytical findings identified less impaired audiovisual integration capability (MU) in schizophrenia, so multisensory training programs are predicted to be more effective for patients, as they could comparably benefit from the information presented in different modalities. In support of our predictions, recent studies comparing the outcomes of audiovisual and unisensory training programs found that a short-term audiovisual training led to significantly stronger cognitive improvement (e.g., emotion identification,

TBW) in schizophrenia patients, and the improvements lasted a week or even longer. [86,87] Based on our meta-regression findings, those MSI tasks that sensitively elicited patients' less impaired MU procedure, such as distance judgement, lexical decision and speech-in-noise recognition, may also lead to satisfactory training outcomes. Another promising topic for training program design is to examine the transfer effects of MSI tasks, [88] that is, whether the MSI avenue is cost-efficient in enhancing any different but malleable cognitive domains. Our other reported statistics are also informative for clinical practice. For example, stimuli with high social or linguistic complexity were found to predict significantly stronger symptom-performance correlations, indicating that such complex stimuli may be used to modulate the outcomes of audiovisual training [89] and to evaluate symptom severity during early screening and prognosis of schizophrenia. Since research concerning the clinical interventions of MSI training on the recovery of patients' cognitive functioning is still emerging, substantially more studies are urgently needed.

Our study also showed certain implications for academic research. Future studies on this topic should evaluate the MSI components in greater details, including the audiovisual processing (AVP), adaptation (MM) and integration benefits (MU). Of the three principles in MSI, [90] the spatial and causal inference principles have attracted much less attention than the temporal principle in existing schizophrenia-oriented MSI research. Additionally, the clinical relevance (e.g., symptom measures, illness duration, medication) and experimental manipulation (e.g., stimuli complexity, MSI tasks, measuring indexes) to each of these components also need to be explored. For example, the influence of medication on patients' MSI performance has been mentioned as a limitation in most studies, [12,40,68,71] but there are still few studies specifically controlled and reported the effects of medication. Of all samples included in this study, almost all patients were under medication, some of whom were even taking multiple antipsychotic drugs. It is highly possible that behavior-symptom correlations in MSI tasks are altered across patients with or without taking medications, and even with different types of medications. More limitations and implications for primary studies are provided in eDiscussion (supplementary File S1).

There were important limitations to the current meta-analysis. We synthesized behavioral and neural effect sizes of audiovisual MSI performance in schizophrenia, which have introduced high heterogeneity in aggregated analyses. Although we have controlled the heterogeneity by multilevel models, subgroup analyses and moderator meta-regression, there may still be unaccountable heterogeneity. Moreover, our correlation analyses were highly constrained by the data report in the original articles. Most included studies only reported the significant correlational coefficients and omitted the non-significant ones. Due to the noticeably high variance in patients' performance, future MSI studies in the schizophrenia population are encouraged to enroll more participants and provide the exact correlational statistics. Third, our current meta-analysis emphasized the clinical relevance of MSI ability on schizophrenic symptoms, but some other dimensions, such as cognitive ability and social functioning, may provide more direct implications on how MSI capability influences real-life living of schizophrenia patients. Of all 46 quantitatively reviewed studies, however, only 15 (32.61%) measured these dimensions. Moreover, different studies used various scales or test batteries, [5,33,40,68,73] and few have established correlations between MSI performance and the score of such measures. In this case, it was infeasible to quantitatively synthesize correlational findings between MSI ability and scores other than symptoms in the current study. We encourage future studies to continue establishing the critical links between MSI and higher-order abilities, especially for schizophrenia patients. Finally, our meta-analysis provided less information for neural activation in passive viewing paradigms, and for individuals with first psychotic episodes or early-stage psychoses, due to the lack of inclusion.

In conclusion, the current study, for the first time, quantitatively synthesized effect sizes in a systematic way on audiovisual MSI in schizophrenia, which not only revealed between-group differences in a series of behavioral and neural metrics, but also illustrated the strengths of the clinical correlations with great details. Moderate-level dysfunctions were identified in the overall audiovisual MSI capability in schizophrenia, but such impairments mainly existed in audiovisual processing (AVP) and adaptation (MM), rather than the integration itself (MU). Furthermore, extremely weak

clinical correlations were observed between schizophrenic symptoms and patients' MSI performance. Clinical characteristics (e.g., BPRS score) were predictive of overall between-group differences, and stimuli with complex social or linguistic characteristics could help observe stronger clinical correlations. Our aggregated findings of audiovisual MSI, which involved a less impaired subordinate integration procedure, warrant more sophisticated investigations in future studies and provide important clinical insights towards the design of cognitive intervention programs, which will hopefully improve the quality of life for this population.

**Acknowledgments:** None.

**Authors' contributions:** **Enze Tang:** Conceptualization, Data extraction, Formal analysis, Literature search, Methodology, Software, Visualization, Writing-original draft. **Nuo Chen:** Data extraction, Formal analysis, Literature search, Writing-review & editing. **Jian Li:** Data examination, Writing-review & editing. **Hongwei Ding:** Conceptualization, Data examination, Funding acquisition, Supervision, Writing-review & editing. **Hui Chen:** Conceptualization, Funding acquisition, Supervision, Writing-review & editing.

**Conflicts of Interest.:** The authors declare no conflict of interest.

**Funding:** This work was supported by grants from Science and Technology Innovation 2030- "Brain Science and Brain-like Research" Major Project (No.2022ZD0210800), National Natural Science Foundation of China (No.32171046), Fundamental Research Funds for the Central University (226-2024-00118), Emerging Enhancement Technology under Perspective of Humanistic Philosophy supported by National Office for Philosophy and Social Sciences (No.20&ZD045) awarded to Hui Chen; and by grants from Major Program of National Office for Philosophy and Social Sciences of China (No.18ZDA293) awarded to Hongwei Ding.

## References

1. Haß K, Sinke C, Reese T, et al. Enlarged temporal integration window in schizophrenia indicated by the double-flash illusion. *Cogn Neuropsychiatry*. 2017;22(2):145-158. doi:10.1080/13546805.2017.1287693
2. Noel JP, Stevenson RA, Wallace MT. Atypical audiovisual temporal function in autism and schizophrenia: Similar phenotype, different cause. *Eur J Neurosci*. 2018;47(10):1230-1241. doi:10.1111/ejn.13911
3. Vanes LD, White TP, Wigton RL, Joyce D, Collier T, Shergill SS. Reduced susceptibility to the sound-induced flash fusion illusion in schizophrenia. *Psychiatry Res*. 2016;245:58-65. doi:10.1016/j.psychres.2016.08.016
4. Wallace MT, Stevenson RA. The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia*. 2014;64:105-123. doi:10.1016/j.neuropsychologia.2014.08.005
5. Mayer AR, Hanlon FM, Teshiba TM, et al. An fMRI study of multimodal selective attention in schizophrenia. *Br J Psychiatry*. 2015;207(5):420-428. doi:10.1192/bjp.bp.114.155499
6. Wood SM, Potts GF, Hall JF, Ulanday JB, Netsiri C. Event-related potentials to auditory and visual selective attention in schizophrenia. *Int J Psychophysiol*. 2006;60(1):67-75. doi:10.1016/j.ijpsycho.2005.05.005
7. Gao T, Wang X, Cen H, et al. Cross-modal associative memory impairment in schizophrenia. *Neuropsychologia*. 2023;191:108721. doi:10.1016/j.neuropsychologia.2023.108721
8. Ghaneirad E, Saenger E, Szycik GR, et al. Deficient audiovisual speech perception in schizophrenia: an ERP study. *Brain Sci*. 2023;13(6):970. doi:10.3390/brainsci13060970
9. Ross LA, Saint-Amour D, Leavitt VM, Molholm S, Javitt DC, Foxe JJ. Impaired multisensory processing in schizophrenia: Deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophr Res*. 2007;97(1-3):173-183. doi:10.1016/j.schres.2007.08.008
10. Mangelinckx C, Belge JB, Maurage P, Constant E. Impaired facial and vocal emotion decoding in schizophrenia is underpinned by basic perceptivo-motor deficits. *Cogn Neuropsychiatry*. 2017;22(6):461-467. doi:10.1080/13546805.2017.138234
11. Van den Stock J, de Jong SJ, Hodiament PP, de Gelder B. Perceiving emotions from bodily expressions and multisensory integration of emotion cues in schizophrenia. *Soc Neurosci*. 2011;6(5-6):537-547. doi:10.1080/17470919.2011.568790
12. Zvyagintsev M, Parisi C, Chechko N, Nikolaev AR, Mathiak K. Attention and multisensory integration of emotions in schizophrenia. *Front Hum Neurosci*. 2013;7:7. doi:10.3389/fnhum.2013.00674
13. Liu T, Pinheiro AP, Zhao Z, Nestor PG, McCarley RW, Niznikiewicz M. Simultaneous face and voice processing in schizophrenia. *Behav Brain Res*. 2016;305:76-86. doi:10.1016/j.bbr.2016.01.039
14. Senkowski D, Moran JK. Early evoked brain activity underlies auditory and audiovisual speech recognition deficits in schizophrenia. *Neuroimage Clin*. 2022;33:102909. doi:10.1016/j.nicl.2021.102909

15. Gao C, Green JJ, Yang X, Oh S, Kim J, Shinkareva SV. Audiovisual integration in the human brain: a coordinate-based meta-analysis. *Cereb Cortex*. 2023;33(9):5574-5584. doi:10.1093/cercor/bhac443
16. Ding H, Zhang Y. Speech prosody in mental disorders. *Annu Rev Linguist*. 2023;9:335-355. doi:10.1146/annurev-linguistics-030421-065139
17. Adámek P, Langová V, Horáček J. Early-stage visual perception impairment in schizophrenia, bottom-up and back again. *Schizophrenia*. 2022;8(1):27. doi:10.1038/s41537-022-00237-9
18. Javitt DC, Sweet RA. Auditory dysfunction in schizophrenia: integrating clinical and basic features. *Nat Rev Neurosci*. 2015;16(9):535-550. doi:10.1038/nrn4002
19. Li Z, Huang J, Hung KSY, et al. Cerebellar hypoactivation is associated with impaired sensory integration in schizophrenia. *J Abnorm Psychol*. 2021;130(1):102-111. doi:10.1037/abn0000636
20. Pearl D, Yodashtkin-Porat D, Katz N, et al. Differences in audiovisual integration, as measured by McGurk phenomenon, among adult and adolescent patients with schizophrenia and age-matched healthy control groups. *Compr Psychiatry*. 2009;50(2):186-192. doi:10.1016/j.comppsy.2008.06.004
21. Stevenson RA, Park S, Cochran C, et al. The associations between multisensory temporal processing and symptoms of schizophrenia. *Schizophr Res*. 2017;179:97-103. doi:10.1016/j.schres.2016.09.035
22. Surguladze SA, Calvert GA, Brammer MJ, et al. Audio-visual speech perception in schizophrenia: An fMRI study. *Psychiatry Res Neuroimaging*. 2001;106(1):1-14. doi:10.1016/S0925-4927(00)00081-0
23. Feldman JL, Dunham K, Cassidy M, Wallace MT, Liu Y, Woynaroski TG. Audiovisual multisensory integration in individuals with autism spectrum disorder: a systematic review and meta-analysis. *Neurosci Biobehav Rev*. 2018;95:220-234. doi:10.1016/j.neubiorev.2018.09.020
24. Gröhn C, Norgren E, Eriksson L. A systematic review of the neural correlates of multisensory integration in schizophrenia. *Schizophr Res Cogn*. 2022;27:25. doi:10.1016/j.scog.2021.100219
25. Tseng HH, Bossong MG, Modinos G, et al. A systematic review of multisensory cognitive-affective integration in schizophrenia. *Neurosci Biobehav Rev*. 2015;55:444-452. doi:10.1016/j.neubiorev.2015.04.019
26. Bremner AJ, Lewkowicz DJ, Spence C. *The multisensory approach to development*. In: Bremner AJ, Lewkowicz DJ, Spence C, eds. *Multisensory Development*. Oxford, UK: Oxford University Press; 2012:1-26.
27. Hershenson M. Reaction time as a measure of intersensory facilitation. *J Exp Psychol*. 1962;63:289-293. doi:10.1037/h0039516
28. Miller J. Divided attention: Evidence for coactivation with redundant signals. *Cogn Psychol*. 1982;14(2):247-279. doi:10.1016/0010-0285(82)90010-x
29. de Boer-Schellekens L, Stekelenburg JJ, Maes JP, Van Gool AR, Vroomen J. Sound improves diminished visual temporal sensitivity in schizophrenia. *Acta Psychol*. 2014;147:136-142. doi:10.1016/j.actpsy.2013.06.013
30. Martin B, Giersch A, Huron C, van Wassenhove V. Temporal event structure and timing in schizophrenia: preserved binding in a longer "now". *Neuropsychologia*. 2013;51(2):358-371. doi:10.1016/j.neuropsychologia.2012.07.002
31. Szycik GR, Ye Z, Mohammadi B, et al. Maladaptive connectivity of Broca's area in schizophrenia during audiovisual speech perception: An fMRI study. *Neuroscience*. 2013;253:274-282. doi:10.1016/j.neuroscience.2013.08.041
32. Gao C, Weber CE, Shinkareva SV. The brain basis of audiovisual affective processing: evidence from a coordinate-based activation likelihood estimation meta-analysis. *Cortex*. 2019;120:66-77. doi:10.1016/j.cortex.2019.05.016
33. Balz J, Roa Romero Y, Keil J, et al. Beta/Gamma oscillations and event-related potentials indicate aberrant multisensory processing in schizophrenia. *Front Psychol*. 2016;7:1896. doi:10.3389/fpsyg.2016.01896
34. Roa Romero Y, Keil J, Balz J, Niedeggen M, Gallinat J, Senkowski D. Alpha-band oscillations reflect altered multisensory processing of the McGurk illusion in schizophrenia. *Front Hum Neurosci*. 2016;10:41. doi:10.3389/fnhum.2016.00041
35. Stone DB, Urrea LJ, Aine CJ, Bustillo JR, Clark VP, Stephen JM. Unisensory processing and multisensory integration in schizophrenia: a high-density electrical mapping study. *Neuropsychologia*. 2011;49(12):3178-3187. doi:10.1016/j.neuropsychologia.2011.07.017
36. Thaler NS, Strauss GP, Sutton GP, et al. Emotion perception abnormalities across sensory modalities in bipolar disorder with psychotic features and schizophrenia. *Schizophr Res*. 2013;147(2-3):287-292. doi:10.1016/j.schres.2013.04.001
37. Zvyagintsev M, Parisi C, Mathiak K. Temporal processing deficit leads to impaired multisensory binding in schizophrenia. *Cogn Neuropsychiatry*. 2017;22(5):361-372. doi:10.1080/13546805.2017.1331160
38. Foucher JR, Lacambre M, Pham BT, Giersch A, Elliott MA. Low time resolution in schizophrenia: lengthened windows of simultaneity for visual, auditory and bimodal stimuli. *Schizophr Res*. 2007;97:118-127. doi:10.1016/j.schres.2007.08.013
39. Müller VI, Kellermann TS, Seligman SC, Turetsky BI, Eickhoff SB. Modulation of affective face processing deficits in schizophrenia by congruent emotional sounds. *Soc Cogn Affect Neurosci*. 2014;9(4):436-444. doi:10.1093/scan/nss107



40. Vogel BD, Bruck C, Jacob H, Eberle M, Wildgruber D. Effects of cue modality and emotional category on recognition of nonverbal emotional signals in schizophrenia. *BMC Psychiatry*. 2016;16:9. doi:10.1186/s12888-016-0913-7
41. Zhou HY, Cui XL, Yang BR, et al. Audiovisual temporal processing in children and adolescents with schizophrenia and children and adolescents with autism: evidence from simultaneity-judgment tasks and eye-tracking data. *Clin Psychol Sci*. 2021;10(3):482-498. doi:10.1177/21677026211031543
42. Lema YY, Gamo NJ, Yang K, Ishizuka K. Trait and state biomarkers for psychiatric disorders: Importance of infrastructure to bridge the gap between basic and clinical research and industry. *Psychiat Clin Neurosci*. 2018;72(7):482-489. doi:10.1111/pcn.12669
43. Garrido-Vásquez P, Jessen S, Kotz SA. Perception of emotion in psychiatric disorders: on the possible role of task, dynamics, and multimodality. *Soc Neurosci*. 2011;6(5-6):515-536. doi:10.1080/17470919.2011.620771
44. Lin Y, Ding H, Zhang Y. Multisensory integration of emotion in schizophrenic patients. *Multisens Res*. 2020;33(8): 865-901. doi:10.1163/22134808-bja10016
45. Thoenes S, Oberfeld D. Meta-analysis of time perception and temporal processing in schizophrenia: differential effects on precision and accuracy. *Clin Psychol Rev*. 2017;54:44-64. doi:10.1016/j.cpr.2017.03.007
46. Zhou HY, Cai XL, Weigl M, Bang P, Cheung EF, Chan RC. Multisensory temporal binding window in autism spectrum disorders and schizophrenia spectrum disorders: a systematic review and meta-analysis. *Neurosci Biobehav Rev*. 2018;86:66-76. doi:10.1016/j.neubiorev.2017.12.013
47. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. doi:10.1371/journal.pmed.1000097
48. Hornix BE, Havekes R, Kas MJH. Multisensory cortical processing and dysfunction across the neuropsychiatric spectrum. *Neurosci Biobehav Rev*. 2019;97:138-151. doi:10.1016/j.neubiorev.2018.02.010
49. Aslam S, Emmanuel P. Formulating a researchable question: a critical step for facilitating good clinical research. *Indian J Sex Transm Dis Aids*. 2010;31(1):47-50. doi:10.4103/0253-7184.69003
50. Moran JK, Keil J, Masurovsky A, Gutwinski S, Montag C, Senkowski D. Multisensory processing can compensate for top-down attention deficits in schizophrenia. *Cereb Cortex*. 2021;31(12):5536-5548. doi:10.1093/cercor/bhab177
51. Nikitova N, Keane BP, Demmin D, Silverstein SM, Uhlhaas PJ. The audio-visual abnormalities questionnaire (AVAQ): development and validation of a new instrument for assessing anomalies in sensory perception in schizophrenia spectrum disorders. *Schizophr Res*. 2019;209:227-233. doi:10.1016/j.schres.2019.03.016
52. Curtis MT, Ren X, Coffman BA, Salisbury DF. Attentional M100 gain modulation localizes to auditory sensory cortex and is deficient in first-episode psychosis. *Hum Brain Mapp*. 2023;44(1):218-228. doi:10.1002/hbm.26067
53. Stekelenburg JJ, Maes JP, Van Gool AR, Sitskoorn M, Vroomen J. Deficient multisensory integration in schizophrenia: an event-related potential study. *Schizophr Res*. 2013;147(2-3):253-261. doi:10.1016/j.schres.2013.04.038
54. Hirano Y, Nakamura I, Tamura S. Abnormal connectivity and activation during audiovisual speech perception in schizophrenia. *Eur J Neurosci*. 2024;59(8):1918-1932. doi:10.1111/ejn.16183
55. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Commun Health*. 1998;52(6):377-384. doi:10.1136/jech.52.60.377
56. Wells GA, Shea B, O'Connell D, et al. *The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses*. Ottawa Health Research Institute. Available from: [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.asp](http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp)
57. Viechtbauer W, Cheung MW. Outlier and influence diagnostics for meta-analysis. *Res Synth Methods*. 2010;1(2):112-125. doi:10.1002/jrsm.11
58. Pustejovsky JE. *clubSandwich: cluster-robust (sandwich) variance estimators with small-sample corrections*. 2022. Available from: URL: <https://CRAN.R-project.org/package=clubSandwich>
59. Pustejovsky JE, Tipton E. Meta-analysis with robust variance estimation: expanding the range of working models. *Prev Sci*. 2022;23(3):425-438. doi:10.1007/s11121-021-01246-3
60. Hedges LV, Tipton E, Johnson MC. Robust variance estimation in meta-regression with dependent effect size estimates. *Res Synth Methods*. 2010;1(1):39-65. doi:10.1002/jrsm.5
61. Tipton E. Small sample adjustments for robust variance estimation with meta-regression. *Psychol Methods*. 2015;20(3):375-393. doi:10.1037/met0000011
62. Radua J, Mataix-Cols D, Phillips ML, et al. A new meta-analytic method for neuroimaging studies that combines reported peak coordinates and statistical parametric maps. *Eur Psychiatry*. 2012;27(8):605-611. doi:10.1016/j.eurpsy.2011.04.001



63. Albajes-Eizaguirre A, Solanes A, Vieta E, Radua J. Voxel-based meta-analysis via permutation of subject images (PSI): theory and implementation for SDM. *Neuroimage*. 2019;186:174-184. doi:10.1016/j.neuroimage.2018.10.077
64. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol*. 2008;61(10):991-996. doi:10.1016/j.jclinepi.2007.11.010
65. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-634. doi:10.1136/bmj.315.7109.629
66. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*. 2000;56(2):455-463. doi:10.1111/j.0006-341x.2000.00455.x
67. Rosenthal R. *Meta-analytic procedures for social research*. Thousand Oaks, CA: SAGE; 1991.
68. Sass K, Heim S, Sachs O, et al. Neural correlates of semantic associations in patients with schizophrenia. *Eur Arch Psychiatry Clin Neurosci*. 2014;264(2):143-154. doi:10.1007/s00406-013-0425-0
69. Smucny J, Rojas DC, Eichman LC, Tregellas JR. Neural effects of auditory distraction on visual attention in schizophrenia. *PLoS One*. 2013;8(4):e60606. doi:10.1371/journal.pone.0060606
70. Straube B, Green A, Sass K, Kirner-Veselinovic A, Kircher T. Neural integration of speech and gesture in schizophrenia: evidence for differential processing of metaphoric gestures. *Hum Brain Mapp*. 2013;34(7):1696-1712. doi:10.1002/hbm.22015
71. Stone DB, Coffman BA, Bustillo JR, Aine CJ, Stephen JM. Multisensory stimuli elicit altered oscillatory brain responses at gamma frequencies in patients with schizophrenia. *Front Hum Neurosci*. 2014;8:788. doi:10.3389/fnhum.2014.00788
72. Surguladze S, Rossell S, Rabe-Hesketh S, David AS. Cross-modal semantic priming in schizophrenia. *J Int Neuropsychol Soc*. 2002;8(7):884-892. doi:10.1017/S1355617702870023
73. Sanfratello L, Aine C, Stephen J. Neuroimaging investigations of dorsal stream processing and effects of stimulus synchrony in schizophrenia. *Psychiatry Res Neuroimaging*. 2018;278:56-64. doi:10.1016/j.psychres.2018.05.005
74. Wynn JK, Jahshan C, Green MF. Multisensory integration in schizophrenia: a behavioral and event-related potential study. *Cogn Neuropsychiatry*. 2014;19(4):319-336. doi:10.1080/13546805.2013.866892
75. de Jong JJ, Hodiamont PP, de Gelder B. Modality-specific attention and multisensory integration of emotions in schizophrenia: reduced regulatory effects. *Schizophr Res*. 2010;122(1-3):136-143. doi:10.1016/j.schres.2010.04.010
76. de Gelder B, Vroomen J, de Jong SJ, Masthoff ED, Trompenaars FJ, Hodiamont P. Multisensory integration of emotional faces and voices in schizophrenics. *Schizophr Res*. 2005;72(2-3):195-203. doi:10.1016/j.schres.2004.02.013
77. Zhou HY, Lai IYS, Hung KSY, et al. Audiovisual temporal processing in adult patients with first-episode schizophrenia and high-functioning autism. *Schizophrenia*. 2022;8(1):9. doi:10.1038/s41537-022-00284-2
78. Jones SA, Noppeney U. *Multisensory integration and causal inference in typical and atypical populations*. In: Gu Y, Zaidel A, eds. *Advances of Multisensory Integration in the Brain*. Adv Exp Med Biol. Singapore: Springer; 2024;1437:59-75. doi:10.1007/978-981-99-7611-9\_4
79. Roa Romero Y, Keil J, Balz J, Gallinat J, Senkowski D. Reduced frontal theta oscillations indicate altered crossmodal prediction error processing in schizophrenia. *J Neurophysiol*. 2016;116(3):1396-1407. doi:10.1152/jn.00096.2016
80. Brown C, Cromwell RL, Filion D, Dunn W, Tollefson N. Sensory processing in schizophrenia: missing and avoiding information. *Schizophr Res*. 2002;55:187-195. doi:10.1016/s0920-9964(01)00255-9
81. van den Boogert F, Klein K, Spaan P, et al. Sensory processing difficulties in psychiatric disorders: a meta-analysis. *J Psychiatr Res*. 2022;151:173-180. doi:10.1016/j.jpsychires.2022.04.020
82. Zhou HY, Yang HX, Cui XL, et al. Self-reported sensory responsiveness patterns in typically-developing and early-onset schizophrenia adolescents: its relationship with schizotypal and autistic traits. *J Psychiatr Res*. 2020;131:255-262. doi:10.1016/j.jpsychires.2020.10.002
83. Dondé C, Mondino M, Brunelin J, Haesebaert F. Sensory-targeted cognitive training for schizophrenia. *Expert Rev Neurother*. 2019;19(3):211-225. doi: 10.1080/14737175.2019.1581609
84. Bechi M, Bosia M, Spangaro M, et al. Visual and audio emotion processing training for outpatients with schizophrenia: an integrated multisensory approach. *Neuropsychol Rehabil*. 2018;28(7):1131-1144. doi:10.1080/09602011.2016.1240698
85. Scoriels L, Genaro LT, Keffer S, et al. Changes in emotion processing and social cognition with auditory versus visual neuroscience-informed cognitive training in individuals with schizophrenia. *Schizophr Res*. 2022;241:267-274. doi: 10.1016/j.schres.2022.02.013
86. Jeong JW, Kim HT, Lee SH, Lee H. Effects of an audiovisual emotion perception training for schizophrenia: a preliminary study. *Front Psychiatry*. 2021;12:522094. doi:10.3389/fpsy.2021.522094

87. Zerr M, Freihorst C, Schütz H, et al. Brief sensory training narrows the temporal binding window and enhances long-term multimodal speech perception. *Front Psychol.* 2019;10:2489. doi:10.3389/fpsyg.2019.02489
88. Powers AR III, Hillock-Dunn A, Wallace MT. Generalization of multisensory perceptual learning. *Sci Rep.* 2016;6:23374. doi:10.1038/srep23374
89. De Nier MA, Gupta PB, Baum SH, Wallace MT. Perceptual training enhances temporal acuity for multisensory speech. *Neurobiol Learn Mem.* 2018;147:9-17. doi:10.1016/j.nlm.2017.10.016
90. Stein BE, Stanford TR. Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci.* 2008;9(4):255-266. doi:10.1038/nrn2331

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.