Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Multimodal speech perception in noise after multisensory perceptual training in autistic children and adolescents

Laura Möde

moede.laura@mh-hannover.de

Hannover Medical School

Erfan Ghaneirad

Hannover Medical School

Gregor R. Szycik

Hannover Medical School

Hans Worthmann

Hannover Medical School

Stefan Bleich

Hannover Medical School

Anna Borgolte

Hannover Medical School

Article

Keywords: Multimodal integration, autism spectrum disorder, audiovisual training, speech in noise, simultaneity judgement, temporal binding window

Posted Date: March 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9010946/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on October 12th, 2026. See the published version at https://doi.org/10.1038/s41598-026-75297-3.

Abstract

Autistic children and adolescents exhibit alterations in multisensory processing of simple audiovisual stimuli and complex speech stimuli. Multisensory perceptual training has been demonstrated to enhance audiovisual simultaneity judgment in typically developed (TD) adults, thereby facilitating speech comprehension in noise. In this study, we hypothesized that multisensory training enhances audiovisual integration and speech comprehension in noise in autistic children and adolescents. A brief simultaneity judgment task (SJT) with feedback was selected as the training paradigm and administered to 19 autistic participants (aged 8–17). Simple, audiovisual integration and speech comprehension were assessed before training (T0), immediately after training (T1), seven (T2) and 14 days (T3) after training. For simple audiovisual stimuli, simultaneity perception was shown to be more precise immediately after multisensory training with the effects lasting over a period of seven days. In a Word Recognition Task (WRT), no significant improvements were demonstrated following the training. This study suggests that multisensory perceptual training enhances simultaneity perception in autistic youth with the effects lasting at least seven days. Subsequent studies must elucidate which variables moderate training outcomes in order to determine how training paradigms should be designed to enhance speech comprehension and increase training effects in autistic individuals.

Introduction

In order to adequately perceive and interpret environmental cues, information from different sensory modalities must be combined via multimodal integration (MI). This process enables the perception of a coherent, holistic environment[1, 2]. In comparison with unimodal processing, MI facilitates faster and more accurate perception, localization, and discrimination by combining complementary information and integrating redundant information [2]. MI is therefore a fundamental component of our sensory system, providing the foundation for development and maintenance of intact higher perceptual, cognitive, motor, and learning processes [1, 3, 4].

This is among other factors, of central importance for language comprehension and segmentation, as the detection, recognition and processing of linguistic stimuli involves the combination and integration of auditory (e.g., content and prosody of voice, background noises) and visual stimuli (e.g., lip movements and facial expressions) [5]. MI is essential for the resolution of ambiguous situations [6, 7]. In situations where acoustic information is deprived, the visual information plays a pivotal role in resolving ambiguities, thereby facilitating speech comprehension and communicative abilities [8, 9]. MI warrants correct perception even in the presence of multiple concurrent conversations and noises that interfere with each other in social situations. Which stimuli are integrated into one unified percept is largely determined by the temporal proximity of the corresponding stimuli[10]. Research indicates that stimuli presented in close temporal proximity are more likely to be integrated into a single perception, thereby increasing neural and behavioral responses [1].

The temporal binding window (TBW), which is defined as the temporal tolerance range in which stimuli are highly likely to be perceived as simultaneous and, consequently, integrated into a unified percept [6], is a critical measure in this context reflecting the accuracy and strength of MI [11]. This tolerance range is imperative for the adequate perception of our environment and offers ecological advantages, as stimuli from different sensory modalities are transmitted at varying velocities, resulting in the arrival at the sensory organs with a temporal lag [12]. Consequently, the point of subjective simultaneity, defined as the stimulus onset asynchrony (SOA) at which the majority of individuals perceive two stimuli as simultaneous, deviates from zero. Instead, compelling research shows that individuals predominantly perceive two stimuli as simultaneous when the visual stimulus slightly precedes the auditory stimulus for review, see 13]. The TBW's size is highly dynamic and varies depending on factors such as the characteristic of the stimuli, requirements of the task and age of the studied subjects [6, 14, 15]. The TBW is smallest for simple audiovisual, non-verbal stimuli (e.g., beeps combined with flashes of light) and significantly larger for complex audiovisual stimuli like speech[16].

The TBW is subject to developmental processes and changes over the course of a lifetime following a U-shaped trajectory [1]. As part of multisensory perpetual narrowing, its size decreases, allowing for more complex, selective, and reliable integration of multisensory stimuli [1, 14]. At birth, the ability to integrate stimuli from multiple modalities is limited to a rudimentary and immature tendency, The refinement and specialization of MI abilities during development is contingent upon experience and ecological adaption to environmental conditions, which is also reflected in a gradual narrowing of the TBW [4]. This developmental process commences during infancy, continues throughout middle childhood and adolescence, which represent significant transitional phases [1, 14]. In children, intact multisensory processing has been demonstrated to be associated with in a variety of domains. These include speech, reading, and executive function, as well as socially relevant skills such as empathy, social reciprocity, and cognition, emotional recognition, and regulation [4]. In elderly adults, there is a decline in MI abilities that is concomitant with a widening of the TBW [15] and results in difficulties in speech perception, especially under noisy conditions [11, 17].

Alterations in multisensory perception have been reported in individuals diagnosed with schizophrenia or autism [18, 19]. MI deficits have been shown to contribute to less precise perception. This is due to an increased likelihood of incorrectly processing stimuli from different sources together [19], heightened cognitive demands of stimulus processing [6], and an increased risk of experiencing the environment as inadequately filtered and thus overwhelmingly stimulating [20]. However, the majority of research on autism has historically centered on aspects of social interaction and communication, as well as restrictive and repetitive behaviors. In clinical practice and scientific research, sensory aspects played a subordinate role and have been rather considered as a side effect of social difficulties [21]. However, recent findings suggest that deviations in sensory processing may already be evident in infancy and can predict the subsequent development of autistic symptoms or later autism diagnoses [22, 23]. It is surprising that MI has received so little attention in this context, despite the fact that various authors

have argued that multisensory processing appears to play a central role in the development and maintenance of complex social and cognitive abilities. This assertion suggests that a deeper understanding of MI in autism may be central to elucidating its development and manifestation [18, 24, 25]. Baum and colleagues (2015) further posit that the central role of MI alterations and popular theories on autism are not mutually exclusive. Rather, they contend that MI deficits represent an important addition to established approaches. The present findings have been summarized in the "sensory-first account of autism," which postulates that changes in sensory processing are a critical determinant of autism. This is due to the fact that early deviations in sensory processing have cascading effects on the development of higher perceptual, cognitive, motor, and social communication skills [21, 25, 26].

A number of studies have indicated a potential correlation between challenges related to social interaction and communicative skills in individuals with autism and alterations in language processing and comprehension that can be attributed to deficits in audiovisual integration [24, 27]. MI alterations are also evident regarding the TBW, which is enlarged in autistic children and adolescents compared to typically developing (TD) children and adolescents [28–31]. In addition, prior studies indicate a link between TBW widening and symptom severity in autism support [32–34]. Alterations in multisensory processing, particularly involving audiovisual stimuli, appear to be more pronounced in younger autistic individuals. The multisensory temporal narrowing observed in TD children and adolescents is not equally evident in autistic children or adolescents but instead seems to occur only in adulthood [33, 35–37]. As highlighted in a recent review by Ampollini and colleagues (2024), the question of whether multisensory perpetual narrowing in autistic individuals is simply delayed, follows different developmental pathways, or is explicable by the formation of compensatory mechanisms during adolescence remains unresolved.

In TD adults, research has demonstrated that multisensory perceptual training can lead to a reduction in TBW size within a single, brief training session. The training effects persist for a duration of seven days, thereby enhancing not only multisensory temporal acuity but also word recognition in noise [15, 38–40]. Further, it has been observed that the magnitude of the TBW prior to training is associated with the extent of the subsequent reduction in TBW size, leading to improved audiovisual integration [38, 41]. This observation underscores the potential of multisensory training as a promising intervention to enhance language comprehension [40]. Given that autistic children and adolescents exhibit an enlarged TBW [28–31], multisensory training appears to be particularly promising intervention for these participants. Some authors even argue that improving MI could serve as a significant component in mitigating the communication and social challenges associated with autism [26, 33]. Only two previous studies have examined the effects of multisensory perceptual training in a small sample of autistic children and adolescents, providing inconclusive results and highlighting that to date no research has been conducted on whether multisensory perceptual training, can improve speech comprehension in noise for this specific study population[42, 43].

The current study aims to address this research gap. To date, the efficacy of multisensory training in enhancing audiovisual integration and, consequently, language comprehension in autistic children and

adolescents remains uncertain. The question of whether multisensory training could serve as a starting point for developing interventions in the context of cognitive remediation, possibly helping to compensate for a potential delay in MI development in autistic individuals, remains unresolved. Therefore, in the present study audiovisual training was conducted to reduce the audiovisual tolerance range of multisensory synchrony perception. Further, it was examined whether multisensory training has the capacity to improve the integration of complex speech stimuli in autistic children and adolescents. In accordance with the extant research, it was hypothesized that improvements would be observed in audiovisual simultaneity perception and that a decrease would be noted in TBW width in a Simultaneity Judgment Task (SJT) as a result of the completion of a brief, multisensory perceptual training consisting of a modified SJT with feedback. In addition, it was hypothesized that there would be an enhancement of speech comprehension subsequent to the training as indicated by higher word recognition scores in a Word Recognition Task (WRT). It was hypothesized that the training effects would visible immediately apparent after the training (T0) and be sustained for a period of seven days (T2) compared to their initial state (T0) but would revert to their initial state after 14 days (T3).

Results

A total of 19 autistic children and adolescents (M_{age} = 12.4 years, *SD* = 2.8, range = 8–17; females = 4) participated in the study. Within the study, German translations of the Social Communication Questionnaire (SCQ) [44] ($SCQ_{current}$: *M* = 15.8, *SD* = 5.2; $SCQ_{lifetime}$: *M* = 18.8, *SD* = 7.0; $Cut\text{-}off_{Autism\ spectrum}$ ≥15) as well as the children and adolescent versions of the Autism Spectrum Quotient (AQ-Child; AQ-Adol) [45, 46] (AQ-Child: *M* = 93.7, *SD* = 6.8, *Cut-off* : ≥76; AQ-Adol: *M* = 29.4, *SD* = 8.7; *Cut-off* ≥ 30) were administered to briefly assess current symptom severity. The mean value AQ-Adol is marginally below the cut-off value, a finding that is most likely due to the presence of two outliers with individual AQ-Adol sum scores of 13 and 17. These outliers may have resulted from a limited capacity for introspection. Since both participants also had a medically confirmed diagnosis of autism in accordance with the currently valid gold standard, which could be confirmed by clinical observation, the participants were nevertheless included in the study. The German version of the vocabulary intelligence subtest of Cattell's Fluid Intelligence Test, Scale 2 (CFT 20-R) [47] was conducted to obtain an overview of the participants' vocabulary (*M* = 103, *SD* = 12.0).

- Simultaneity Judgement Task (SJT)

We calculated the TBW for each participant by fitting a Gaussian distribution to the data derived from the SJT. Since the goodness of fit varied greatly between participants (adj. R^2 ranging from − 0.3 to 0.97), we did not use individual TBWs as an outcome measure for the training effect, but SOA conditions as a performance measure. For this reason, TBWs were only reported at group level where goodness of fit was satisfactory (mean adj. R^2_{T0} =0.83, mean adj. R^2_{T1} = 0.96, mean adj. R^2_{T2} =0.98 and mean adj. R^2_{T3} =0.93). In accordance with the 70% criterion, TBW was 100.9 ms at T0, 77.6 ms at T2, and 96.7 ms at T3. At T1, the 70% criterion was not met, so it was not possible to determine the TBW (Fig. 1). Descriptively,

the curve fit demonstrated a reduction in TBW seven days after the training and a widening of TBW after 14 days.

The linear mixed-effects model for SJT revealed a significant main effect of point of measurement (POM) ($F(3, 586.9) = 6.19$, $p < 0.01$) and SOA ($F(10, 586.9) = 44.81$, $p < 0.001$), but no significant POM x SOA interaction ($F(30, 586.9) = 1.23$, $p = 0.19$) (Fig. 2). Relative to baseline (T0) Tukey-adjusted pairwise comparisons revealed significant differences compared to T1 (Δ = -0.076, *95% CI* [-0.125, -0.028], $p < 0.01$) as well as T2 (Δ = -0.068, *95% CI* [-0.116, -0.020], $p < 0.05$) but no significant differences between T0 and T3 ($p = 0.36$). The post-hoc tests indicated no significant differences between the remaining combinations of POM (all $p > 0.05$). Regarding SOA conditions, it was demonstrated that shorter SOAs (0-100 ms) did not differ significantly from each other (all $p > 0.05$). Significant differences were first observed between SOA 0 ms and SOA 125 ms ($p = 0.0004$) and at even greater SOA intervals (all $p < 0.001$). The findings indicate that the participants did not detect a difference between shorter SOAs (up to 100 ms) and simultaneous trials (SOA 0 ms). The stimuli were only perceived as asynchronous when longer SOAs of at least 125 ms were presented. In line with our hypotheses, these results also indicate that the percentage of synchrony perception was decreased immediately after the multisensory perceptive training (T1) as well as 7 days (T2) after completing the training compared to baseline (T0). However, after 14 days (T3), there were no longer any significant differences compared to T0, indicating that audiovisual integration ability returns to its initial level, with no more training effects visible after 14 days.

- Word Recognition Task (WRT)

The mixed-effects model of WRT revealed a significant main effect of condition (($F(3, 189.25) = 1114.02$, $p < 0.01$) but neither a main effect of POM ($F(3, 189.13) = 1.72$, $p = 0.16$) nor a significant condition x POM interaction ($F(9, 188.25) = 2.36$, $p = 0.26$). Tukey adjusted pairwise comparisons showed comparable word recognition rates when participants saw a frozen image and heard the speaker (unimodal NN) compared to when they saw and heard a video of the speaker (multimodal NN), when no noise was provided (Δ = 1.92, *95% CI* [-2.815, 6.65], $p = 0.72$). Higher proportions of word recognition were demonstrated in the multimodal condition without noise (multimodal NN) in comparison to multimodal (multimodal − 12 dB SNR) (Δ = 66.232, *95% CI* [61.538, 70.925], $p < 0.001$) and unisensory noise conditions (unimodal − 12 dB SNR) (Δ = 78.725, *95% CI* [74.117, 83.331], $p < 0.001$). In the unimodal condition without noise (unimodal NN), participants understood significantly more words correctly than in the multimodal (multimodal − 12 dB SNR) (Δ = 68.14, *95% CI* [63.455, 72.843], $p < 0.001$) and unimodal conditions (unimodal − 12 dB SNR) (Δ = 80.642, *95% CI* [76.035, 85.25], $p < 0.001$) in noise. Comparing the two noise conditions, as expected, rates of word recognition were higher in the multimodal condition (multimodal − 12 dB SNR) compared to the unimodal condition (unimodal − 12 dB SNR) (Δ = 12.493, *95% CI* [8.054, 16.933], $p < 0.001$) (Fig. 3). Hence, as expected, word recognition was higher in conditions without noise. In addition, participants recognized more words correctly in the multimodal condition than in the unimodal condition in noise. Contrary to our hypotheses and although descriptively a slight

improvement in word recognition was visible after multisensory training, no significant increase in word recognition scores was demonstrated between T0 and T1 or T0 and T2.

Discussion

In the present study, we examined whether a short, multisensory perceptive training would enhance multisensory perception in autistic children and adolescents, thereby leading to improvements in speech comprehension. The findings of this study suggest that improving simultaneity perception in autistic children and adolescents is achievable through a single, brief perceptual training session. In addition, the decrease in TBW width after seven days in comparison with the baseline measurement offers preliminary evidence that such training can lead to a reduction in TBW size in autistic children and adolescents. However, given the absence of a reduction in TBW size immediately following the training, these results must be interpreted with particular caution. Nevertheless, our findings align with a series of studies that have demonstrated that brief multisensory training is sufficient to induce a substantial reduction in simultaneity perception and TBW size in TD adults [15, 38–40]. Initial studies observing patients with schizophrenia [48], as well as autistic children and adolescents [42, 43], also point to the usefulness of such training in clinical populations, which is consistent with our findings. The findings of the present study suggest that the training effects persist for a duration of seven days but diminish after 14 days. Conceptualizing prospective interventions, it is essential to acknowledge the necessity of reiterating the training after 14 days at the latest, although earlier repetition may be advisable to preserve training effects. The optimal frequency and duration of such training to achieve the best possible effects has not yet been sufficiently determined in either TD adults or clinical populations and therefore needs to be specified in future studies.

Despite the observation of a descriptive decrease in TBW size following the training, a high heterogeneity and variability in goodness of fit regarding the TBW curve fit was demonstrated across participants. Consequently, it is challenging to draw definite conclusions about the overall population of autistic individuals. The present study suggests that the examination of distinct subpopulations (e.g., children vs. adolescent, lower vs. higher overall and verbal IQ or symptom severity and specific symptom manifestations) would provide a more profound comprehension of the usefulness of multisensory perceptual training in autism, thus, allowing for the development of more specific interventions, which is in line with the findings of previous studies [42, 43].

Notably, the percentage of trials perceived as synchronous immediately after training (T1) fell below the 70% threshold, even for synchronous trials or small SOAs, which are typically perceived as predominantly synchronous. As a result, the calculation of TBW size immediately after the training was not possible. Given that multisensory training does not appear to have a significant impact on the outcomes of these trials in other study populations [15, 38–40], it is unlikely that these effects are directly or exclusively attributable to alterations in perception. Various explanations are possible for these surprising findings. On the one hand, the relatively long duration of the first session (approximately 120 minutes), in conjunction with the absence of a longer break following the training, may have resulted in diminished

attention among the participants. This, in turn, could have led to less precise responses, consequently resulting in unanticipated response pattern. The manifestation of the anticipated training effects after seven days, despite their absence immediately following the training, could also be explained by attention-related limitations. In addition, other potential explanations should be considered that relate to alterations in the response behavior of autistic children and adolescents in contrast to other study populations. In this context, previous studies have demonstrated that autistic children exhibit more conservative response behavior, i.e. rather responding with “no” when unsure, thereby demonstrating distinct response behavior from TD participants [49]. Furthermore, previous studies indicate that autistic individuals show increased reactivity to errors [50] and greater performance impairment reflected by rushed responses and increased error-rates as a result of frustration caused by negative feedback [51]. In relation to the present study, it can be speculated that the participants may have been more inclined to report two stimuli as asynchronous when they were uncertain, particularly following the presentation of negative feedback in the training paradigm. Furthermore, heightened reactivity to negative feedback may also have contributed to hasty responses in this study, which could explain the imprecise and less differentiated response pattern immediately after the training. Future studies should aim to reduce attention related limitations by shortening sessions or implementing longer breaks. Moreover, subsequent studies should control for possible alterations in the response behavior to ensure that the MI abilities of autistic children are assessed correctly and without bias. One feasible approach involves the estimation of individual response criteria, as proposed by Quinde-Zlibut and colleagues (2020) or the individual adaptation of training paradigms.

To the best our knowledge, our study is among the first to examine whether multisensory perceptual training benefits speech comprehension under noisy conditions in autistic children and adolescents. Contrary to our predictions, however, we did not observe any substantial enhancement in speech comprehension following the training. This findings contrasts with the results of studies on TD adults [39] and schizophrenia patients [48], which have demonstrated that a brief, multisensory training can enhance word recognition under noisy conditions. These findings suggest that training effects yield improvements in complex multisensory integrative processes such as speech perception. In contrast, the findings of Kunnath and colleagues (2025) revealed no significant differences in word recognition scores at the group level between trained and untrained participants. However, they did observe that a greater reduction in TBW size after multisensory training was associated with improved speech perception under noisy conditions. They concluded that multisensory training can have a positive effect on speech processing. The utilization of fNIRS revealed that these enhancements may be facilitated by the upregulation of multisensory cortical networks including the left temporal gyrus (MTG), the left superior temporal gyrus (STG) and visual cortices that exhibited low baseline activity. In a study on autistic youth, Feldman and colleagues (2023) obtained analogous results to those reported in our study. Using the McGurk illusion, they demonstrated that multisensory training does not improve audiovisual integration in their sample of autistic youth. In summary, the available evidence suggests that multisensory training may enhance speech comprehension in TD adults and clinical samples; however, the findings remain inconclusive. This phenomenon may be partially attributed to the presence of

significant interindividual variability in response to multisensory training and thus improvements in audio-visual integration abilities following the training [40, 42]. Feldman and colleagues (2023) suggest that the effects of multisensory training may be contingent on nonverbal IQ. In their study training effects generalized to untrained speaker or syllables solely in a subgroup of autistics with above average nonverbal IQ. The absence of substantial enhancement in speech comprehension in the present study may be ascribed, in part, to the presence of considerable interindividual heterogeneity. This observation underscores once more the necessity for the development of training paradigms that are tailored to specific subgroups or even individuals. Moreover, a prior study by Cross and colleagues (2022) underscores that autistic children exhibit fundamental alterations in the processing of multisensory stimuli and the development of multisensory integrative abilities in comparison to TD children. The authors posit that linguistic and non-linguistic audiovisual stimuli follow disparate developmental trajectories in autistic individuals and are processed by distinct cortical networks. The processing of simple audiovisual stimuli necessitates the early activation of cross-sensory regions within the visual and auditory cortices. In contrast, the processing of linguistic stimuli involves the engagement of extensive networks of hierarchically organized brain areas and the implementation of multiple integrative mechanisms. The findings of this study suggest that there may have been no observable effects on speech comprehension in our study because enhancing simultaneity perception for simple, audiovisual stimuli may not be sufficient to improve the integration of complex speech stimuli in autistic children and adolescents.

In addition, two recent studies [52, 53] point out that a distinction must be made between the so-called temporal window of integration (TWI) and the temporal window of synchrony (TWS) since they represent conceptually and terminologically different constructs. Jertberg and colleagues (2025) posit that the TWI corresponds to the temporal window in which two stimuli from different modalities are integrated into one perception by means of MI processes [9]. In contrast, they propose that the TWS merely represents the time window within which two stimuli are perceived as synchronous, which is neither necessary nor sufficient for MI. The present study utilizes the SJT, which has been demonstrated to allow for the determination of the TWS exclusively. Consequently, the question of the extent to which the study design allows for the drawing of conclusions regarding the influence of improved MI ability on speech comprehension arises. Given the fact that the existing paradigms only enhance simultaneity perception, it is possible that no improvement in speech comprehension was demonstrated in the present study because the training did not improve MI but only simultaneity perception. Furthermore, this argumentation also provides an explanatory model for the ambiguous findings in previous studies. In most cases, no clear distinction was made between TWI and TWS, which may have led to contradictory results and should be taken into account in future studies [52, 53]. Hence, subsequent studies should critically examine whether SJT can be replaced by other paradigms (e.g., based on the McGurk illusion) [54]. Such paradigms could facilitate the direct determination of TWI and improvement of MI abilities. This, in turn, may enable more definitive conclusions to be drawn regarding the efficacy of multisensory training in improving speech comprehension.

When interpreting the results of this study, several methodological limitations must be taken into account. Since this study was primarily concerned with testing the applicability of multisensory training in autistic children and adolescents, the sample size was small. There was no comparison group of TD controls or a control group of autistic children and adolescents who did not undergo multisensory training, which limits the comparability of these results. Given that the present study offers preliminary evidence of the potential of multisensory training in autistic children and adolescents, it appears reasonable to include a TD control group to facilitate more direct conclusions in future studies. Even more importantly, RCTs that employ adapted and revised paradigms should be conducted as these studies are crucial for acquiring more nuanced insights and derive more specific interventions for autistic children and adolescents.

In addition, the paradigms employed in this study were designed to align with those of previous studies to ensure high comparability across studies [39, 48]. Consequently, the paradigms employed in this study had been primarily tested in adult samples. The selected SOA range was more limited than in the study by Feldman and colleagues (2023) and the − 12 dB SNR condition in the WRT may have posed challenging conditions for the participants as indicated by a lower rate of correct responses in the WRT compared to TD adults[39] and autistic adults [37]. However, particularly this SNR was selected as it has been demonstrating its optimal performance in inducing multisensory enhancement with regard to speech perception [8]. Subsequent research should examine whether adapting the level of difficulty to the study population or even individuals, e.g., by extending the intervals between SOAs, may facilitate enhancements in training success.

Beyond that, it remains unclear whether children and adolescents with ASD show similar or distinct changes in neural activity patterns in multisensory cortical networks compared to TD individuals lowing audiovisual training. The utilization of imaging techniques in autistic children and adolescents may have the potential to yield more comprehensive insights into the impact of multisensory, perceptive training in this specific study population. Examining alterations in neural activation and individual differences more thoroughly, may thus help to tailor future training more accurately to individual needs, which, based on the results of our study, is urgently needed.

The present study demonstrated that single, brief multisensory perceptual training session can enhance simultaneity perception in autistic children and adolescents. This finding highlights the potential of multisensory training as a promising starting point for the development of cognitive interventions. Our study showed that, similar to TD adults, the training effects persist for a duration of seven days, subsequently returning to baseline level after a period of 14 days. The training did not result in the anticipated enhancement of language comprehension. Subsequent studies must elucidate which variables mediate these effects and which neural processes underlie them to develop subgroup-specific interventions that contribute not only to improving simple audiovisual processing, but also speech and thus communication-related skills in autistic individuals. Such approaches are necessary to broaden the understanding of the cascading effects multisensory processing alterations have on higher cognitive,

perceptual and motor skills, which is crucial to develop appropriate interventions for autistic children and adolescents.

Methods

- Participants

Participants were recruited via flyers at Hannover Medical School and the Lüneburg Institute for Training in Child and Adolescent Behavioral Therapy (Lüneburger Ausbildungsinstitut für Kinder- und Jugendlichenverhaltenstherapie, LAKIJU-VT) as well as specialized autism outpatient centres (Autismus Ambulanz Spektral, Autismus Ambulanz der Lebenshilfe Alfeld | Leine e.V.). The assessment and experiments were conducted either at Hannover Medical School or on site of the outpatient centers and took place between January 2024 and March 2025.

The participants had normal or corrected-to-normal vision, normal hearing and spoke German fluently. None of the participants had a history of genetic diseases, current neurological disorders or dyslexia. Only those who had already received a confirmed diagnosis of either childhood autism, atypical autism or Asperger's syndrome according to ICD-10 criteria [55] were eligible to participate in the study. For participants to be included, diagnostic testing had to comprise at least the following psychometric instruments, conducted by a qualified examiner: German versions of the Autism Diagnostic Observation Schedule (ADOS-2) [56] and the Autism Diagnostic Interview-Revised (ADI-R) [57] as well as standardized intelligence diagnostic (e.g., WISC-V) [58]. Within the study, German translations of the Social Communication Questionnaire (SCQ) [44] as well as the children and adolescent versions of the Autism Spectrum Quotient (AQ-Child; AQ-Adol) [45, 46] were administered to briefly assess current symptom severity. The German version of the vocabulary intelligence subtest of Cattell's Fluid Intelligence Test, Scale 2 (CFT 20-R) [47] was conducted to obtain an overview of the participants' vocabulary. The participants conducted the experiments in a quiet, enclosed and darkened room on a 15' Lenovo B51-80 80lm laptop (screen 1366 × 768 pixel resolution, viewing distance of approximately 60cm).

The Ethics Committee of Hannover Medical School provided approval for the study, and all participant's legal guardians gave written informed consent in accordance with the Declaration of Helsinki. The participants/ their parents received a small monetary reimbursement for study participation. The study was registered at the German Clinical Trials Register (DRKS00032847, date of registration: 25.01.2024).

- Stimuli and design

The study design entailed two established experimental paradigms to examine audio-visual integration, namely the Simultaneity Judgement Task (SJT) and the Word Recognition Task (WRT). Beyond that, a modified version of the SJT with feedback was included as the multisensory training paradigm. Presentation Software was used to present stimuli.

The subjects participated in three sessions. In the initial session, the participants and their parents completed the aforementioned questionnaires. Subsequently, the participants completed the two paradigms (SJT and WRT) immediately before (T0) and immediately following the multisensory training (T1). The participants underwent the multisensory training only once during the first session. Follow-up measurements took place 7 (T2) and 14 days (T3) after the multisensory training.

Simultaneity Judgement Task (SJT)

The SJT was conducted to examine simple, audio-visual processing. In this task, participants were instructed to indicate via button press whether a visual and an auditory stimulus were presented simultaneously or with a time lag. Each trial started with a white fixation cross presented against a black background for 1000 ms (2 deg. of the visual arc). A white ring was then presented as the visual stimulus in the middle of the screen for one refresh cycle (16.6ms, 6° of the visual arc). A 1850Hz beep sound was chosen as the auditory stimulus and shown for 8 ms. After each trial, the fixation cross was presented again for 1000 ms. stimulus onset asynchronies (SOA) ranged between 0 and 250 ms with intervals of 25 ms, resulting in eleven SOA conditions. There were 15 trials for each SOA condition, yielding a total of 165 trials, which were presented in a pseudorandom order. In accordance with previous studies, only visual-leading conditions were applied to enhance ecological validity [16, 39, 59] (Fig. 4).

Word Recognition Task (WRT)

Complex audio-visual integration processes were examined by using a modified version of a speech perception task, initially designed by Ross and colleagues [8]. The modified version used was utilized and tested in various previous studies [39, 60]. The stimuli consisted of German disyllabic nouns selected from the CELEX lexical database [61]. Only nouns that, according to the Mannheim frequency measure (MannMln), occur at least once per one million words were included. A male native speaker of German with linguistic experience spoke the stimuli, which were then videotaped. Each video clip had a duration of 2 seconds. The speaker's face was shown in a full frontal view (720 × 576 pixels, 25° of the visual arc). To enable the analysis of unimodal integration processes in contrast to MI processes, the paradigm comprised a unimodal condition (auditory stimulus only) as well as a multimodal condition (audio-visual stimuli). For the unimodal condition, instead of the recorded video tape, a static image of the speaker's face was shown while the audio stimulus was presented. For the multimodal condition, the original video clip was used with the auditory and the visual stimuli presented simultaneously. To allow for the examination of multisensory improvement in speech perception, the audio stream was mixed either with no background noise (0 dB, NN) or white noise with a signal-to-noise ratio (SNR) of -12 dB. Hence, resulting in four conditions (unimodal NN, unimodal − 12 dB SNR, multimodal 0 dB, multimodal − 12 dB SNR). For each condition, 20 words were presented, resulting in a total of 80 trials per point of measurement. The stimuli were presented in a pseudorandomized order. To mitigate the impact of

learning effects, different words were selected for each point of measurement. Following the presentation of each stimulus, the participants were requested to verbally report the word they comprehended, which was then documented by the experimenter. In the event that the participants did not understand the word, they were invited to either guess the word or state that they did not understand the word being said. The responses were only marked as correct, if the whole word was recognized and stated exactly as it was presented. Subsequent to the presentation of each response, the participants were instructed to press the space bar, which initiated the following trial (Fig. 5).

Multisensory training

In this study, a multisensory training paradigm was utilized, which has previously been employed in several prior studies [38, 39, 48]. The experiment comprised the SJT outlined above, augmented by feedback concerning the accuracy of the response provided following each trial. The provided feedback encompassed both written German phrase (“correct” or “incorrect”) and visual feedback in the form of a green, happy emoji or a red, sad emoji. SOA conditions ranged from 25 to 150 ms in 25 ms intervals. Synchronous (SOA = 0 ms, 60 trials total) and asynchronous SOA conditions (SOA = 25–150 ms, 10 trials per SOA) were presented in equal proportions to guarantee an equal likelihood of synchronous and asynchronous trials, thereby preventing biases in response patterns (Fig. 6).

Data analysis

The statistical analysis was conducted using R statistical software (v2025.09.1 + 401) and MATLAB software (vR2023a).

In order to estimate TBW, a separate Gaussian distribution was calculated for each participant at each time of measurement using the nonlinear least square method with mean, standard deviation and amplitude set as free parameters. The TBW was estimated based on the 70% criterion. Consequently, all SOA conditions that were perceived as simultaneous in at least 70% of the trials, were incorporated in the TBW [42, 43, 48]. To assess training effects on SJT and WRT performance, linear mixed-effect models were conducted. For the SJT, the percentage of trials perceived as simultaneous was selected as the response variable. Point of measurement (POM; 4 levels: T0-T3) and SOA (11 levels: 0-150 ms with 25 ms intervals) were included as fixed effects and participant was incorporated as a random intercept. For the WRT, the number of correctly recognized words was the dependent variable. The fixed effects included POM (4 levels, T0-T3) and condition (unimodal NN, unimodal − 12 dB SNR, multimodal 0 dB, multimodal − 12 dB SNR) whereas participant was incorporated as a random intercept. Residuals were inspected visually using Q-Q plots as well as residual-versus-fitted plots to ensure that the assumptions of normal distribution and homoscedasticity were not violated. Values above or below 1.5 IQR were considered outliers and excluded from further analyses. One participant was excluded from the analysis due to an implausible response pattern in the SJT. At T0 and T1, this participant reported perceiving at least 75% of trials as synchronous in each SOA condition, while at T2 and T3, he reported perceiving a

maximum of 25% of trials as synchronous in all SOA conditions. At all measurement points, there was no differentiation in this participant's response behavior depending on the SOA condition, as would have been expected. Two other participants were excluded from the analysis because they showed a completely random response pattern, which is most likely due to either a lack of understanding of the task or a lack of motivation. Restricted maximum likelihood (REML) was used for estimation of model parameters. Fixed effects were tested using a Type-III F-test with Kenward-Roger approximation to estimate degrees of freedom. Subsequently, post-hoc tests were performed using Tukey adjusted pairwise comparisons.

Declarations

Acknowledgements

We gratefully acknowledge the contributions of our colleague Prof. Dr. Stefan Bleich, who was instrumental in initiating this study and sadly passed away before its completion.

We thank all participants for their time and effort in participation. We thank the Lüneburger Ausbildungsinstitut für Kinder- und Jugendlichenverhaltenstherapie (LAKIJU-VT), the Autismus Ambulanz Spektral and the Autismus Ambulanz der Lebenshilfe Alfeld | Leine e.V. for their support in recruiting participants and conducting data collection in their outpatient clinics.

Research Funding

This research did not receive funding.

Author Contributions

LM participated in the coordination and conception of the study design, performed the measurement, conducted analysis and interpretation of the data and drafted the manuscript. AB and EG participated in the coordination and conception of the study and supported the data interpretation. HW and SB participated in the conception and coordination of the study. SB passed away shortly before submission. GS conceived of the study, participated in design and coordination of the study and supported data interpretation.

Data Availability

The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Additional Information

To avoid ableist language, we opt for identity-first (autistic person) or person first language (person on the autism spectrum), as these terms are preferred by the majority of English-speaking autistic individuals.

Conflict of Interest

The authors declare no competing interests.

References

1. Murray, M. M., Lewkowicz, D. J., Amedi, A. & Wallace, M. T. Multisensory Processes: A Balancing Act across the Lifespan. *Trends Neurosc*. **39**, 567–579. 10.1016/j.tins.2016.05.003 (2016).
2. Murray, M. M. & Wallace, M. T. in *The neural bases of multisensory processes*. (eds Murray, M. M. & Wallace, M. T.) (CRC, 2011).
3. Zhou, H., Cheung, E. F. & Chan, R. C. Audiovisual temporal integration: Cognitive processing, neural mechanisms, developmental trajectory and potential interventions. *Neuropsychologia* **140**, 107396. 10.1016/j.neuropsychologia.2020.107396 (2020).
4. Ampollini, S., Ardizzi, M., Ferroni, F. & Cigala, A. Synchrony perception across senses: A systematic review of temporal binding window changes from infancy to adolescence in typical and atypical development. *Neurosc Biobehav Rev.* **162**, 105711. 10.1016/j.neubiorev.2024.105711 (2020).
5. Rosenblum, L. D. Speech Perception as a Multimodal Phenomenon. *Curr. Dir. Psychol. Sci.* **17**, 405–409. 10.1111/j.1467-8721.2008.00615.x (2008).
6. Wallace, M. T. & Stevenson, R. A. The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia* **64**, 105–123. 10.1016/j.neuropsychologia.2014.08.005 (2014).
7. Wallace, M. T., Woynaroski, T. G. & Stevenson, R. A. MI as a Window into Orderly and Disrupted Cognition and Communication. *Annu. Rev. Psychol.* **71**, 193–219. 10.1146/annurev-psych-010419-051112 (2020).
8. Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C. & Foxe, J. J. Do You See What I Am Saying? Exploring Visual Enhancement of Speech Comprehension in Noisy Environments. *Cereb. Cortex*. **17**, 1147–1153. 10.1093/cercor/bhl024 (2006).
9. Sumby, W. H. & Pollack, I. Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* **26**, 212–215. 10.1121/1.1907309 (1954).
10. Lewkowicz, D. J., Schmuckler, M. & Agrawal, V. The multisensory cocktail party problem in adults: Perceptual segregation of talking faces on the basis of audiovisual temporal synchrony. *Cognition* **214** 10.1016/j.cognition.2021.104743 (2021).
11. Pepper, J. L. & Nuttall, H. E. Age-related changes to MI and audiovisual speech perception. *Brain Sci.* **13** 10.3390/brainsci13081126 (2023).
12. Musacchia, G. & Schroeder, C. E. Neuronal mechanisms, response dynamics and perceptual functions of multisensory interactions in auditory cortex. *Hear. Res.* **258**, 72–79. 10.1016/j.heares.2009.06.018 (2009).
13. Van Eijk, R. L., Kohlrausch, A., Juola, J. F. & van de Par Audiovisual synchrony and temporal order judgments: effects of experimental method and stimulus type. *Percept. Psychophys*. **70**, 955–968.

10.3758/pp.70.6.955 (2008).

14. Hillock-Dunn, A. & Wallace, M. T. Developmental changes in the multisensory temporal binding window persist into adolescence. *Dev. Sci.* **15**, 688–696. 10.1111/j.1467-7687.2012.01171.x (2012).
15. Setti, A. et al. Improving the efficiency of MI in older adults: Audio-visual temporal discrimination training reduces susceptibility to the sound-induced flash illusion. *Neuropsychologia* **61**, 259–268. 10.1016/j.neuropsychologia.2014.06.027 (2014).
16. Stevenson, R. A. & Wallace, M. T. Multisensory temporal integration: task and stimulus dependencies. *Exp. Brain Res.* **227**, 249–261. 10.1007/s00221-013-3507-3 (2013).
17. Weissgerber, T., Müller, C., Stöver, T. & Baumann, U. Age differences in speech perception in noise and sound localization in individuals with subjective normal hearing. *Front. Psychol.* **13**, 845285. 10.3389/fpsyg.2022.845285 (2022).
18. Baum, S. H., Stevenson, R. A. & Wallace, M. T. Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder. *Prog Neurobiol.* **134**, 140–160. 10.1016/j.pneurobio.2015.09.007 (2015).
19. Zhou, H. et al. Multisensory temporal binding window in autism spectrum disorders and schizophrenia spectrum disorders: A systematic review and meta-analysis. *Neurosci. Biobehav Rev.* **86**, 66–76. 10.1016/j.neubiorev.2017.12.013 (2018).
20. Sartorato, F., Przybylowski, L. & Sarko, D. K. Improving therapeutic outcomes in autism spectrum disorders: Enhancing social communication and sensory processing through the use of interactive robots. *J. Psychiatr Res.* **90**, 1–11. 10.1016/j.jpsychires.2017.02.004 (2017).
21. Russo, N. et al. A cascading effects model of early sensory development in autism. *Psychol. Rev.* **10**, 1037rev0000558 (2025).
22. Estes, A. et al. Behavioral, cognitive, and adaptive development in infants with autism spectrum disorder in the first 2 years of life. *J. Neurodev Disord*. **7** (1), 24. 10.1186/s11689-015-9117-6 (2015).
23. Turner-Brown, L. M., Baranek, G. T., Reznick, J. S., Watson, L. R. & Crais, E. R. The First Year Inventory: a longitudinal follow-up of 12-month-old to 3-year-old children. *Autism* **17**(5), (2013). 527 – 40; 10.1177/1362361312439633.
24. Stevenson, R. A. et al. The cascading influence of multisensory processing on speech perception in autism. *Autism* **22**, 609–624. 10.1177/1362361317704413 (2018).
25. Robertson, C. E. & Baron-Cohen, S. Sensory perception in autism. *Nat. Rev. Neurosci.* **18**, 671–684. 10.1038/nrn.2017.112 (2017).
26. Cascio, C. J., Woynaroski, T., Baranek, G. T. & Wallace, M. T. Toward an interdisciplinary approach to understanding sensory function in autism spectrum disorder. *Autism Res.* **9**, 920–925. 10.1002/aur.1612 (2016).
27. Smith, E., Zhang, S. & Bennetto, L. Temporal synchrony and audiovisual integration of speech and object stimuli in autism. *Res. Autism Spectr. Disord*. **39**, 11–19. 10.1016/j.rasd.2017.04.001 (2017).

28. de Boer-Schellekens, L., Eussen, M. & Vroomen, J. Diminished sensitivity of audiovisual temporal order in autism spectrum disorder. *Front. Integr. Neurosci.* **7**, 8. 10.3389/fnint.2013.00008 (2013).
29. Noel, J. P., Niear, M., Stevenson, R., Alais, D. & Wallace, M. Atypical rapid audio-visual temporal recalibration in autism spectrum disorders: Audiovisual temporal recalibration in ASD. *Autism Res.* 10; (2016). 10.1002/aur.1633
30. Stevenson, R. A. et al. Multisensory temporal integration in autism spectrum disorders. *J. Neurosci.* **34**, 691–697. 10.1523/JNEUROSCI.3615-13.2014 (2014).
31. Woynaroski, T. G. et al. Multisensory Speech Perception in Children with Autism Spectrum Disorders. *J. Autism Dev. Disord*. **43**, 2891–2902. 10.1007/s10803-013-1836-5 (2013).
32. Brandwein, A. B. et al. Neurophysiological indices of atypical auditory processing and MI are associated with symptom severity in autism. *J. Autism Dev. Disord*. **45**, 230–244. 10.1007/s10803-014-2212-9 (2015).
33. Feldman, J. I. et al. Audiovisual MI in individuals with autism spectrum disorder: A systematic review and meta-analysis. *Neurosci. Biobehav Rev.* **95**, 220–234. 10.1016/j.neubiorev.2018.09.020 (2018).
34. Kawakami, S., Uono, S., Otsuka, S., Zhao, S. & Toichi, M. Everything has its time: Narrow temporal windows are associated with high levels of autistic traits via weaknesses in MI. *J. Autism Dev. Disord*. **50**, 1561–1571. 10.1007/s10803-018-3762-z (2020).
35. Ainsworth, K. & Bertone, A. Audiovisual temporal binding window narrows with age in autistic individuals. *Autism Res.* **16**, 355–363. 10.1002/aur.2860 (2023).
36. Beker, S., Foxe, J. J. & Molholm, S. Ripe for solution: Delayed development of multisensory processing in autism and its remediation. *Neurosci. Biobehav Rev.* **84**, 182–192. 10.1016/j.neubiorev.2017.11.008 (2018).
37. Borgolte, A. et al. Enhanced attentional processing during speech perception in adult high-functioning autism spectrum disorder: An ERP-study. *Neuropsychologia* 161, 108022; (2021). 10.1016/j.neuropsychologia.2021.108022
38. Powers, A. R., Hillock, A. R. & Wallace, M. T. Perceptual Training Narrows the Temporal Window of Multisensory Binding. *J. Neurosci.* **29**, 12265–12274. 10.1523/JNEUROSCI.3501-09.2009 (2009).
39. Zerr, M. et al. Brief Sensory Training Narrows the Temporal Binding Window and Enhances Long-Term Multimodal Speech Perception. *Front. Psychol.* **10**, 2489. 10.3389/fpsyg.2019.02489 (2019).
40. Kunnath, A. J., Bertisch, H. S., Kim, A. S., Gifford, R. H. & Wallace, M. T. Effects of multisensory simultaneity judgment training on the comprehension and cortical processing of speech in noise: a randomized controlled trial. *Sci. Rep.* **15**, 12956. 10.1038/s41598-025-96121-4 (2025).
41. Stevenson, R. A., Wilson, M. M., Powers, A. R. & Wallace, M. T. The effects of visual training on multisensory temporal processing. *Exp. Brain Res.* **225**, 479–489. 10.1007/s00221-012-3387-y (2013).
42. Feldman, J. I. et al. A Randomized Controlled Trial for Audiovisual Multisensory Perception in Autistic Youth. *J. Autism Dev. Disord*. **53**, 4318–4335. 10.1007/s10803-022-05709-6 (2023).

43. Feldman, J. I. et al. Plasticity of Temporal Binding in Children with Autism Spectrum Disorder:A Single Case Experimental Design Perceptual Training Study. *Res. Autism Spectr. Disord*. **74** 10.1016/j.rasd.2020.101555 (2020).
44. Bölte, S., Poustka, F., Rutter, M. & Bölte, S. *Fragebogen zur sozialen Kommunikation: Autismus Screening; FSK* (Huber, 2006).
45. Auyeung, B., Baron-Cohen, S., Wheelwright, S. & Allison, C. The autism spectrum quotient: Children's version (AQ-Child). *J. Autism Dev. Disord*. **38**, 1230–1240. 10.1007/s10803-007-0504-z (2008).
46. Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R. & Wheelwright, S. The autism-spectrum quotient (AQ)—adolescent version. *J. Autism Dev. Disord*. **36**, 343–350. 10.1007/s10803-006-0073-6 (2006).
47. Weiß, R. H. *CFT 20-R: Grundintelligenztest skala 2-revision* (Hogrefe, 2006).
48. Ghaneirad, E., Schütz, H., Möde, L. & Anna, B. R, S. G. Multisensory simultaneity-judgment training improves speech comprehension in patients with schizophrenia *Neuropsychol Rehabil.;* (in press).
49. Quinde-Zlibut, J. M. et al. Elevated Thresholds for Light Touch in Children With Autism Reflect More Conservative Perceptual Decision-Making Rather Than a Sensory Deficit. *Front. Hum. Neurosci.* **14**, 122. 10.3389/fnhum.2020.00122 (2020).
50. Möde, L. et al. Cognitive control in adults with high-functioning autism spectrum disorder: a study with event-related potentials. *Front. Psychiatry*. **14**, 1180827. 10.3389/fpsyt.2023.1180827 (2023).
51. Ghosn, F. et al. The effects of reward and frustration on the task performance of autistic children and adolescents. *Res. Dev. Disabil.* **140**, 104567. 10.1016/j.ridd.2023.104567 (2023).
52. Parise, C. V., Parise, E. & Parise, A. Perceiving audiovisual synchrony: A quantitative synthesis of simultaneity and temporal order judgments from 185 studies. *Neurosci. Biobehav Rev.* **181**, 106449. 10.1016/j.neubiorev.2025.106449 (2026).
53. Jertberg, R. M., Soto-Faraco, S., van Wassenhove, V. & Van der Burg, E. Temporal Window of Integration XOR Temporal Window of Synchrony. *Multisens. Res.* **38**, 353–367. 10.1163/22134808-bja10171 (2025).
54. McGurk, H. & Macdonald, J. Hearing lips and seeing voices. *Nature* **264**, 746–748. 10.1038/264746a0 (1976).
55. WHO. *The ICD-10 classification of mental and behavioural disorders* (World Health Organization, 1993).
56. Poustka, L. et al. Hogrefe,. Diagnostische beobachtungsskala für autistische störungen-2 (ADOS-2). Deutschsprachige Fassung der Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) (Lord, C. et al.) (2015).
57. Bölte, S., Rühl, D., Schmötzer, G. & Poustka, F. *ADI-R. Diagnostisches Interview für Autismus-revidiert*Huber, (2006).
58. Raiford, S. E. *The Wechsler intelligence scale for children—fifth edition integrated* 303–332 (The Guilford Press, 2018). Flanagan, D. P. & McDonough, E. M.

59. Powers Iii, A. R., Hillock-Dunn, A. & Wallace, M. T. Generalization of multisensory perceptual learning. *Sci. Rep.* **6**, 23374. 10.1038/srep23374 (2016).
60. Borgolte, A. et al. Neural correlates of audiovisual integration in schizophrenia–an ERP study. *Front. Psychiatry*. **15** 10.3389/fpsyt.2024.1492266 (2024).
61. Baayen, R. H., Piepenbrock, R. & Gulikers, L. The CELEX lexical database (release 2). *Distributed by the linguistic data consortium, University of Pennsylvania* (1995).

Figures

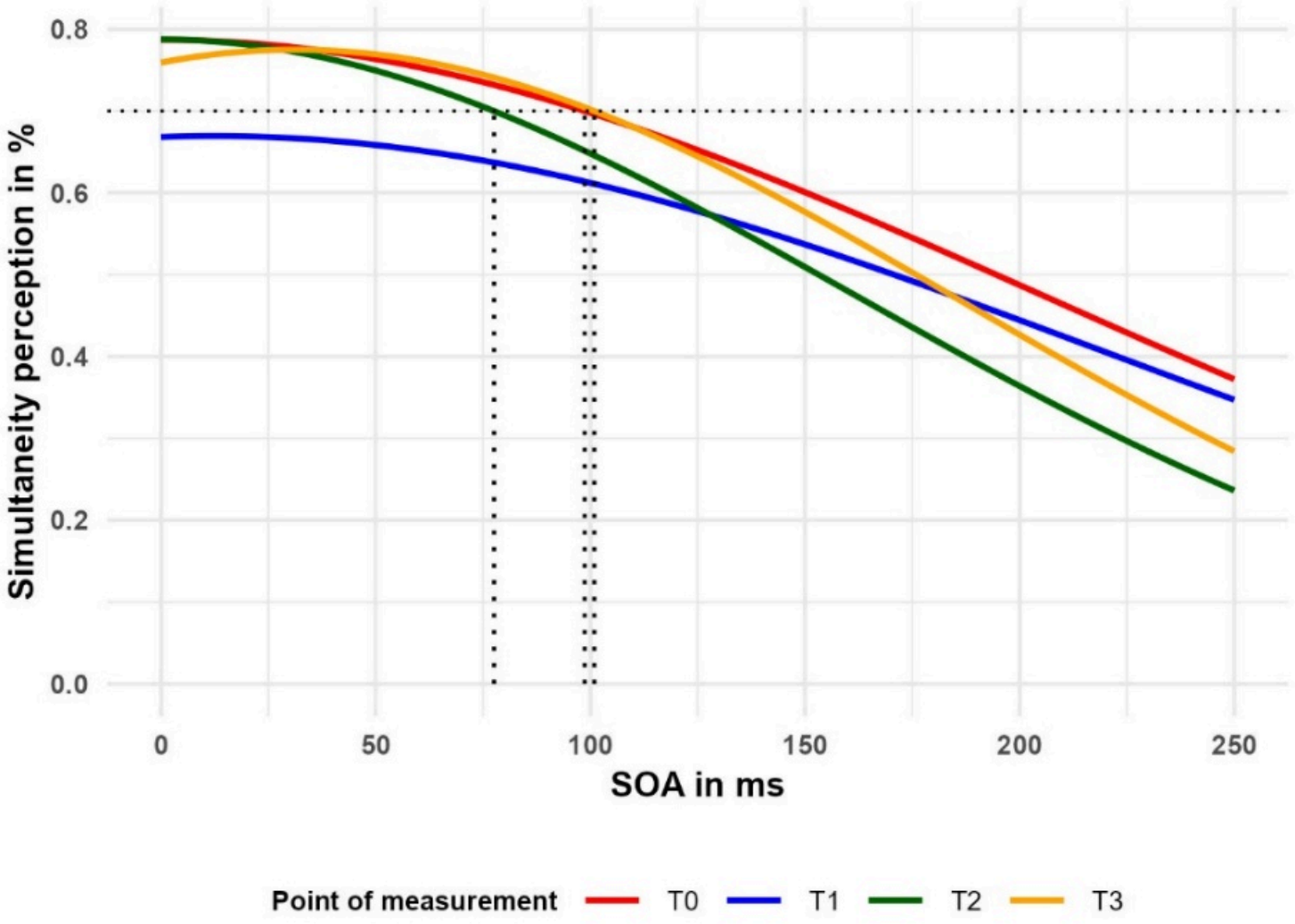

Figure 1

Gaussian fits of simultaneity perception according to stimulus onset asynchrony (SOA) across points of measurement (T0-T3) and temporal binding window according to the 70%-criterion.

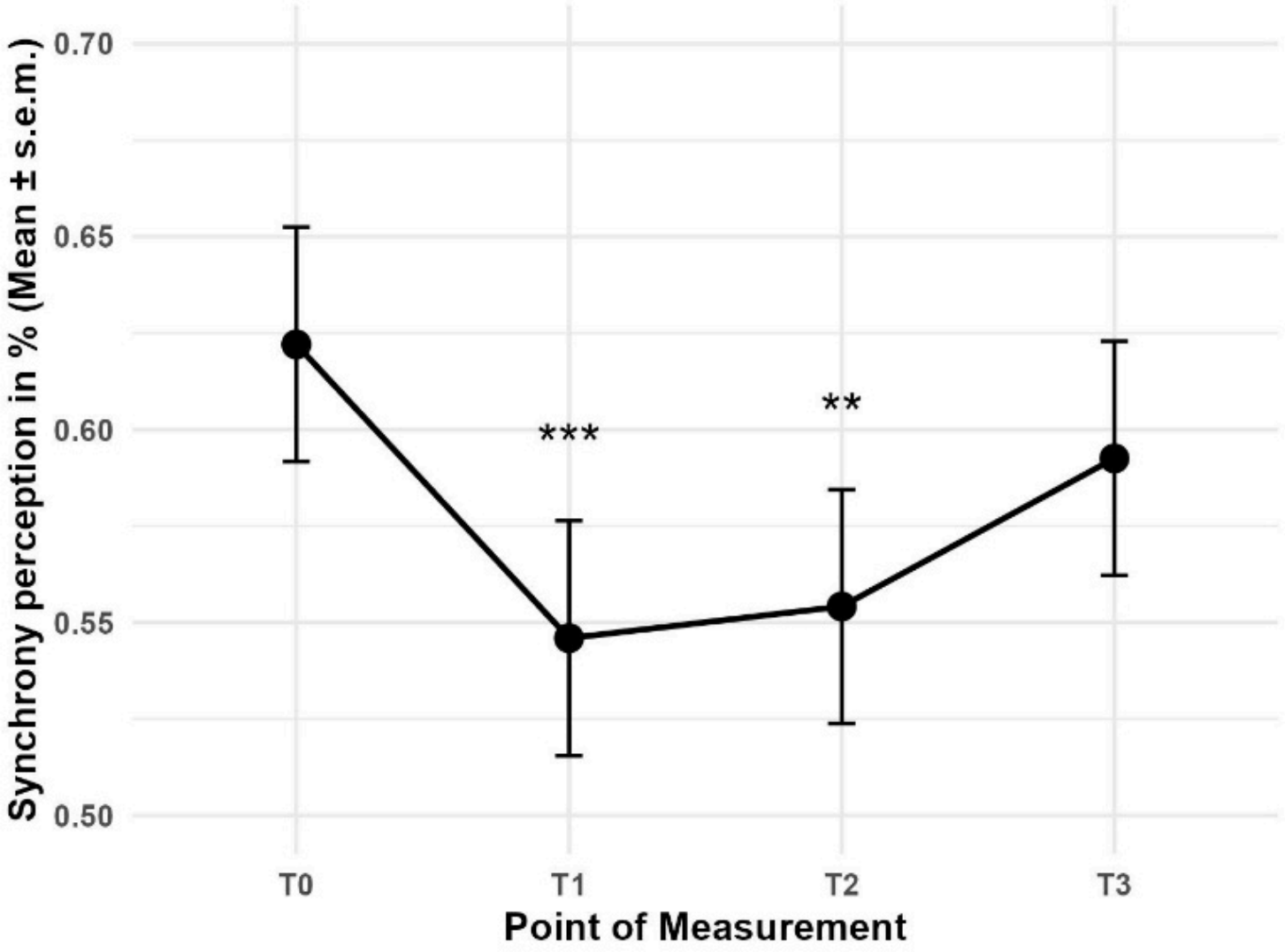

Figure 2

Synchrony perception in % (Mean + SD) across points of measurement (T0-T3) and all stimulus onset asynchronies (SOA) conditions. Asterisks indicate a significant reduction in synchrony perception percentage in T1 and T2 relative to baseline but not in T3 compared to baseline (*** p < .001, **p<.05).

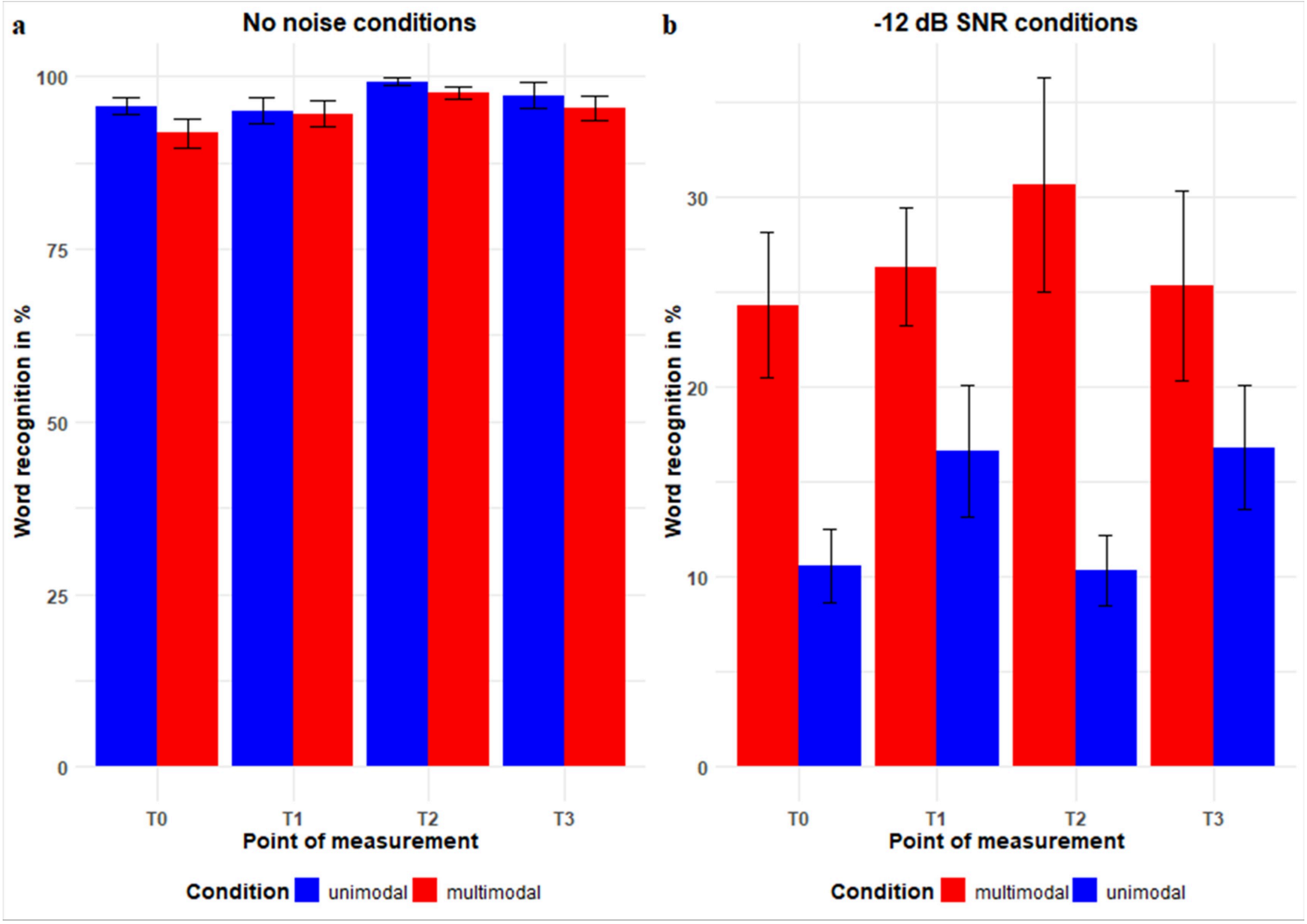

Figure 3

Word recognition in percent across points of measurement (T0-T3) in **a** unimodal and multimodal trials with no noise (NN) and **b** unimodal and multimodal trials with -12 dB signal-to-noise ratio (SNR). Error bars depict the standard error of the mean (s.e.m).

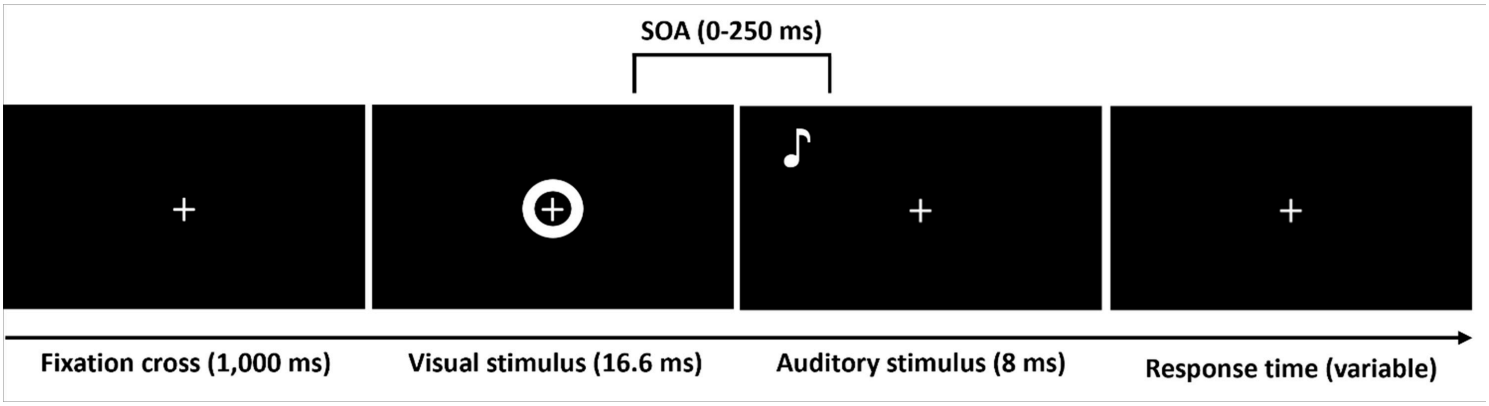

Figure 4

Experimental design of the Simultaneity Judgement Task.

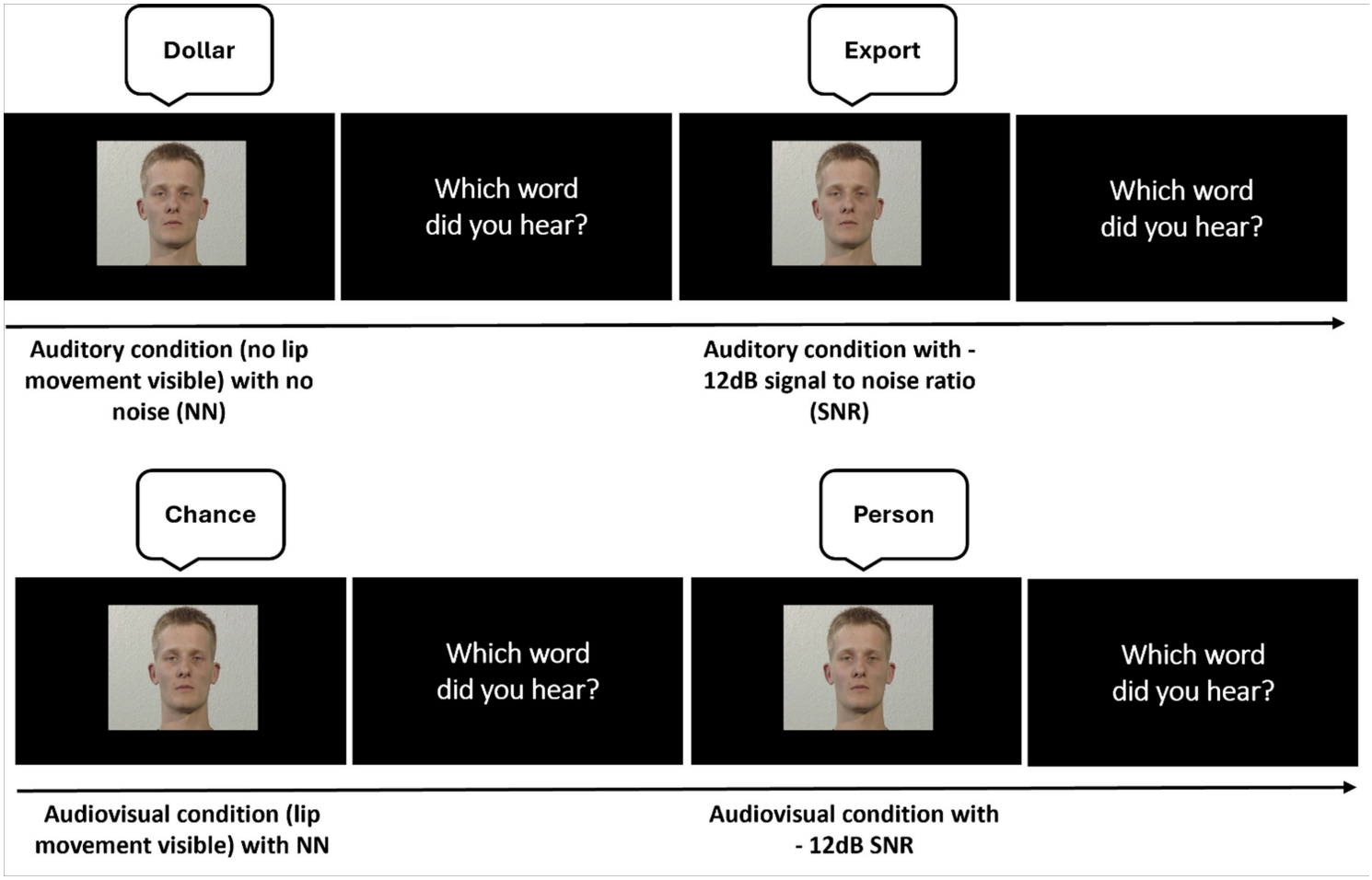

Figure 5

Experimental design of the Word Recognition Task with the four conditions depicted (unimodal no noise (NN), unimodal -12 dB signal-to-noise ratio (SNR), multimodal 0 dB, multimodal -12 dB SNR).

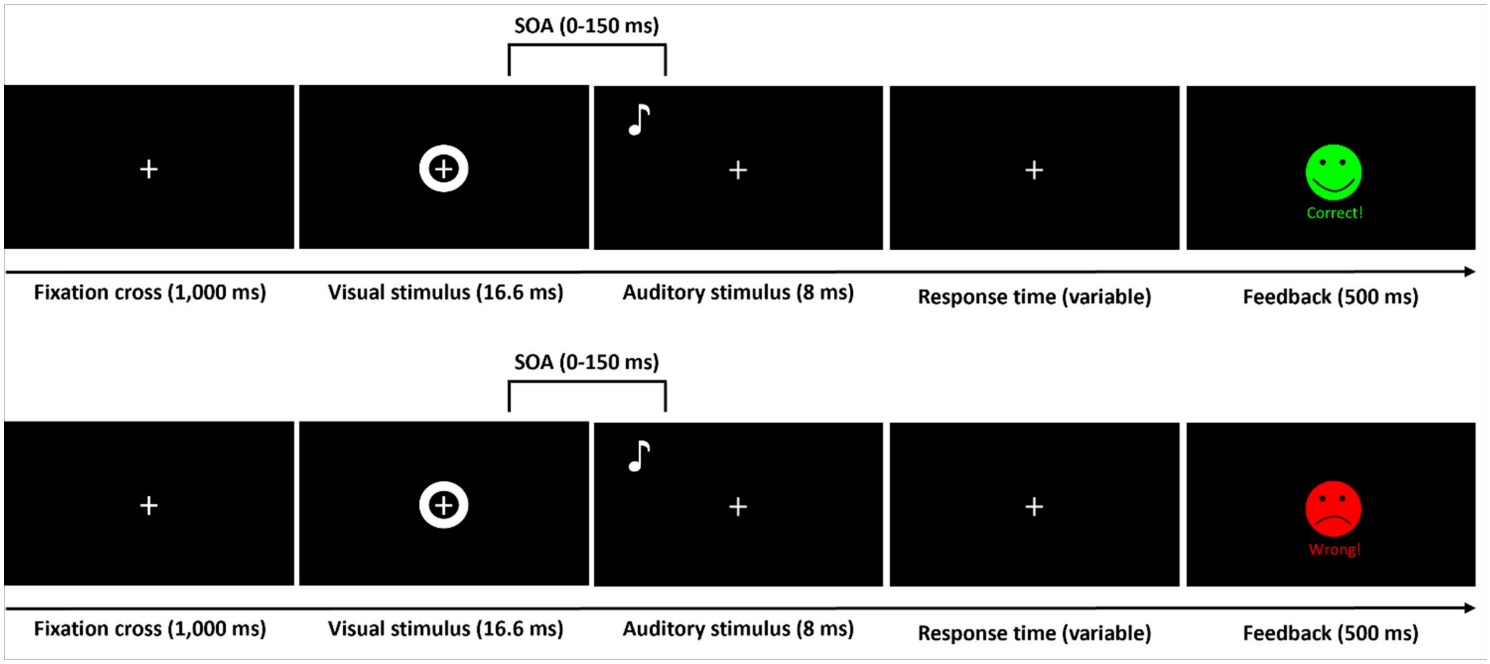

Figure 6

Experimental design of the multisensory training paradigm. A single correct trial is followed by an erroneous trial.