Differences between autistic and non-autistic individuals in audiovisual speech integration: a systematic review and meta-analysis

Published version at: https://www.sciencedirect.com/science/article/pii/S0149763424002562
To link to this article DOI: http://dx.doi.org/10.1016/j.neubiorev.2024.105787

Publisher: Elsevier

Neuroscience and Biobehavioral Reviews 164 (2024) 105787

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Differences between autistic and non-autistic individuals in audiovisual speech integration: A systematic review and meta-analysis

Robert M. Jertberg [a,*], Frederik J. Wienicke [b], Krystian Andruszkiewicz [a], Sander Begeer [a], Bhismadev Chakrabarti [e,f,g], Hilde M. Geurts [c,d], Ralph de Vries [h], Erik Van der Burg [a,c]

[a] Department of Clinical and Developmental Psychology, Vrije Universiteit Amsterdam, The Netherlands and Amsterdam Public Health Research Institute, Amsterdam, the Netherlands
[b] Department of Clinical Psychology, Behavioural Science Institute, Radboud University, Nijmegen, the Netherlands
[c] Department of Psychology, Universiteit van Amsterdam, the Netherlands
[d] Leo Kannerhuis (Youz/Parnassiagroup), the Netherlands
[e] Centre for Autism, School of Psychology and Clinical Language Sciences, University of Reading, UK
[f] India Autism Center, Kolkata, India
[g] Department of Psychology, Ashoka University, India
[h] Medical Library, Vrije Universiteit, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Autism spectrum disorder
Speech
Audiovisual integration
McGurk
Multisensory

ABSTRACT

Research has indicated unique challenges in audiovisual integration of speech among autistic individuals, although methodological differences have led to divergent findings. We conducted a systematic literature search to identify studies that measured audiovisual speech integration among both autistic and non-autistic individuals. Across the 18 identified studies (combined N = 952), autistic individuals showed impaired audiovisual integration compared to their non-autistic peers ($g = 0.69$, 95 % CI [0.53, 0.85], $p < .001$). This difference was not found to be influenced by participants' mean ages, studies' sample sizes, risk-of-bias scores, or paradigms employed. However, a subgroup analysis suggested that child studies may show larger between-group differences than adult ones. The prevailing pattern of impaired audiovisual speech integration in autism may have cascading effects on communicative and social behavior. However, small samples and inconsistency in designs/analyses translated into considerable heterogeneity in findings and opacity regarding the influence of underlying unisensory and attentional factors. We recommend three key directions for future research: larger samples, more research with adults, and standardization of methodology and analytical approaches.

The integration of information from different sensory streams that correspond in meaningful ways (e.g., temporal or spatial alignment) is a powerful means of reducing ambiguity and enhancing perception (Sumby and Pollack, 1954; Van der Burg et al., 2008; van Ee et al., 2009; Vroomen and de Gelder, 2000). The interactions between the auditory and visual domains are an area of particular interest due to their importance to the processing of speech stimuli. The integration of perceptually distinct auditory components of speech (known as *phonemes*) with their visual counterparts (known as *visemes*) can lead to significant improvements in speech perception, particularly in noisy/ambiguous contexts (Erber, 1969; Irwin and DiBlasi, 2017; Sumby and Pollack, 1954). Furthermore, the integration of incongruent audiovisual speech stimuli can produce powerful illusions, such as the McGurk/MacDonald effect (Mcgurk and Macdonald, 1976). Here, the presentation of a mismatched phoneme and viseme pair leads participants to either perceive a third syllable that lies between them, known as a fusion, or to "hear what they see," in a phenomenon known as visual capture. In this way, audiovisual integration can lead to both enhancement and distortion of language perception.

Significant differences exist in how prone individuals are to integrating sight and sound (Magnotti and Beauchamp, 2018), which are particularly pronounced when comparing the general population to certain clinical groups, such as individuals with autism[1] (i.e., autism spectrum disorder, see Stevenson et al., 2014). In autism, atypical

* Correspondence to: Vrije Universiteit Amsterdam, De Boelelaan 1105, Amsterdam 1081 HV, the Netherlands.
E-mail address: jertbergiii@gmail.com (R.M. Jertberg).

[1] We alternate between person-first and identity-first language to address the varying preferences of the international autism community, as recommended in Buijsman et al. (2023).

https://doi.org/10.1016/j.neubiorev.2024.105787
Received 22 December 2023; Received in revised form 15 May 2024; Accepted 24 June 2024
Available online 28 June 2024

multisensory processing has often been speculated to be important to broader behavioral differences (Brock et al., 2002). It is theorized that basic multisensory differences might have a cascading influence via speech perception on the higher level social factors commonly viewed as the core features of autism (Stevenson et al., 2018). Some "sensory first" theories even posit that sensory differences are most essential to the condition, due to the fact that they precede and predict higher level ones (Cascio et al., 2016; Falck-Ytter and Bussu, 2023; Robertson and Baron-Cohen, 2017). If these theories are correct, studying multisensory integration (MSI) in autism could be crucial to our understanding of the condition.

While MSI (in this case, primarily audiovisual interactions) can greatly benefit speech perception throughout the lifetime, it may play an especially important role during early language learning (Mason et al., 2019). Considering that many autistic individuals demonstrate delays in language acquisition and persistently atypical processing of speech stimuli (Brignell et al., 2018; Mody and Belliveau, 2013), it is plausible that differences in multisensory processing may reach far beyond basic sensory perception via an influence on language. What is more, the precise types of noisy environments in which autistic individuals struggle the most with speech perception (Alcántara et al., 2004; Fadeev et al., 2023; Mamashli et al., 2017) are the same in which visual cues afford the greatest cross-modal advantage (MacLeod and Summerfield, 1987; Sumby and Pollack, 1954). This provides further evidence for a unique relationship between multisensory processing differences and autism. At the convergence of these lines of research into language, MSI, and autism is a clear scientific imperative: determining the degree to which multisensory processing of speech stimuli is affected in autism and how this might translate into higher order differences.

Two meta-analyses have spoken to this question on different levels. Feldman et al. (2018) investigated audiovisual integration among autistic and non-autistic individuals broadly, finding differences across stimulus types and an indication that integration of speech stimuli may be especially affected in autism. These differences were particularly pronounced in younger samples. Zhang et al. (2019) focused specifically on the McGurk/MacDonald effect in individuals with and without autism. They found that autistic individuals demonstrated diminished susceptibility to the illusion, consistent with the conclusions of Feldman et al. (2018). On the other hand, at odds with the findings of Feldman et al. (2018), they reported a positive correlation between the difference in MSI magnitude and age. This led Zhang et al. (2019) to conclude that autistic people do not improve in their ability to integrate audiovisual speech stimuli, resulting in a growing disparity between them and non-autistic individuals, who continue to develop this ability into adulthood. However, the studies included in their meta-analysis had a rather limited age range, with the mean ages of their samples ranging from 7.9 to 14.9 years old, with the exception of a single study including adults. This is problematic in light of other research that suggests differences between autistic and non-autistic children in the ability to integrate audiovisual speech stimuli may be resolved during adolescence (Foxe et al., 2015; Taylor et al., 2010). These inconsistencies raise questions about the influence of age on potential differences in audiovisual integration of speech stimuli in autism.

Additionally, contradictory findings arise both within McGurk/MacDonald studies and when comparing them to other paradigms. Given that McGurk/MacDonald experiments make up the entirety of the Zhang et al. (2019) meta-analysis and the vast majority of the linguistic studies included in the Feldman et al. (2018) meta-analysis, this complicates their interpretations. While the McGurk/MacDonald effect is often considered a hallmark of MSI, its illusory nature depends upon a violation of the cross-modal congruence of normal, everyday speech stimuli. This begs the question of how much can be inferred from such a contrived paradigm relative to more realistic designs like speech in noise experiments (e.g., Sumby and Pollack, 1954), which more directly emulate the type of noisy environments autistic people struggle with the most in real life. One study that compared McGurk/MacDonald and speech in noise paradigm performance with the same sample found no relationship between susceptibility to the illusion and the ability to use visual cues to understand speech embedded in noise, challenging the assumption that McGurk/MacDonald is an ecologically valid measure of audiovisual speech integration (Van Engen et al., 2017).

Furthermore, a review by Magnotti and Beauchamp (2018) posited that between-group differences in the McGurk/MacDonald effect have been inflated in autism research due to small sample sizes and publication bias. They noted a high degree of variability in the findings comparing autistic and non-autistic participants (ranging from 45 % more fusions among non-autistic individuals to 10 % fewer, with no difference detected in the study with the largest sample). Given that McGurk/MacDonald studies make up the majority of the multisensory speech processing literature, this raises questions about the potential influence of sample size on findings. However, the review by Magnotti and Beauchamp (2018) derived conclusions from observed trends in the data and statistical modelling assuming a publication bias rather than a formal meta-analytic approach, which could offer further evidence in favor (or against) their hypothesis.

Taken together, these considerations motivate a review that focuses specifically on the unique features of audiovisual speech processing in autism but extends beyond the McGurk/MacDonald effect, including both illusory and more ecologically valid paradigms. Sub-group analyses are necessary to address the inconsistencies in findings relative to paradigm type, age, and sample size. This undertaking is both practically important to understanding the impediments to communication that autistic individuals face and theoretically important due to the crucial role MSI plays in sensory-first theories of autism. If differences in MSI disappear as autistic individuals reach maturity, it would suggest that they are at most precursors to later behavioral differences (rather than persistent underlying causes of them), and the targets for therapeutic interventions may need to be adjusted accordingly. Multisensory training has been shown to be very effective (Stevenson et al., 2013), even in older adults (Setti et al., 2014), and may serve as a useful tool in improving the quality of life of autistic individuals, but it is imperative that we understand at what ages such an intervention would be of value.

To this end, we conducted a meta-analysis investigating audiovisual integration of speech stimuli in autism. We examined first whether autistic individuals differ from those without autism in their ability to integrate speech information across the senses. We then delved deeper into with which paradigms and at which ages such a difference might be detectable, with attention to enhancement of perception when integrating sight and sound and distortion of perception in illusory paradigms like the McGurk/MacDonald effect. Finally, we investigated what influence sample size might have on any potential disparity in the magnitude of the effect between autistic and non-autistic individuals.

1. Methods

1.1. Design

This study was pre-registered in the PROSPERO international prospective register of systematic reviews (CRD42022339828) and is reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Page et al., 2021; see online Supplementary Table S.1 for the PRISMA checklist). The extracted effect size data of the included studies as well as the R code of the analyses are available at https://osf.io/fvw4t/?view_only=3fab3d0612824a198716a51ec1c574c2.

1.2. Search strategy

In collaboration with a medical information specialist (RdV), we conducted systematic searches in the online databases PubMed, Embase.com, PsycINFO, Web of Science, and Scopus, from inception to October 9th, 2023. For these searches, we used an extensive search string including synonyms for the following terms as index terms or free-text words: "Pervasive Child Development Disorders", "Autism", "Asperger", "Visual streams", "Auditory streams", and "Speech" (see online Supplementary Table S.2 for the full search string applied in PubMed; the others are also available upon request). Additionally, references of the identified articles were screened for missed relevant publications. No language or date restrictions were applied.

1.3. Study selection

Relevant studies measured audiovisual integration of speech stimuli using visual stimuli to enhance/distort perception of auditory stimuli (or vice versa) and compared non-autistic individuals and individuals diagnosed with autism. Studies were excluded if they were conference proceedings or investigated the relationship between autistic traits (rather than official diagnoses by licensed medical health professionals) and integration. Two independent raters (RJ and FW) screened the titles and abstracts of all potentially relevant studies for eligibility. For all studies that could not definitely be excluded, full texts were requested and checked against the in- and exclusion criteria by the two independent raters. At each step, disagreement between the raters was resolved through consensus. If consensus could not be reached, a third rater was consulted. Fig. 1 depicts the search and selection process.

1.4. Data extraction

Two independent reviewers (RJ and FW) extracted the following data: a) study authors, b) paradigm type, c) mean age, range, and standard deviation (for each group), d) diagnosis type, e) mean IQ, range, and standard deviation (for each group), f) sample size (for each group), g) effect type, h) effect direction, i) other statistical test in lieu of differences in means (*F* or *t* statistic preferably, but Mann-Whitney U test in one case, and regression coefficient in another, all converted to Cohen's *d*). If consensus could not be reached, a third reviewer was consulted.

1.5. Risk-of-bias assessment

Risk of bias of included studies was evaluated by two raters (RJ and KA) using an adapted version of the Newcastle-Ottawa risk of bias assessment tool for case/control studies (Wells et al., 2014). The Newcastle-Ottawa scale is the preferred risk of bias assessment tool for non-randomized studies according to the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2023). It involves a

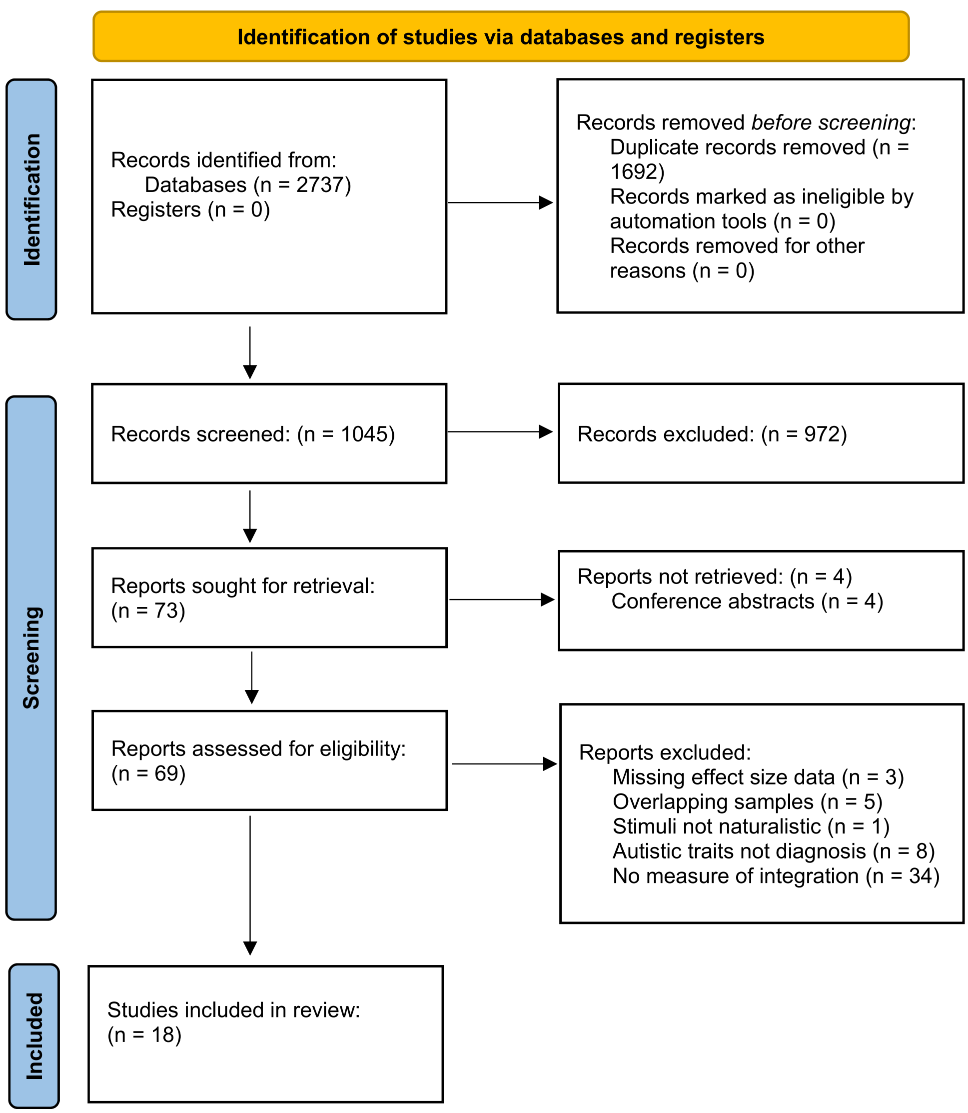

Fig. 1. PRISMA 2020 Flow Diagram for Systematic Reviews. © Reproduced with permission of the PRISMA Group, which encourages sharing and reuse for non-commercial purpose.

rating system in which studies are given higher scores for ensuring that case and control groups are representative, well-defined, and well-matched. Further information on our adaptation of the scale can be found in online Supplementary Text S.1.

1.6. Data analysis

Effect sizes of included studies were pooled using the random effects model since studies were expected to differ in their underlying variance. The analyses used Hedge's *g* because this measure has been argued to be more robust than other effect sizes (such as Cohen's *d*) when sample sizes are small or differ between groups (Cooper et al., 2009). A Hedge's *g* value of ≤ 0.32 is considered small, 0.33–0.55 moderate, and ≥ 0.56 large (Lipsey and Wilson, 1993). To calculate Hedge's *g* for each study, we used the reported means, standard deviations, and sample sizes of the autistic and non-autistic groups on the task (e.g., McGurk/MacDonald task). If this information was not reported in the publication, we based effect sizes on the *F*, *t*, or Mann-Whitney *U* statistics, or regression coefficient, in this order. If there was no single pre-specified primary outcome or task assessing audiovisual integration reported, we extracted the effect size information for all tasks and pooled these in the main analyses. If data for participants diagnosed with autism spectrum disorder and Asperger's were reported separately, they were pooled, as these conditions are no longer distinguished in diagnostic practice (American Psychiatric Association, 2013). Heterogeneity was assessed by the I^2 statistics and prediction intervals. Additionally, we conducted subgroup analyses using the same meta-analytic approach as previously described, comparing studies which examined either adults or children and studies which used the McGurk/MacDonald or speech in noise paradigm. Furthermore, we investigated whether effect sizes varied as a function of participants' mean ages or studies' sample sizes using meta-regression analyses. Similarly, we investigated the robustness of our findings via a meta-regression of the risk-of-bias scores. Additionally, we conducted a post-hoc sensitivity analysis excluding one study (Irwin et al., 2023) due to the fact that behavioral data was only reported for a subgroup of participants (for whom demographic information was not reported or made available upon request) because of technical and performance issues. Potential publication bias was examined by contour-enhanced funnel plots and Egger's test of the intercept. All analyses were conducted using R (version 4.0.3; R Core Team, 2020) and the meta and metafor packages (Balduzzi et al., 2019; Viechtbauer, 2010).

2. Results

2.1. Included studies

Fig. 1 shows the systematic literature search, which resulted in a total of 2737 records (1045 after deduplication). After screening of the studies' titles and abstracts, 972 were excluded and 73 were reviewed in full text. Of these, 55 were excluded, with the most common reasons for exclusion being not reporting a measure of audiovisual integration ($k = 34$) or examining participants with no official autism diagnosis ($k = 8$). One study (Williams et al., 2004) was excluded as the visual stimuli were animated avatars that were not as naturalistic as the videos used in other studies. In instances where overlapping samples were identified, we excluded the study with the smaller sample. Ultimately, 18 studies remained for inclusion in the meta-analysis. Table 1 provides an overview of the study characteristics of the 18 included studies comprising a total of 952 participants.

2.2. Bias assessment

The outcomes of the risk of bias assessment are shown in Table 1. Considerable risk of bias was detected, with a mean rating of 3.6 on a scale of 0 (worst)–6 (best). This is largely due to lack of information on either selection or definition of controls (16/18 studies). Many studies also failed to qualify for one or more comparability points as they did not screen for hearing/vision problems or because they matched on group means rather than individually (10/18 studies). As shown in Fig. 2, the contour-enhanced funnel plot showed a minor degree of asymmetry. However, this asymmetry was negligible, as indicated by the non-significant Egger's test of the intercept ($\beta 0 = 0.71$, $SE = 0.12$, $p = .899$).

2.3. Meta-analysis

Results of the individual studies and the pooled effect size are presented in Fig. 3. Across the 18 included studies, the mean difference in audiovisual integration between individuals with autism and controls was $g = 0.69$ (95 % CI [0.53, 0.85], $p < .001$, $I^2 = 26.66$ %, 95 % PI [0.28, 1.10]), indicating that autistic individuals showed impaired audiovisual integration. Effect size seemed unaffected by the risk of bias scores ($\beta = -0.05$ (95 % CI [-0.24, 0.13], $p = .542$). Additionally, when excluding one study (Irwin et al., 2023), which was subject to technical difficulties recording responses and did not report demographic information for the subgroup of participants whose behavioral data were analyzed, a similar pattern of results emerged ($g = 0.70$ (95 % CI [0.53, 0.87], $p < .001$, $I^2 = 30.41$ %, 95 % PI [0.25, 1.14]). Furthermore, the difference in audiovisual integration was neither significantly influenced by the sample sizes of studies ($\beta = -0.001$, (95 % CI [-0.005, 0.005], $p = .989$) nor the mean age of participants ($\beta = -0.02$, 95 % CI [-0.05, 0.004], $p = .099$). There was, however, a statistically significant difference in the magnitude of the disparity in MSI between autistic and non-autistic people when comparing studies involving children (mean age < 18, $k = 16$, $g = 0.75$, 95 % CI [0.58, 0.94]) to studies with adults (mean age ≥ 18, $k = 2$, $g = 0.29$, 95 % CI [-1.95, 2.53]; $Q = 5.51$, $p = .019$). Lastly, there was no difference in effect sizes between studies using either audiovisual distortion (i.e., McGurk/MacDonald and phonemic restoration paradigms; $k = 13$, $g = 0.72$, 95 % CI [0.49, 0.96]) or audiovisual enhancement (i.e., speech in noise paradigm; ($k = 3$, $g = 0.71$, 95 % CI [0.10, 1.31]; $Q = 0.01$, $p = .929$); similar findings emerged when including studies which used both types of paradigms (see online Supplementary Table S.3).

3. Discussion

3.1. Summary of findings

In this meta-analysis, we sought to determine whether autistic individuals differ from non-autistic ones in their ability to integrate audiovisual speech stimuli and the influences paradigm, age, and sample size might have on such a difference. Overall, we detected a moderate to strong effect of diagnostic group, with autistic participants showing diminished audiovisual integration relative to their peers. This finding appeared robust to control for the risk of bias of included studies. When comparing paradigm types, we found no difference between approaches. In meta-regressions, neither mean age nor mean sample size was found to influence the magnitude of the difference between groups. However, in a sub-group analysis comparing adult studies to child ones, it was found that the child studies reported larger differences between groups.

The main finding, that autistic individuals showed less integration of audiovisual speech stimuli, resonates with prior meta-analyses (Feldman et al., 2018; Zhang et al., 2019). Where our results differ from prior findings and break new ground is in the outcomes of our subgroup and meta-regression analyses. Our meta-analysis was the first to conduct a sub-group analysis comparing audiovisual enhancement (i.e. the gain afforded in speech in noise paradigms) to distortion (i.e. the McGurk/MacDonald effect and phonemic restoration). While methodological differences led to greater diversity and, therefore, less clarity in studies measuring distortion of perception, the magnitude of effects detected therein did not differ from those measuring enhancement of perception. Our meta-analysis was also the first to formally test the

Table 1
Characteristics of included studies.

Study	Paradigm	N ASD	N TD	Age ASD	Age TD	IQ ASD	IQ TD	ROB Score	Hearing/vision check	Unisensory control	Visual attention control
Bebko et al. (2014)	McGurk	30	19	11.21 (2.56) **	10.21 (2.74)	97.1 (17.64) **	99.05 (9.51)	4	Parental report	No significant differences	Experimenter judged whether looking at screen
DePape et al. (2012)	McGurk	27	27	15 (1.7)	14.6 (2)	NR	NR	1	Hearing tested	No significant differences	None
Feldman, Conrad, et al. (2022)	McGurk	18	18	12.58 (3.33)	11.67 (2.43)	105.8 (13.2)*	111.4 (13.2)*	3	Parental report	-Difference in V (not A) -Unisensory differences not included in model.	None
Feldman, Tu, et al. (2022)	McGurk	20	20	9.71 (2.37)	9.98 (2.25)	103.5 (12.8)*	119.2 (9.3)*	3	Screening at entry	No significant differences	None
Feng, Lu, Wang, et al. (2021)	McGurk	26	26	6.07 (0.81)	5.76 (0.88)	101.5 (11.24)	104.54 (13.1)	4	None	None	-Eye tracking -Differences partially explain integration differences
Feng et al. (2023)	McGurk	30	30	5.65 (0.77)	5.73 (0.51)	111 (11.82)	111.3 (10.73)	3	None	None	-Eye tracking -Difference not factored into analyses
Foxe et al. (2015)	Speech in Noise	84	142	10.82 (2.86)	11.32 (3.4)	102.2 (17.66)	109.99 (14.06)	4	Tested	-Differences in both A/V -Difference in integration above unisensory differences	-Eye tracking -No significant differences
Iarocci et al. (2010)	McGurk	12	12	10.6 (2.8)	10.32 (3.3)	NR	NR	5	Parental report	-Differences in V (not A) -No differences in integration above lipreading	None
Irwin et al. (2011)	Speech in Noise	13	13	9.08 (NR)	9.16 (NR)	NR	NR	4	Parental report	-Difference in V (not A) -Unisensory differences not included in model.	-Eye-tracking -Non-fixation removed (more for autism)
Irwin et al. (2011)	McGurk	13	13	9.08 (NR)	9.16 (NR)	NR	NR	4	Parental report	Difference in V (not A) Unisensory differences not included in model.	-Eye tracking -Non-fixation removed (more for autism)
Irwin et al. (2023)	Phonemic Restoration	13	32	NR	NR	NR	NR	3	Screened at entry	-Difference in pixelated condition, reflecting A sensitivity in the absence of discernible visual articulatory information. -Unisensory differences not included in model.	None
Keane et al. (2010)	McGurk	9	9	30 (9)	30 (8)	107 (22)	114 (18)	4	All normal or corrected to normal except one with autism.	No significant differences	None
Mongillo et al. (2008)	McGurk	15	21	13.73 (NR)	13.44 (NR)	96.1 (21.8)	102.1 (19)	4	None	None	None
Saalasti et al. (2012)	McGurk	16	16	32 (10)	32 (8)	113 (14)	118 (9)	4	Self-report	No significant differences	-Eye tracking -No significant differences
Smith and Bennetto (2007)	Speech in Noise	18	19	15.84 (2.17)	16.08 (2.04)	108.1 (14.23)	112.37 (8.81)	4	Tested	-Differences in V (not A) -Difference in integration above unisensory differences	None
Stevenson et al. (2014)	McGurk	32	32	11.8 (3.2)	12.3 (2.3)	NR	NR	3	Self-report	No significant differences.	None
Stevenson et al. (2017)	Speech in Noise	25	37	10.8 (3.4)	11.1 (3.4)	96.4 (20)	NR	4	None	-Differences in both A/V -No difference in integration above unisensory differences (collapsed across phoneme and whole word conditions)	None
Stevenson et al. (2018)	McGurk	38	38	12.3 (3.1)	11.1 (2.7)	NR	NR	4	None	None	None

(continued on next page)

Table 1 (*continued*)

Study	Paradigm	N ASD	N TD	Age ASD	Age TD	IQ ASD	IQ TD	ROB Score	Hearing/vision check	Unisensory control	Visual attention control
Stevenson et al. (2018)	Speech in Noise	38	38	12.3 (3.1)	11.1 (2.7)	NR	NR	4	None	-Difference in V (not A) -No difference in integration above unisensory differences	None
Taylor, Isaac, and Milne (2010)	McGurk	24	30	12.61 (2.4)	11.79 (2.49)	NR	NR	6	None	-Difference in V (not A). -Difference in integration above unisensory differences	None

Mean (standard deviation) reported for age and IQ. | * indicates non-verbal measures of IQ. | ** mean and standard deviation pooled between autism and Asperger's groups | A = auditory, V= visual | NR: not reported

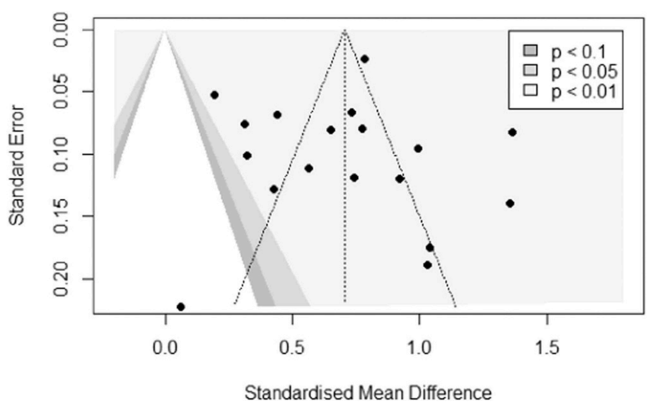

Fig. 2. Contour-Enhanced Funnel Plot of Identified Studies. Note: Statistical significance of studies is represented by the grey shaded regions, while the white shaded regions represent *p* values >.10.

hypothesis that small samples might be leading to an inflated estimate of between-group differences in MSI in autism, as outlined in Magnotti and Beauchamp (2018). While we did not find statistical evidence that sample size influenced the magnitude of the difference between groups directly, we did confirm that the samples throughout the literature were considerably smaller than their power analyses called for (average N=52.9, compared to 100–450), making it difficult to properly evaluate their claims with the existing data. However, contrasting with the assumptions of Magnotti and Beauchamp (2018), we did not find evidence for publication bias. So, although this meta-analysis does not rule out the possibility of overestimation of group differences, it also does not provide evidence in favor of their hypothesis.

Our results also spoke to the contradictory previous findings regarding the possible influence of age on differences between groups. The null findings of our meta-regression involving age differed from those of the meta-analyses by Feldman et al. (2018) (wherein larger differences between groups in MSI generally were found at younger ages) as well as Zhang et al. (2019) (which found the opposite with McGurk/MacDonald studies specifically). However, this may be because the vast majority of studies we included (16/18) focused on a narrow age range of children, and the small samples of the adult studies left us underpowered to detect an overall association with age. This possibility is tentatively supported by the difference found in our subgroup analysis comparing adult studies to those involving children, which found a larger difference between groups in the latter. However, with only two adult studies, it is crucial to temper the conclusions drawn from this analysis.

Still, it is worth noting that, against the grain of the prevailing trend with children, neither study involving adults (Saalasti et al., 2012; Keane et al., 2009) detected a difference between groups in the magnitude of the McGurk/MacDonald effect. Zhang et al. (2019) deviated from the conclusions of the original authors of the former by reporting a difference between groups and did not include the latter, perhaps explaining the discrepancy between their findings and ours (as well as their disagreement with those of Feldman et al., 2018). What is more, the two studies (Foxe et al., 2015; Taylor, Isaac, and Milne, 2010)

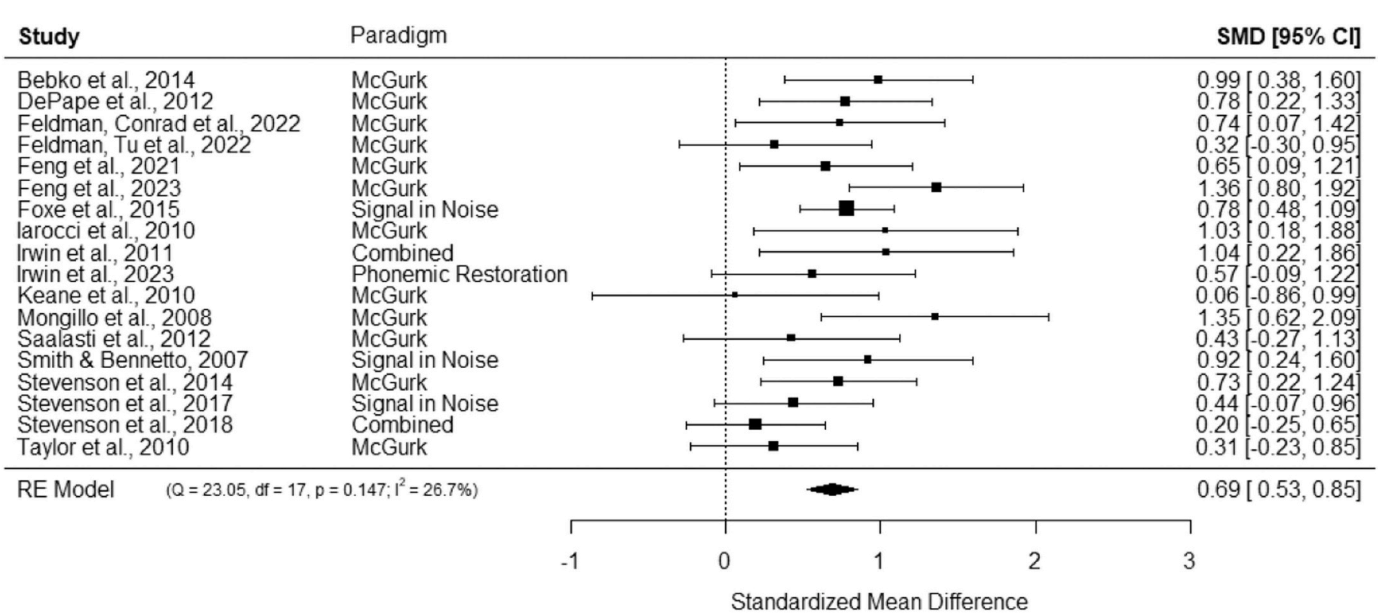

Fig. 3. Forest Plot of Identified Studies and Pooled Effect Size.

that divided children into age groups found that the differences detected in younger children were resolved in the eldest age group, which they took as evidence that autistic children catch up to their non-autistic peers in their ability to integrate the sights and sounds of language during adolescence. Together, the findings of our age analyses and patterns in the literature are, therefore, more in line with those of Feldman et al. (2018); however, the data are too limited for firm conclusions to be drawn about the influence of age on MSI differences between autistic and non-autistic individuals.

3.2. Strengths and limitations

A strength of our meta-analysis is its ability to focus on a specific area of audiovisual integration previous studies have suggested might be uniquely affected in autism. This allowed us to evaluate the sum of evidence regarding the influence of relevant variables outside the scope of prior reviews, namely the potential effects of sample size, paradigm type, and age across audiovisual speech integration paradigms. We were also able to include relevant studies released subsequently to any prior reviews (Feldman, Conrad, et al., 2022; Feldman, Tu, et al., 2022; Feng et al., 2023; Feng, Lu, Wang et al., 2021; Irwin et al., 2023) in this burgeoning field of research. Finally, our meta-analysis employed a risk of bias assessment. This convention is very common in meta-analyses focusing on clinical randomized controlled trials, but less standard in the fields of cognitive psychology and sensory perception. Including a risk of bias assessment is essential to mitigating the risk of decisions in sampling and design from individual studies distorting insight into the underlying differences between groups. Measuring its influence allowed us to provide evidence that differences in participant selection and comparability were not systematically influencing our results.

However, there are also a number of limitations of the current meta-analysis. We chose to only include studies with participants who had received an autism diagnosis. While this allowed us to synthesize the existing evidence regarding integration of speech stimuli in a clinical group, it did not allow insight into potential differences across the broader autism phenotype and beyond DSM classifications. Additionally, our risk of bias assessment was not ideally tailored to the type of experiments included. The Newcastle-Ottawa Scale is designed primarily for case/control studies investigating the influence of factors to which individuals were exposed on their likelihood to develop a given condition. In contrast, our studies compared cases to controls in performance on tasks. This meant that we needed to adapt the tool by omitting the exposure section, rendering us unable to glean as much information about specific studies as would have been possible if there was a better fit between our risk of bias assessment tool and the designs of the experiments included.

Regarding subgroup analyses, the high degree of variability across study types and limited number of studies in certain categories (e.g., adult samples and enhancement paradigms) meant that they were hamstrung in their ability to detect potential differences. This issue was compounded by small sample sizes across designs resulting in low statistical power overall. In the case of the age analyses, study-level proxies (i.e. the relationship between mean age and effect size) lack resolution that may be important to fully exploring the influence of age on integration. For example, Foxe et al. (2015) and Taylor et al. (2010) show differences for particular age subgroups but not others, nuances that are lost when collapsing into a single mean age and pooled effect size per study.

Furthermore, many studies differed in their approach to scoring, analyzing, and reporting data, and several were missing the necessary data for effect size calculation. For example, two studies (Feng, Lu, Fang, et al., 2021; Woynaroski et al., 2013) had to be excluded from our analyses due to their decision to leave out participants who did not experience the McGurk/MacDonald illusion. The rate at which this occurred differed between groups, and other studies did not adopt this convention, because it risks discarding the very difference of interest. Another lacked demographic information for the relevant participants (Irwin et al., 2023), meaning it could not be included in the age meta-regression, and the balance of age/IQ between groups is unknown. Additionally, most McGurk/MacDonald studies did not control for unisensory performance. For the sake of consistency in the type of data pooled, we took the raw audiovisual scores for all McGurk/MacDonald studies. This meant sacrificing the efforts of several researchers to account for unisensory performance in their models and better isolate audiovisual integration. Finally, not all studies reported each syllable response. Several only reported fusion rates, despite visual capture also reflecting a form of MSI, and differences between groups being known to vary according to which response types are compared (Zhang et al., 2019). We extracted the rate of phoneme consistent responses, when possible, because it reflects the inverse of both types of integration, and therefore the full extent of differences in MSI. However, this was not possible for studies that only reported fusion rates.

3.3. Conclusions/implications

This meta-analysis has revealed one very clear trend: autistic individuals, at least as children, show markedly different audiovisual processing profiles. They are characterized by an attenuated benefit of integration in speech in noise paradigms and reduced susceptibility to the McGurk/MacDonald effect. This pattern of findings supports the premise of sensory first theories, at least at early developmental stages, and implicates MSI as a key problem area for autistic children. The findings of Feldman, Conrad, et al. (2022) in particular provide evidence for this line of thinking by linking susceptibility to the McGurk/MacDonald effect to "core" autistic traits such as social communication skills in both autistic and non-autistic individuals. This being said, surveying the literature on this topic has revealed considerable diversity in experimental design, data analysis, and reporting. This heterogeneity in approaches not only sheds some light on the heterogeneity of findings; it also makes it difficult to ascertain the degree to which numerous confounding variables may have played a role.

From a design perspective, a clear difference among studies was the experimenters' approaches to accounting for unisensory processing abilities. First of all, it is problematic that several studies did not report any efforts to test for normal (or corrected-to-normal) hearing/vision across cohorts (see Table 1), given that autistic individuals suffer a higher occurrence rate of several visual processing disorders (Reynolds and Culican, 2023) as well as hearing impairment (Demopoulos and Lewine, 2016). As such, there is the possibility that in some studies, differences in sight and hearing might contribute to ostensible deficits in MSI that are better explained on the unisensory level. That being said, most studies did at least use self-report to screen for vision/hearing problems and still tended to find differences between groups (see Table 1).

Beyond basic vision and hearing impairments, some studies detected differences between groups in the ability to recognize phonemes/visemes in auditory/visual-only trials, as well as differences in attention to the stimuli (using eye tracking). Both unisensory reliability and attention are established contributors to MSI (Alais and Burr, 2004; Hirst et al., 2018; Talsma et al., 2010; Van der Burg et al., 2012) that were shown to predict audiovisual performance in some studies included, but were not measured in many of them (see Table 1 for an overview). In one study (Iarocci et al., 2010), the difference between groups in MSI disappeared once controlling for a disparity in lipreading ability between groups, although most experiments found group effects on integration over and above unisensory or attentional differences. Accordingly, despite the room for these potential confounds to play a role in the findings, the general agreement across paradigms still provides evidence for a difference in MSI that is not entirely explained by lower-level differences in unisensory perception or attention, at least among children. This conclusion is most strongly supported by the convergence of evidence showing group effects even after controlling for at least one of these factors. To better isolate differences in MSI, it is crucial that future research adheres to this convention and controls for underlying unisensory and attentional differences.

In addition to our primary findings, our subgroup analyses and meta-regressions also have broad implications and provide greater insight into the nuances of the differences between groups in MSI. Firstly, the non-significant difference in group effect magnitude between enhancement (speech in noise) and distortion (McGurk/MacDonald) paradigms suggests that both approaches may be similarly sensitive to differences between autistic and non-autistic individuals when it comes to MSI. However, it should be noted that speech in noise experiments are much less common in the literature, with the paradigm being employed in only 5/18 studies included. When one considers the greater ecological validity they offer, as well as the inconsistency in findings seen with the McGurk/MacDonald effect (even when using the same stimuli, see Getz and Toscano, 2021), it seems clear that further research utilizing the speech in noise paradigm is warranted. However, the McGurk/MacDonald effect remains a simple and effective alternative to studying these differences, as is shown by the agreement of the results the two paradigms produce.

In addition to paradigm type, we also found no influence of sample size on the magnitude of differences between groups. However, smaller samples naturally tend to result in greater variability, and a poorer estimation of the true effect size. The moderate degree of heterogeneity we detected across studies ($I^2 = 26.66$ %, 95 % PI [0.28, 1.10]) may be an indication of this imprecision, given the generally small samples. While it is likely that the aforementioned methodological inconsistency also plays a role here, recruitment of larger samples is a clear way to address one possible source of the heterogeneity in outcomes.

Finally, our two-pronged investigation of the influence of age on differences in MSI between groups provides some greater insight into the issue, but certainly no resolution. While the results of our sub-group analysis and null findings of the adult studies included (Keane et al., 2010; Saalasti et al., 2012) weigh in favor of the hypothesis that the differences between autistic and non-autistic children in MSI are largely resolved by adulthood (see Foxe et al., 2015; Taylor et al., 2010), the volume of data available (combined N=50 for adults vs. 902 for child studies) is insufficient to draw firm conclusions. The non-significant influence of mean age in our meta-regression emphasizes the importance of tempering conclusions from this analysis. Accordingly, while this meta-analysis does not provide strong evidence that age plays a significant role in the differences studied, it certainly motivates more extensive research with older samples. When a greater volume of data investigating audiovisual speech processing in autistic and non-autistic adults becomes available, this would be an ideal topic for an individual participant data meta-analysis. This would provide greater sensitivity to differences that may be lost when using group means (as discussed in the limitations section) and speak to the contrasting results of prior meta-analyses.

In conclusion, while the sum of the evidence suggests that differences exist in the ability of autistic and non-autistic individuals to integrate the sights and sounds of speech, questions remain as to the degree to which unisensory/attentional factors explain them, as well as the influence of age upon them. We propose three key suggestions for future research to address these remaining questions and inconsistencies: further research into adults, larger samples, and greater standardization of experimental methods and analyses (especially in controlling for confounding factors). On the former note, adults make up the vast majority of those with autism, but a small minority of the research focus. To understand the degree to which multisensory differences persist throughout the lifetime, more extensive research with adults is necessary. This is crucial to determining when interventions involving multisensory training may be beneficial to autistic individuals. Regarding the latter two points, while the diversity in findings with regard to autism is often attributed to the heterogeneity of the autism spectrum, the lack of standardization in procedures and the room for variability with small samples surely exacerbate this issue. Greater methodological consistency and larger samples should produce more similar findings and, with them, a clearer understanding of autism.

Author note

The authors have no conflict of interest to declare. Funding provided by ZonMw (60-63600-98-834).

Data Availability

The extracted effect size data of the included studies as well as the R code of the analyses are available at http://osf.io/fvw4t/?view_only=3fab3d0612824a198716a51ec1c574c2.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.neubiorev.2024.105787.

References

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

Alcántara, J.I., Weisblatt, E., Moore, B., Bolton, P., 2004. Speech-in-noise perception in high-functioning individuals with autism or Asperger's syndrome. J. Child Psychol. Psychiatry Allied Discip. *45*, 1107–1114. https://doi.org/10.1111/j.1469-7610.2004.t01-1-00303.x.

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders: DSM-5 (5th Edition). https://doi.org/10.1176/appi.books.9780890425596.

Balduzzi, S., Rücker, G., Schwarzer, G., 2019. How to perform a meta-analysis with R: a practical tutorial. Evid. Based Ment. Health *22* (4), 153–160. https://doi.org/10.1136/ebmental-2019-300117.

Bebko, J.M., Schroeder, J.H., Weiss, J.A., 2014. The McGurk effect in children with autism and asperger syndrome. Autism Res. *7* (1), 50–59. https://doi.org/10.1002/aur.1343.

Brignell, A., Morgan, A.T., Woolfenden, S., Klopper, F., May, T., Sarkozy, V., Williams, K., 2018. A systematic review and meta-analysis of the prognosis of language outcomes for individuals with autism spectrum disorder. Autism Dev. Lang. Impair. *3*, 2396941518767610 https://doi.org/10.1177/2396941518767610.

Brock, J., Brown, C.C., Boucher, J., Rippon, G., 2002. The temporal binding deficit hypothesis of autism. Dev. Psychopathol. *14* (2), 209–224. https://doi.org/10.1017/S0954579402002018.

Buijsman, R., Begeer, S., Scheeren, A.M., 2023. Autistic person' or 'person with autism'? Person-first language preference in Dutch adults with autism and parents. Autism.: Int. J. Res. Pract. *27* (3), 788–795. https://doi.org/10.1177/13623613221117914.

Cascio, C.J., Woynaroski, T., Baranek, G.T., Wallace, M.T., 2016. Toward an interdisciplinary approach to understanding sensory function in autism spectrum disorder. Autism Res.: Off. J. Int. Soc. Autism Res. *9* (9), 920–925. https://doi.org/10.1002/aur.1612.

Cooper, H., Hedges, L., Valentine, J., 2009. The handbook of research synthesis and meta-analysis. Second Ed.) Vol. 113.

Demopoulos, C., Lewine, J.D., 2016. Audiometric profiles in Autism Spectrum Disorders: does subclinical hearing loss impact communication? Autism Res.: Off. J. Int. Soc. Autism Res. *9* (1), 107–120. https://doi.org/10.1002/aur.1495.

DePape, A.-M.R., Hall, G.B.C., Tillmann, B., Trainor, L.J., 2012. Auditory processing in high-functioning adolescents with Autism Spectrum Disorder. PloS One *7* (9), e44084. https://doi.org/10.1371/journal.pone.0044084.

van Ee, R., van Boxtel, J.J.A., Parker, A.L., Alais, D., 2009. Multisensory congruency as a mechanism for attentional control over perceptual selection. J. Neurosci. *29* (37), 11641–11649. https://doi.org/10.1523/JNEUROSCI.0873-09.2009.

Erber, N.P., 1969. Interaction of audition and vision in the recognition of oral speech stimuli. J. Speech Hear. Res. *12* (2), 423–425.

Fadeev, K.A., Goyaeva, D.E., Obukhova, T.S., Ovsyannikova, T.M., Shvedovskiy, E.F., Nikolaeva, A.Y., Davydova, E.Y., Stroganova, T.A., Orekhova, E.V., 2023. Difficulty with speech perception in the background of noise in children with autism spectrum disorders is not related to their level of intelligence. Clin. Psychol. Spec. Educ. *12* (1), 180–212. https://doi.org/10.17759/cpse.2023120108.

Falck-Ytter, T., Bussu, G., 2023. The sensory-first account of autism. Neurosci. Biobehav. Rev. *153*, 105405 https://doi.org/10.1016/j.neubiorev.2023.105405.

Feldman, J.I., Dunham, K., Cassidy, M., Wallace, M.T., Liu, Y., Woynaroski, T.G., 2018. Audiovisual multisensory integration in individuals with autism spectrum disorder: a systematic review and meta-analysis. Neurosci. Biobehav. Rev. *95*, 220–234. https://doi.org/10.1016/j.neubiorev.2018.09.020.

Feldman, J.I., Conrad, J.G., Kuang, W., Tu, A., Liu, Y., Simon, D.M., Wallace, M.T., Woynaroski, T.G., 2022. Relations between the McGurk effect, social and communication skill, and autistic features in children with and without autism.

J. Autism Dev. Disord. *52* (5), 1920–1928. https://doi.org/10.1007/s10803-021-05074-w.

Feldman, J.I., Tu, A., Conrad, J.G., Kuang, W., Santapuram, P., Woynaroski, T.G., 2022. The impact of singing on visual and multisensory speech perception in children on the autism spectrum. Multisens. Res. *36* (1), 57–74. https://doi.org/10.1163/22134808-bja10087.

Feng, S., Lu, H., Wang, Q., Li, T., Fang, J., Chen, L., Yi, L., 2021. Face-viewing patterns predict audiovisual speech integration in autistic children. Autism Res.: Off. J. Int. Soc. Autism Res. *14* (12), 2592–2602. https://doi.org/10.1002/aur.2598.

Feng, S., Lu, H., Fang, J., Li, X., Yi, L., Chen, L., 2021. Audiovisual speech perception and its relation with temporal processing in children with and without autism. Read. Writ.: Interdiscip. J. psyh. https://doi.org/10.1007/s11145-021-10200-2.

Feng, S., Wang, Q., Hu, Y., Lu, H., Li, T., Song, C., Fang, J., Chen, L., Yi, L., 2023. Increasing audiovisual speech integration in autism through enhanced attention to mouth. Dev. Sci. *26* (4), e13348 https://doi.org/10.1111/desc.13348.

Foxe, J.J., Molholm, S., Del Bene, V.A., Frey, H.-P., Russo, N.N., Blanco, D., Saint-Amour, D., Ross, L.A., 2015. Severe multisensory speech integration deficits in high-functioning school-aged children with Autism Spectrum Disorder (ASD) and their resolution during early adolescence. Cereb. Cortex (N. Y., N. Y.: 1991) *25* (2), 298–312. https://doi.org/10.1093/cercor/bht213.

Getz, L.M., Toscano, J.C., 2021. Rethinking the McGurk effect as a perceptual illusion. Atten. Percept. Psychophys. *83* (6), 2583–2598. https://doi.org/10.3758/s13414-021-02265-6.

Hirst, R.J., Stacey, J.E., Cragg, L., Stacey, P.C., Allen, H.A., 2018. The threshold for the McGurk effect in audio-visual noise decreases with development. Article 1. Sci. Rep. *8* (1) https://doi.org/10.1038/s41598-018-30798-8.

Iarocci, G., Rombough, A., Yager, J., Weeks, D.J., Chua, R., 2010. Visual influences on speech perception in children with autism. Autism.: Int. J. Res. Pract. *14* (4), 305–320. https://doi.org/10.1177/1362361309353615.

Irwin, J., DiBlasi, L., 2017. Audiovisual speech perception: a new approach and implications for clinical populations. Lang. Linguist. Compass *11* (3), 77–91.

Irwin, J., Harwood, V., Kleinman, D., Baron, A., Avery, T., Turcios, J., Landi, N., 2023. Neural and behavioral differences in speech perception for children with autism spectrum disorders within an audiovisual context. J. Speech, Lang. Hear. Res.: JSLHR *66* (7), 2390–2403. https://doi.org/10.1044/2023_JSLHR-22-00661.

Irwin, J., Tornatore, L.A., Brancazio, L., Whalen, D.H., 2011. Can children with autism spectrum disorders 'Hear' a speaking face? Child Dev. 82 (5), 1397–1403. https://doi.org/10.1111/j.1467-8624.2011.01619.x.

Keane, B.P., Rosenthal, O., Chun, N.H., Shams, L., 2010. Audiovisual integration in high functioning adults with autism. Res. Autism Spectr. Disord. *4* (2), 276–289. https://doi.org/10.1016/j.rasd.2009.09.015.

Lipsey, M.W., Wilson, D.B., 1993. The efficacy of psychological, educational, and behavioral treatment: confirmation from meta-analysis. Am. Psychol. *48* (12), 1181–1209. https://doi.org/10.1037/0003-066X.48.12.1181.

MacLeod, A., Summerfield, Q., 1987. Quantifying the contribution of vision to speech perception in noise. Br. J. Audiol. *21* (2), 131–141. https://doi.org/10.3109/03005368709077786.

Magnotti, J.F., Beauchamp, M.S., 2018. Published estimates of group differences in multisensory integration are inflated. PLOS ONE *13* (9), e0202908. https://doi.org/10.1371/journal.pone.0202908.

Mamashli, F., Khan, S., Bharadwaj, H., Michmizos, K., Ganesan, S., Garel, K.-L.A., Hashmi, J.A., Herbert, M.R., Hämäläinen, M., Kenet, T., 2017. Auditory processing in noise is associated with complex patterns of disrupted functional connectivity in autism spectrum disorder. Autism Res.: Off. J. Int. Soc. Autism Res. *10* (4), 631–647. https://doi.org/10.1002/aur.1714.

Mason, G.M., Goldstein, M.H., Schwade, J.A., 2019. The role of multisensory development in early language learning. J. Exp. Child Psychol. *183*, 48–64. https://doi.org/10.1016/j.jecp.2018.12.011.

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

Mody, M., Belliveau, J.W., 2013. Speech and language impairments in autism: insights from behavior and neuroimaging. North Am. J. Med. Sci. *5* (3), 157–161.

Mongillo, E., Irwin, J., Whalen, D., Klaiman, C., Carter, A., Schultz, R., 2008. Audiovisual processing in children with and without autism spectrum disorders. J. Autism Dev. Disord. *38* (7), 1349–1358. https://doi.org/10.1007/s10803-007-0521-y.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ *372*, n71. https://doi.org/10.1136/bmj.n71.

Reynolds, M., Culican, S.M., 2023. Visual Autism. Children *10* (4), 606. https://doi.org/10.3390/children10040606.

Robertson, C.E., Baron-Cohen, S., 2017. Sensory perception in autism. Article 11. Nat. Rev. Neurosci. *18* (11) https://doi.org/10.1038/nrn.2017.112.

Saalasti, S., Kätsyri, J., Tiippana, K., Laine-Hernandez, M., von Wendt, L., Sams, M., 2012. Audiovisual speech perception and eye gaze behavior of adults with asperger syndrome. J. Autism Dev. Disord. *42* (8), 1606–1615. https://doi.org/10.1007/s10803-011-1400-0.

Setti, A., Stapleton, J., Leahy, D., Walsh, C., Kenny, R.A., Newell, F.N., 2014. Improving the efficiency of multisensory integration in older adults: audio-visual temporal discrimination training reduces susceptibility to the sound-induced flash illusion. Neuropsychologia *61*, 259–268. https://doi.org/10.1016/j.neuropsychologia.2014.06.027.

Smith, E.G., Bennetto, L., 2007. Audiovisual speech integration and lipreading in autism. J. Child Psychol. Psychiatry, Allied Discip. *48* (8), 813–821. https://doi.org/10.1111/j.1469-7610.2007.01766.x.

Stevenson, R.A., Wilson, M.M., Powers, A.R., Wallace, M.T., 2013. The effects of visual training on multisensory temporal processing. Exp. Brain Res. *225* (4), 479–489. https://doi.org/10.1007/s00221-012-3387-y.

Stevenson, R.A., Siemann, J.K., Schneider, B.C., Eberly, H.E., Woynaroski, T.G., Camarata, S.M., Wallace, M.T., 2014. Multisensory temporal integration in autism spectrum disorders. J. Neurosci. *34* (3), 691–697. https://doi.org/10.1523/JNEUROSCI.3615-13.2014.

Stevenson, R.A., Baum, S.H., Segers, M., Ferber, S., Barense, M.D., Wallace, M.T., 2017. Multisensory speech perception in autism spectrum disorder: from phoneme to whole-word perception. Autism Res.: Off. J. Int. Soc. Autism Res. *10* (7), 1280–1290. https://doi.org/10.1002/aur.1776.

Stevenson, R.A., Segers, M., Ncube, B.L., Black, K.R., Bebko, J.M., Ferber, S., Barense, M. D., 2018. The cascading influence of multisensory processing on speech perception in autism. Autism.: Int. J. Res. Pract. *22* (5), 609–624. https://doi.org/10.1177/1362361317704413.

Sumby, W.H., Pollack, I., 1954. Visual contribution to speech intelligibility in noise. J. Acoust. Soc. Am. *26* (2), 212–215. https://doi.org/10.1121/1.1907309.

Talsma, D., Senkowski, D., Soto-Faraco, S., Woldorff, M.G., 2010. The multifaceted interplay between attention and multisensory integration. Trends Cogn. Sci. *14* (9), 400–410. https://doi.org/10.1016/j.tics.2010.06.008.

Taylor, N., Isaac, C., Milne, E., 2010. A comparison of the development of audiovisual integration in children with autism spectrum disorders and typically developing children. J. Autism Dev. Disord. *40* (11), 1403–1411. https://doi.org/10.1007/s10803-010-1000-4.

Van der Burg, E., Olivers, C.N.L., Bronkhorst, A.W., Theeuwes, J., 2008. Pip and pop: nonspatial auditory signals improve spatial visual search. J. Exp. Psychol.: Hum. Percept. Perform. *34* (5), 1053–1065. https://doi.org/10.1037/0096-1523.34.5.1053.

Van der Burg, E., Olivers, C.N.L., Theeuwes, J., 2012. The attentional window modulates capture by audiovisual events. PloS One *7* (7), e39137. https://doi.org/10.1371/journal.pone.0039137.

Van Engen, K.J., Xie, Z., Chandrasekaran, B., 2017. Audiovisual sentence recognition not predicted by susceptibility to the McGurk effect. Atten. Percept. Psychophys. *79* (2), 396–403. https://doi.org/10.3758/s13414-016-1238-9.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. *36*, 1–48. https://doi.org/10.18637/jss.v036.i03.

Vroomen, J., de Gelder, B., 2000. Sound enhances visual perception: cross-modal effects of auditory organization on vision. J. Exp. Psychol.: Hum. Percept. Perform. *26* (5), 1583–1590.

Wells, G., Wells, G., Shea, B., Shea, B., O'Connell, D., Peterson, J., Welch, Losos, M., Tugwell, P., Ga, S.W., Zello, G., Petersen, J., 2014. The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses. https://www.semanticscholar.org/paper/The-Newcastle-Ottawa-Scale-(NOS)-for-Assessing-the-Wells-Wells/c293fb316b6176154c3fdbb8340a107d9c8c82bf.

Williams, J.H.G., Massaro, D.W., Peel, N.J., Bosseler, A., Suddendorf, T., 2004. Visual-auditory integration during speech imitation in autism. Res. Dev. Disabil. *25* (6), 559–575. https://doi.org/10.1016/j.ridd.2004.01.008.

Woynaroski, T.G., Kwakye, L.D., Foss-Feig, J.H., Stevenson, R.A., Stone, W.L., Wallace, M.T., 2013. Multisensory speech perception in children with autism spectrum disorders (psyh.). J. Autism Dev. Disord. *43* (12), 2891–2902. https://doi.org/10.1007/s10803-013-1836-5.

Zhang, J., Meng, Y., He, J., Xiang, Y., Wu, C., Wang, S., Yuan, Z., 2019. McGurk effect by individuals with autism spectrum disorder and typically developing controls: a systematic review and meta-analysis (psyh.). J. Autism Dev. Disord. *49* (1), 34–43. https://doi.org/10.1007/s10803-018-3680-0.