Title

Atypical attention to voice in toddlers and pre-schoolers with autism spectrum disorder is related to unimpaired cognitive abilities. An ERP study.

RUNNING HEAD: P3a is reduced in ASD without developmental delay, but intact in ASD with DD

Authors

Alessandra Piatti [a, *]

Sara Van der Paelt [a]

Petra Warreyn [a]

Herbert Roeyers [a]

a: Department of Experimental Clinical and Health Psychology, Ghent University, Henri Dunantlaan 2, Ghent, Belgium.

* Corresponding author: alessandra.piatti@ugent.be

Funding statement

This research was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement (No 642996, 2015; Brainview).

Ethical Committee approval

The study protocol was approved at Ghent University by the Ethische Commissie Psychologie en Pedagogische Wetenschappen, approval code: 2016/85

Declaration of Competing Interest

The authors declare there is no actual or potential conflict of interest in relation to this article.

Author contributions

AP designed the study, collected and analysed EEG/ERP data, conducted and interpreted statistical analyses, and drafted the manuscript. SVdP designed the study, collected and analysed behavioural data, and reviewed and edited the manuscript. PW designed the study, and reviewed and edited the manuscript. HR designed the study, and reviewed and edited the manuscript.

Abstract

Background: Sound perception in autism spectrum disorder (ASD) is usually unimpaired, even when auditory stimuli carry a social value, as it is the case for speech. Nevertheless, orienting to sounds in a speech context might be problematic in some individuals with ASD, which in ERP studies is reflected by a diminished P3a component. As P3 values and cognitive abilities seem to be inversely related under some circumstances, the current study investigates whether diminished attentional orienting to sounds in speech is equally observable in children with ASD with and without developmental delay (DD).

Method: Fifty-one children with typical development or ASD, with or without comorbid DD (ASD/noDD and ASD/DD), aged 1.5 through 4 years took part in a passive auditory oddball task while EEG data were recorded. The paradigm consisted in the presentation of two deviant stimuli (one vowel sound and one complex tone) either in a speech or in a non-speech context.

Results: We found overall more negative MMN voltages in both ASD groups compared to TD. For P3a mean voltages, we found an attenuated response in children ASD/noDD when deviant tones were presented in speech, but not in other conditions. Children with ASD/DD did not differ from TD in P3a mean voltages.

Conclusion: Atypical speech sound processing might be more accentuated in children with ASD/noDD than in their peers with comorbid DD. This finding is interpreted within the theoretical framework of neural adaptation.

Keywords: auditory oddball, MMN, P3a, ASD, developmental delay, neural adaptation.

1. Introduction

Babies are born with an innate ability to connect to other humans. As a matter of fact, starting hours after birth, infants orient more easily to social stimuli like faces (Farroni et al., 2013; Johnson et al., 1991) or human voice (Vouloumanos et al., 2010) than to non-social visual or auditory stimuli. In particular, the ability to orient to human voice, in itself a basic social-communicative skill which is likely to be present in utero (Kisilevsky et al., 2009), shares some of the cognitive underpinnings of more complex communicative abilities, and at the same time has a pivotal role in their development (Mundy, 2018; Sirri et al., 2020).

For this reason, the investigation of anomalies in voice processing has long been an attractive area of study for researchers interested in mapping the early symptoms of autism spectrum disorder (ASD), a neurodevelopmental condition which is namely characterised by deficits in social communication and social interaction, and by restricted, stereotyped behaviours and interests (American Psychiatric Association, 2013).

Typically, the most basic stages of voice sound processing are investigated by means of brain activity data, which do not require overt responses. In particular, event-related potentials (ERP) are useful as they provide a direct measure of brain activity and have high temporal resolution, which is appropriate to tackle perceptual and attentional processes occurring in an order of magnitude of approximately 10 milliseconds (Woodman, 2010).

The experimental task of election for the study of brain responses to voice sounds is the oddball paradigm, during which a given sound is repeatedly presented (henceforth, ‘standard’ sound) and is occasionally substituted by a different, infrequent sound (deviant). In children, up to approximately 10 years of age (Ponton et al., 2000), the ERP response to standard sounds is biphasic and includes a positive deflection peaking between 100 ms and 200 ms after stimulus onset (P1), and a negative deflection between 200 and 300 ms (N2). In addition to the P1-N2 complex, which is mostly linked to perceptual, pre-attentional processing, deviant sounds elicit a positive deflection between 300 and 400 ms after stimulus onset (P3), which reflects attentional orienting.

The analysis of ERP responses to infrequent sounds is generally based on the difference wave, which is computed by subtracting the average response to standard sounds from the average response to deviant sounds. In the difference wave, the main components of interests are the mismatch negativity (MMN), which in children is mostly related to the difference in the P1-N2 complex between deviant and standard stimulus, and the P3 effect, which is the residual P3 component of the ERP to the deviant stimulus after the ERP to the standard stimulus is subtracted. The MMN component is typically observed approximately from 150 ms to 250 ms after stimulus onset. Although the MMN is elicited independently of attentional allocation, attentional process can still influence its latency and amplitude.

The P3 effect in the difference ERP is more closely related to attentional shift and is sensitive to the probability of occurrence of the deviant stimulus, as well as to the extent to which it acoustically differs from standard (Goldstein et al., 2002; Polich, 2007; Wronka et al., 2008). The initial phase of attentional orienting is mostly indexed by the frontal P3, or P3a subcomponent, whereas later stages of attentional shift are reflected by the parietal subcomponent or P3b. The latter is generally observed in active oddball tasks, where it is elicited by target stimuli, which require a behavioural response. In contrast to the P3b subcomponent, the P3a does not differ largely between active and passive paradigms (Polich, 2007; Wronka et al., 2008).

The oddball ERP literature on voice sound perception in ASD focuses mainly on the MMN and P3a components of the mismatch response, and concerns mostly school-aged children, from seven to thirteen years of age, and adults.

The MMN component has been reported as intact or enhanced both in children (Whitehouse & Bishop, 2008; Čeponienė et al., 2003; Lepistö et al., 2005, 2006, 2008; Lindström et al., 2018; Vlaskamp et al., 2017) and in adults with ASD (Fan & Cheng, 2014; Kujala et al., 2007; Lepistö et al., 2007) when the experimental manipulations on voice sounds are exclusively phonetic or phonological, suggesting intact perception of the acoustic features of speech. Remarkably, a reduced MMN was found in children with ASD when the experimental paradigms involved changes in more than one acoustic feature at a time (Lepistö et al., 2008; Vlaskamp et al., 2017), or more subtly, when manipulating a single acoustic feature like pitch resulted in changes at the emotional or semantic level (Charpentier et al., 2018; Lindström et al., 2016; Yu et al., 2015). A reduced MMN in response to

multi-level changes was also found in adolescents (Ludlow et al., 2014) and adults with ASD (Fan & Cheng, 2014). In addition to the paradigm structure, individual differences have been found to influence MMN mean voltages in ASD. Overall, MMN attenuation seems related to comorbid intellectual disability (Chen et al., 2020; Schwartz et al., 2018), younger age (Schwartz et al., 2018), and limited verbal ability (Chen et al., 2020; Matsuzaki et al., 2019), although enhanced MMN might also be related more severe social communication symptoms (Vlaskamp et al., 2017).

As far as the P3a effect is concerned, most studies on speech sound processing found it to be diminished specifically in response to deviant sounds presented amid standard speech sounds, but intact or enhanced response to deviant sounds (including speech sounds) in a series of standard non-speech stimuli, both in children and adults with ASD (Čeponienė et al., 2003; Lepistö et al., 2005, 2006; Lindström et al., 2016, 2018; Whitehouse & Bishop, 2008).

Overall, these findings suggest that, in spite of unimpaired acoustic processing of speech, attentional orienting to sounds in speech is diminished in children with ASD. Nevertheless, the ERP literature on speech sound processing in ASD is biased by the relatively high IQ of the children involved (Chen et al., 2020; Schwartz et al., 2018), and the key finding of a reduced P3a effect in spite of a normal MMN in response to deviant sounds in speech might specifically concern individuals without intellectual disability. Just like for MMN, individual differences involving age, language ability, and IQ might contribute to explaining the heterogeneity of P3a findings as well as experimental manipulations do. However, the relationship between these factors and attentional orienting as indexed by the P3a component may be a non-linear one.

As a matter of fact, a link is known to exist between cognitive abilities and attentional orienting as indexed by the P3a component, although this relationship can be either a direct or an inverse one, depending on the conditions under which the P3a component is elicited, and on which aspects of IQ are taken into account. In typically developing individuals, a larger P3a effect has been reported to correlate with a higher full-scale IQ (Wronka et al., 2013). However, it is remarkable that specific traits of intelligence, namely those related to the freedom from distractibility trait, are inversely related to attentional orienting under some experimental manipulations. Children (Kilpeläinen et al., 1999a, 1999b) and adolescents (Määttä et al., 2005) with low distractibility traits have been shown to display a diminished P3a effect in response to novel sounds compared to their peers with higher distractibility

traits in passive auditory oddball tasks. Adolescents with lower distractibility traits, and attenuated P3a effect, had also higher verbal IQ than their peers with high distractibility scores and enhanced P3a (Määttä et al., 2005).

Thus, enhanced attentional orienting does not homogeneously translate into benefits for the individual, and diminished orienting is in turn not always detrimental. In particular, the diminished P3a to sound change in speech observed in children with ASD and high IQ might be related to a more efficient allocation of cognitive resources, and adaptive processes acting at the neural level. Cognitive compensation in ASD is described as the possibility for affected individuals to achieve an observable performance that is at least superficially similar to the one of neurotypicals (Jones et al., 2014; Livingstone & Happé, 2017; Philip et al., 2012). More recently, within the theorising efforts around ASD pathophysiology, the concept of cognitive compensation has been further elaborated upon and extended to include neural processes that allow for an optimization of an individual's general functioning given some specific physiological constraints. Because this process will not necessarily result in typical or nearly typical brain function, it has been referred to as 'neural adaptation' rather than 'compensation' (Johnson et al., 2015, 2017).

In the current study we investigated the relationship between IQ and auditory processing of speech and non-speech in toddlers with ASD, using a passive oddball paradigm (abridged from Whitehouse & Bishop, 2008), in which deviant speech and non-speech stimuli are presented in a stream of speech or non-speech repeated, standard stimuli, henceforth referred to as background stimuli. The comparison of the same deviant stimuli based on whether they occur in a speech or non-speech background is relevant as previous studies suggest that individuals with ASD orient atypically to speech and non-speech sounds in a speech context, rather than to speech sounds as such. Unlike most of the previous studies, we also included participants with ASD and comorbid developmental delay (ASD/DD) in addition to children with ASD and no developmental delay (ASD/noDD).

The hypothesis is that, compared to TD children, children with ASD/noDD will show a diminished P3a effect to deviant stimuli, specifically for deviant sounds in speech, in spite of intact MMN. On the other hand, we expect a less pronounced difference, if any, in P3a effect between children with ASD/DD and TD children, although MMN might be attenuated in this group, given the comorbidity with developmental delay.

2. Methods

2.1 Test battery

The behavioural test battery included an observation of ASD symptomatology (Autism Diagnostic Observation Scale – 2, henceforth ADOS-2, Lord et al., 2012) and a developmental test of cognitive development (Mullen's Scale of Early Learning, henceforth M-SEL, Mullen et al., 1995).

The ADOS-2 measures autistic traits and is suitable for use from age 12 months through adulthood. Symptoms related to social affect (SA) and restricted, repetitive behaviour (RRB) are assessed on a scale in which higher scores correspond to higher severity. Based on the age and language development of the children, different modules (Module 1 to 3 and Toddler Module) were administered. Total and subscales raw scores were converted to calibrated severity scores (CSS, Esler et al., 2015; Hus et al., 2014; Lord et al., 2012), ranging from 1 (mildest class) to 10 (most severe).

The M-SEL test provides an indicative measure of child development based on an evaluation of their receptive visual, fine motor, receptive language and expressive language abilities. In the current study we calculated the children's developmental quotient (DQ, see also Messinger et al., 2013) – an equivalent measure of full-scale IQ - as the ratio, multiplied by 100, between the mental age corresponding to the test score and actual age. Non-verbal DQ was calculated as the average between the ratio-DQ scores in visual and fine-motor domains, whereas language comprehension and production scores were used to calculate the verbal DQ.

In children without a diagnosis of ASD that were recruited for the control group, the general development was considered as typical if the parents did not express any concern when prompted to do so in a short anamnestic questionnaire. ASD traits pertaining to communication (C), reciprocal social interaction (S) and restricted, repetitive and stereotyped behaviour (R) were screened for by means of the Social Communication Questionnaire (SCQ, Rutter et al., 2003), with reference to their occurrence in the 12 months preceding administration. The screening was considered positive if a child reached or exceeded a total score of 11 out of 39 (Allen et al., 2007).

2.2 Participants

Fifty-five families of children with a clinical or working diagnosis of ASD were recruited from home-guidance centres in Flanders, Belgium, as part of a randomised-controlled trial on the effects of early social-communicative interventions for children with ASD. Those presented here are pre-intervention data. Inclusion criteria were referral from a home-guidance centre and an ADOS-2 CSS score above the cut-off for ASD. Of the initial sample, 14 children (5 ASD/noDD, 9 ASD/DD) refused to wear the EEG cap, 5 children did not produce sufficient artifact-free data (1 ASD/noDD, 4 ASD/DD), and 2 children did not meet the inclusion criteria based on ADOS-2 CSS scores. For one girl with ASD/DD, the family disclosed use of anticonvulsant medication; however, her data were included in the analysed sample to reflect the heterogeneity of the population. Post-hoc tests showed that group differences are not attributable to measures collected from this child (see supplementary materials S6).

The 34 children included in the final sample were assigned to one of two groups based on their ratio DQ score: those with a score of 71 or higher were considered as having no comorbid DD (ASD/noDD), whereas those with a DQ of 70 or lower (World Health Organization, 2018) were considered as having comorbid developmental delay (ASD/DD).

In the final sample, the ASD/noDD group included 12 children (4 girls, mean age: 39.72, SD: 7.54, range: 28.60 – 49.26; mean DQ: 88.29, SD: 12.77, range: 71 – 111; mean ADOS-2 CSS: 6.58, SD; 1.93, range: 4 – 10). The ASD/DD group consisted of 22 children (2 girls, mean age: 34.37, SD: 7.11, range: 21.00 – 48.13; mean DQ: 45.73, SD: 11.77, range: 21 – 65; mean ADOS-2 CSS: 7.90, SD: 1.73, range: 4 – 10).

Partly because of the young age of the participants, a considerable number of them (4 ASD/noDD, 6 ASD/DD) was being assessed upon a diagnostic suspicion of ASD, but had not yet received a clinical diagnosis at the time of their site visit. All the children in the ASD/DD and ASD/noDD groups received a clinical diagnosis of ASD before or during the last of three site visits foreseen by the protocol of the clinical trial to which they took part.

The families of 25 children without clinically relevant peculiarities were recruited through advertising in the local community. 5 TD children refused to start the EEG session, whereas 3 children did not

produce sufficient artifact-free data. This left a final sample of 17 children (5 girls, mean age: 38.21 months, SD: 10.99, range: 20.07 – 57.07; mean SCQ total score: 6.82, SD: 1.85, range: 4 – 9).

As expected, the criterion for group assignation determined a significant difference in DQ between ASD/noDD and ASD/DD ($t(32) = 9.254$, $p = . 0000000001$), which was observed both in the non-verbal ($t(32) = 7.739$, $p = .00000001$), and verbal domains ($t(32) = 7.457$, $p = .00000002$).

Besides the difference in developmental measures between the ASD/DD and ASD/noDD, preliminary tests on demographic and behavioural measures show that the three groups did not differ in sex ratio ($X2(2, N = 51) = 3.625$, $p = .163$) or age ($F(2,48) = 1.550$, $p = .223$). Tests were also run for pairwise comparisons between the two ASD groups, which showed the largest differences. The difference in sex ratio between ASD/DD and ASD/noDD was marginally significant ($X2(1, N = 34) = 3.140$, $p = .076$), and so was the difference in age ($t(32) = -1.880$, $p = .069$).

Finally, difference in symptom severity was preliminarily tested for the two ASD groups. ADOS-CSS scores tended to be higher in ASD/DD than in ASD/noDD children ($U(NASD/DD = 22, NASD/noDD = 12) = 82.00$, $z = 1.783$, $p = .075$), particularly for the SA domain ($U(NASD/DD = 22, NASD/noDD = 12) = 70.50$, $z = 2.198$, $p = .027$). No significant group difference was observed for RRB traits ($U(NASD/DD = 22, NASD/noDD = 12) = 108.00$, $z = .846$, $p = .395$).

The recruitment and experimental procedures were approved by the local Ethical Committee. Parents or guardians gave informed consent conform the Helsinki Declaration, and children were offered age-appropriate books or toys as a compensation for their participation to the study. Table 1 shows a summary of the demographic and relevant behavioural characteristics of the three samples of children.

2.3 Experimental procedure

EEG was recorded at 500 Hz sampling frequency with an active electrodes set-up (actiCHamp, Brain Products, Germany), with 9 recording sites, based on the international 10-20 system (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4). The TP9 site served as the implicit reference, while the ground electrode was placed at the AFz site. Extra recordings were collected from TP10, to allow for offline re-referencing to

the linked mastoids, and EOG was collected from Fp1, AF7 and AF8. Impedance was kept below 40KΩ.

The experimental paradigm was a passive oddball task based on the one used by Whitehouse and Bishop (2008): participants listened to standard and deviant auditory stimuli, presented at 60 dB volume, while sitting independently or on the parent's lap in an electrically shielded room with dimmed lights and watching a silent video. Stimulus duration was 200 ms, including a 5 ms rising and falling phases at the beginning and at the end of each sound, and stimuli were presented with a 700 ms SOA. The deviant stimuli in the oddball were a speech stimulus (vowel /i/, with f0 at 186 Hz and formants at 293 Hz, 2646 Hz, 3571 Hz, and 4130 Hz) and a non-speech stimulus (a tone composed of sine waves with frequencies of 800 Hz, 1600 Hz, 3200, and 4000 Hz), which were presented in a series of standard speech stimuli (vowel /a/, with f0 at 186 Hz and formants at 837 Hz, 1404 Hz, 3050 Hz, and 4204 Hz) or non-speech stimuli (a complex tone composed of sine waves with frequencies of 500 Hz, 1000 Hz, 1500 Hz, and 2000 Hz). Complex tones were generated using Praat software (Boersma & Wennick, 2016), whereas speech stimuli were edited from real voice sounds, pronounced at constant pitch by a trained speaker. The paradigm included thus 4 conditions of deviant stimuli: deviant tone in a background of standard tones (Tone-T), deviant vowel in a background of standard tones (Vowel-T) deviant tone in a background of standard vowels (Tone-V), and deviant vowel in a background of standard vowels (Vowel-V).

Stimuli were presented in a pseudo-random order, with at least two standard stimuli presented after each deviant. A total of 60 deviant stimuli were presented for each of the four conditions (10% occurrence). Standard stimuli included 480 stimuli for each of the two conditions (80% occurrence); however, standard stimuli following deviant were excluded from analysis, leaving 360 standard stimuli per condition.

The experiment consisted of 1200 stimuli, divided into four blocks (2 with speech standard stimuli and 2 with non-speech standard stimuli), the order of which was counterbalanced between participants. The total administration time was 14 minutes.

The speech standard block was heard as the first by 9 TD participants, 6 participants with ASD/noDD, 11 participants with ASD/DD, whereas the remaining children (8 TD, 6 ASD/noDD, and 11 ASD/DD) heard the non-speech block first.

2.4 EEG data processing

EEG was processed using FieldTrip toolbox (Oostenveld et al., 2011). Continuous data were re-referenced at the linked mastoids and band-passed at 0.1-35 Hz. For trial definition, baseline correction was applied for the 100 ms preceding stimulus onset, and the 600 ms epoch after stimulus onset was considered for ERP analysis.

Vertical EOG was computed as the difference between the recording at Fp1 and the average reference, computed over the nine recording sites mentioned above. Horizontal EOG was computed as the difference between the recordings at AF7 and AF8. Trials with a voltage exceeding a range of +/-150 µV or a standard deviation of 55 µV across time-points were considered as contaminated by artifacts and excluded from analysis.

2.5 ERP analysis

Statistical analysis was conducted on the average ERP response to standard stimuli, and on the difference wave (deviant – standard) in response to deviant stimuli, using FieldTrip toolbox (Oostenveld et al., 2011) and in-house scripts for Matlab software (The MathWorks Inc, 2016). Although our hypothesis concerned mean voltages, group differences in latency were also exploratively addressed. The latency of the components of interest was calculated based on fractional area measures (Kiesel et al., 2008), and mean voltages were calculated for 100 ms long time-windows centred at the mean fractional area latency for each measure.

For standard stimuli, the P1 component was identified as a positive deflection occurring between 50 and 250 ms after stimulus onset, and fractional area latency was calculated as the time-point for which the area between the ERP curve and the x axis reached 50% of its value. For the N2 component, the area was defined by a negative deflection between 200 and 400 ms and the x axis.

For deviant stimuli, the MMN component was defined as a deflection towards negative values in the difference ERP between 100 and 300 ms after stimulus onset, and the P3a as a deflection towards positive values in the 250-450 ms time-window after stimulus onset. In order to define the latency of deflections that did not cross the x axis, areas of interest were calculated as the ones delimited by the ERP deflection and a horizontal axis with intercept between the maximum and minimum ERP value in

the relevant time-window. Based on this, fractional area latency was defined as the time-point for which the areas of interest reached 50% of their value.

The final time-windows for primary and secondary ERP components, based on mean fractional area latency, were 110 to 210 ms for P1, 266 to 366 ms for N2, 163 to 263 ms for MMN and 286 ms to 386 ms for P3a. Mean voltages were calculated over the mentioned time-windows for the four components of interest.

Because of the fronto-central distribution of the ERP components of interest, statistical analysis was run on the averaged signal over a region of interest (ROI) consisting of midline electrodes Fz and Cz.

Participants were included in the grand average with a minimum of 12 artifact-free trials per condition. Table 2 shows the mean number of trials produced by the three groups in each condition. No significant difference in number of trials was found within or between groups.

Statistics were conducted for each component of interest using permutation-based mixed design ANOVA (Maris, 2012). Planned comparisons were limited to interactions involving group, and to the main effect of group. For standard stimuli, both for latency and for amplitude, the Stimulus x Group and group comparisons were tested for the P1 and N2 components. For deviant stimuli, the comparisons were Background x Stimulus x Group, Background x Group, Stimulus x Group, and Group. These were calculated for latency and amplitude of the MMN and P3a components. Testing the hypothesis on group differences in fractional area latency required 4 ANOVAs, with Bonferroni-corrected alpha = .0125. The same alpha level was applied to the 4 ANOVAs in which group differences in mean voltages were tested.

3. Results

Mean ERP responses to standard tones, averaged over Fz and Cz sites, are displayed In Figure 1. Figure 2 shows difference ERPs elicited by deviant tones, also averaged over the two frontocentral midline sites.

3.1 Mean fractional area latency of ERP components

Mean fractional area latency values are reported in Table 3 and Table 4. No statistical comparison involving Group was significant at the Bonferroni-corrected level for the primary (Table 5) or difference ERP (Table 6) components of interest. The results of exploratory correlation analysis on mean fractional area latency values are reported the supplementary materials (Table S1 and S2).

3.2 Mean voltages of ERP components

Analysis of P1 and N2 mean voltages to standard tones did not reveal any main effect of group, or stimulus x group interaction (Table 8). Table 7 shows the mean P1 and N2 voltages for the three groups.

The analysis of difference ERP showed a main effect of group for MMN mean voltages ($F(2,48) = 7.545$, $p = .001$, $\alpha = .0125$, $\eta2 = .198$). Post-hoc pairwise comparisons with two-tailed independent-samples t-test show that the group difference is driven by children with ASD/DD showing overall more negative MMN voltages than TD ($t(37) = -4.012$, $p = .0008$, $\alpha = .016$, Cohen's $d = -1.31$). The comparison between ASD/noDD and TD was non-significant in spite of a medium effect size ($t(27) = -1.971$, $p = .055$, $\alpha = .016$, Cohen's $d = -.75$). The comparison between ASD/DD and ASD/noDD was also non-significant ($t(32) = -1.329$, $p = .192$, $\alpha = .016$, Cohen's $d = -.48$). Mean MMN voltages across conditions were -6.18 µV (3.84) for ASD/DD, -4.31 µV (3.69) for ASD/noDD, and -1.54 µV (4.84) for TD. One-sample single-tailed permutation-based t-tests show that the mean MMN voltages averaged across conditions are significantly below 0 for the two ASD groups (ASD/DD: $t(33) = -7.543$, $p = .0001$, $\alpha = .05$, Cohen's $d = -1.60$; ASD/noDD: $t(11) = -4.048$, $p = .001$ $\alpha = .05$, Cohen's $d = -1.16$), but not for the TD group ($t(16) = -1.316$, $p = .098$ $\alpha = .05$, Cohen's $d = -.31$). Mean MMN voltages for all conditions and groups are reported in Table 9. See Table 10 for other omnibus comparisons based on MMN mean voltages.

For P3a mean voltages, there was a significant Background x Stimulus x Group interaction ($F(2,48) = 4.747$, $p = .011$, $\alpha = .0125$, $\eta2 = .089$), which we followed-up by analysing the Background x Group interaction separately for deviant tones ($F(2,48) = 5.333$, $p = .009$, $\alpha = .025$, $\eta2 = .182$) and deviant vowels interaction ($F(2,48) = .203$, $p = .819$, $\alpha = .025$, $\eta2 = .024$). Only the former comparison was

significant, and was followed-up by testing group differences separately for deviant tones in tones ($F(2,48) = .347$, $p = .699$, $\alpha = .025$, $\eta 2 = .014$) and deviant tones in vowels ($F(2,48) = 6.172$, $p = .004$, $\alpha = .025$, $\eta 2 = .205$). The result of the follow-up analysis suggests that the initial 3-ways interaction was driven by group differences in P3a mean voltages in response to deviant tones in vowels (Tone-V). Post-hoc pairwise two-tailed independent samples t-test showed that children with ASD/noDD differed significantly both from children with ASD/DD ($t(32) = -3.396$, $p = .002$, $\alpha = .016$, Cohen's $d = -1.23$) and from TD children ($t(27) = -3.080$, $p = .006$, $\alpha = .016$, Cohen's $d = -1.18$), whereas children with ASD/DD did not differ from TD ($t(37) = .118$, $p = .903$, $\alpha = .016$, Cohen's $d = .04$). Mean P3a voltages elicited by Tone-V stimuli were 6.13 µV (8.72) for TD children, 6.46 µV (8.39) for children with ASD/DD, and -3.15 µV (6.83) for children with ASD/noDD. One-sample single-tailed permutation-based t-tests show that the ASD/DD ($t(33) = 3.610$, $p = .002$, $\alpha = .05$, Cohen's $d = .76$) and TD groups ($t(16) = 2.902$, $p = .005$, $\alpha = .05$, Cohen's $d = .70$) have a consistently positive P3a response to Tone-V stimuli, whereas this was not the case for children with ASD/noDD ($t(11) = -1.601$, $p = .925$, $\alpha = .05$, Cohen's $d = -.46$) . See Table 9 for mean P3a voltages for all conditions and groups, and Table 10 for other omnibus comparisons on P3a mean voltages.

Correlation across mean voltages, age, and behavioural measures was explored and results are reported in Table S3 and S4 (supplementary materials). Because of the marginally significant age and sex difference between ASD/DD and ASD/noDD, the analysis on P3a mean voltages, for which a difference between the two ASD groups was found, we run again while controlling for age. This did not change the result pattern (see section S5 in the supplementary materials). An additional post-hoc test also excluded that group differences were inflated by the inclusion of data from one child taking anticonvulsant medication (section S6 in the supplementary materials).

4. Discussion

In the present study we tested the hypothesis that diminished attentional orienting to sound change in speech is a peculiarity of individuals with ASD without comorbid intellectual disability or developmental delay, who constitute the vast majority of the participants of the studies that addressed the question of attentional orienting in ASD. To this purpose, we recruited children with ASD with and without comorbid developmental delay, as well as typically developing children.

Our main finding concerns the different levels of attentional orienting, indexed by the P3a component, observed in children with ASD/DD and ASD/noDD in spite of their similar perceptual sensitivity. Contrarily to our hypothesis, the different DQ levels in the two ASD groups did not affect their sensory processing of deviant sounds. This is at odds with the meta-analytic results reported by Schwartz et al. (2018), who suggest that MMN is attenuated in participants with ASD and intellectual disability, in younger children, and particularly in response to non-speech sounds. Attenuated MMN voltages in participants with lower DQ emerged also from Chen's et al. meta-analysis (2020), and lower verbal ability was associated with attenuated MMN in the large sample of participants in the MEG study by Matsuzaki et al. (2019). In our sample, the young age of the children, and possibly sensitivity issues non-targeted by the behavioural tests and questionnaires administered, together with the use of a single-feature experimental manipulation, may have contributed to the similar way in which ASD/DD and ASD/noDD differed from TD. Moreover, as we were particularly interested in P3a differences across groups, we limited our analysis to frontocentral midline sites, which may have left lateralisation differences in the topographic distribution of the MMN component undetected. A trend towards an enhanced (more negative) MMN in ASD compared to TD was also observed in puberal children by Whitehouse & Bishop (2008), from whose study we adapted the current paradigm.

The enhanced MMN that we observed in both ASD groups may fit into the predictive coding account of perception in ASD, according to which a weak prior or expectation is formed based on sensory input (Pellicano & Burr, 2012; Van Boxtel & Lu, 2013; van de Cruys et al., 2014). As a consequence, sensory stimuli in ASD are perceived as unexpected to a larger extent than in typical development. Predictive coding would maintain that, due to enhanced perception, the hierarchic attribution of saliency to sensory input be challenged, thereby affecting attentional orienting to relevantly deviant stimuli. However, whilst a reduced P3a component following an atypically negative MMN was observed in children with ASD/noDD in response to deviant tones in speech, this atypicality was not present in children with ASD/DD, who showed typical levels of attentional orienting even in presence of enhanced perception. As one would expect perceptual aberrant precision (Lawson et al., 2014) to challenge ASD/DD children at least to the same extent as ASD/noDD children, it seems appropriate to integrate the predictive coding perspective with a neural adaptation framing of the current findings.

In children with more efficient cognitive functioning, neural adaptation may be involved in the downregulation of the response to challenging stimuli, such as social ones, or those presented in a

social context, thereby contributing to neurophysiological homeostasis. Alternatively, it could be the more efficient allocation of cognitive resources in children who downregulate social stimulus processing to result in a more favourable developmental outcome. Finally, there could also be an interplay between the two processes, with unimpaired cognitive functioning allowing for selective attention and vice-versa (Johnson et al., 2015).

An important consideration is that the two ASD groups differed ADOS-2 CSS scores, particularly in the social affect domain. Group differences in attentional orienting to stimuli in a social context remained significant when this possible confound was controlled for. Nevertheless, the relationship between DQ and symptom severity, well- known in the ASD population (Mayes & Calhoun, 2011; Vivanti et al., 2013), might paradoxically suggest that reduced attentional orienting to stimuli in a social context is related to milder social affect symptoms, similarly to what reported by Herrington et al. (2016), who observed reduced sensitivity to social stimuli in individuals with ASD and a less severe clinical situation (i.e., low anxiety traits).

Finally, a different interpretative framework would suggest linking the different P3a levels observed in the two ASD groups to age rather than DQ differences. As a matter of fact, although neither ASD group differed from the TD group in terms of age, children with ASD/DD tended to be younger than their peers with no developmental delay. Running statistical analysis while controlling for age did not change the observed results, suggesting that group differences in ERP measures are unlikely due to differences in age. Yet, for children with ASD, a negative trend was present in the relationship between age and P3a mean voltages to deviant tones in speech (see Table 3 in the supplementary materials). This suggests that the interrelated character of attentional orienting, verbal ability, and age, may speak to the longitudinal evolution of adaptive processes in ASD, with neural adaptation providing an increasing level of constrained optimisation of cognitive resources over time, albeit leading to more atypical brain function. This would be comparable to the more atypical brain response to hearing their own name found in toddlers than in infants with an elevated likelihood of developing ASD (Arslan et al., 2020), which also suggests that developmental pathways in ASD and TD diverge increasingly over time.

Although the current findings on children with ASD/noDD derive from a small sample, which represents an important limitation, they do resonate with previous reports on similar issues. A

relationship between diminished P3 effect and higher cognitive and linguistic abilities was also found in children and adolescents with ASD taking part in an active oddball task with a speech and a non-speech target deviants occurring amid non-speech standard sounds (Dawson et al., 1988), which would also fit in a neural adaptation account of a reduced P3a in ASD. Finally, it is worth looking at our results together with those of Green et al. (2015), who found that a pre-emptive positive parenting intervention for 7-10 months old siblings of children with ASD likely resulted in the reduction of P1 and P3 voltages in response to vowel change, in the face of a positive behavioural outcome based on social attention measures. On a similar note, the findings of Vlaskamp et al. (2017) point to a possible relationship between MMN attenuation and less severe symptomatology in puberal children with ASD, whilst Gomot et al. (2011) found enhanced P3a in children with higher intolerance to change (see also the relation between P3a and RRB traits in table S3).

Not having collected DQ measures for TD children represents a further limitation of the current study. As a matter of fact, knowing whether or not the relation between higher DQ and attenuated attentional orienting in social contexts is specific to ASD would corroborate the interpretative framework for this finding.

Accentuated atypicalities in participants with ASD and higher cognitive abilities have also been found in specific behavioural domains. For instance, attention to faces (Norbury et al., 2009) and use of pragmatic prosody (DePape et al., 2012), were found to be more aberrant in children with ASD and normal verbal IQ than in their peers with lower verbal IQ.

Implications

The present study adds to the evidence that ASD is a highly heterogeneous and stratified disorder. Brain activity measures, including ERPs, may be helpful in identifying subgroups within the spectrum, and to predict cognitive and behavioural outcome. However, atypical brain measures are not necessarily associated with more severe developmental outcome, as in some cases they may stem from adaptive processes rather than from mere impairment.

References

Allen, C. W., Silove, N., Williams, K., & Hutchins, P. (2007). Validity of the Social Communication Questionnaire in Assessing Risk of Autism in Preschool Children with Developmental Problems. *Journal of Autism and Developmental Disorders*, 37(7), 1272–1278. https://doi.org/10.1007/s10803-006-0279-7

American Psychiatric Association. (2013). Neurodevelopmental disorders. In *Diagnostic and statistical manual of mental disorders*. Fifth edition. https://doi.org/10.1176/appi.books.9780890425596.dsm01

Arslan, M., Warreyn, P., Dewaele, N., Wiersema, J. R., Demurie, E., & Roeyers, H. (2020). Development of neural responses to hearing their own name in infants at low and high risk for autism spectrum disorder. *Developmental Cognitive Neuroscience*, 41, 100739. https://doi.org/10.1016/j.dcn.2019.100739

Boersma, P., & Wennick, D. (2016). *Praat: doing phonetics by computer* [Computer program]. Version 6.0.22, retrieved 15 November 2016 from http://www.praat.org

Čeponienė, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., & Yaguchi, K. (2003). Speech–sound-selective auditory impairment in children with autism: They can perceive but do not attend*. Proceedings of the National Academy of Sciences of the United States of America*, 100(9), 5567–5572. https://doi.org/10.1073/pnas.0835631100

Charpentier, J., Kovarski, K., Houy-Durand, E., Malvy, J., Saby, A., Bonnet-Brilhault, F., Latinus, M., & Gomot, M. (2018). Emotional prosodic change detection in autism Spectrum disorder: An electrophysiological investigation in children and adults. *Journal of Neurodevelopmental Disorders*, 10(1), 28. https://doi.org/10.1186/s11689-018-9246-9

Chen, T.-C., Hsieh, M. H., Lin, Y.-T., Chan, P.-Y. S., & Cheng, C.-H. (2020). Mismatch negativity to different deviant changes in autism spectrum disorders: A meta-analysis. *Clinical Neurophysiology*, 131(3), 766–777. https://doi.org/10.1016/j.clinph.2019.10.031

Dawson, G., Finley, C., Phillips, S., Galpert, L., & Lewy, A (1988). Reduced P3 amplitude of the event-related brain potential: its relationship to language ability in autism. *Journal of Autism and Developmental Disorder*, 18(4), 493–504. https://doi.org/10.1007/BF02211869

DePape, A.M., Chen, A., Hall, G.B.C., & Trainor, L. J. (2012). Use of prosody and information structure in high functioning adults with autism in relation to language ability. *Frontiers in Psychology*, 26(3), 72. https://doi.org/10.3389/fpsyg.2012.00072

Esler, A.N., Hus Bal, V., Guthrie, W., Wetherby, A., Weismer, S.E., & Lord, C. (2015). The Autism Diagnostic Observation Schedule, Toddler Module: Standardized Severity Scores. *Journal of Autism and Developmental Disorders*, 45(9), 2704–2720. https://doi.org/10.1007/s10803-015-2432-7

Fan, Y.T., & Cheng, Y. (2014). Atypical mismatch negativity in response to emotional voices in people with autism spectrum conditions. *PloS One*, 9(7), e102471. https://doi.org/10.1371/journal.pone.0102471

Farroni, T., Chiarelli, A. M., Lloyd-Fox, S., Massaccesi, S., Merla, A., Di Gangi, V., Mattarello, T., Faraguna, D., & Johnson, M. H. (2013). Infant cortex responds to other humans from shortly after birth. *Scientific Reports*, 3, 2851. https://doi.org/10.1038/srep02851

Goldstein, A., Spencer, K. M., & Donchin, E. (2002). The influence of stimulus deviance and novelty on the P300 and novelty P3. Psychophysiology, 39(6), 781–790. https://doi.org/10.1017/S004857720201048X

Gomot, M., Blanc, R., Clery, H., Roux, S., Barthelemy, C., & Bruneau, N. (2011). Candidate Electrophysiological Endophenotypes of Hyper-Reactivity to Change in Autism. *Journal of Autism and Developmental Disorders*, 41(6), 705–714. https://doi.org/10.1007/s10803-010-1091-y

Green, J., Charman, T., Pickles, A., Wan, M.W., Elsabbagh, M., Slonims, V., Taylor, C., Janet McNally, J., Booth, R., Gliga, T., Jones, E.J.H., Harrop, C., Bedford, R., Johnson, M. H., & the BASIS team (2015). Parent-mediated intervention versus no intervention for infants at high risk of autism: a parallel, single-blind, randomised trial. *Lancet Psychiatry*, 2, 133–140. https://doi.org/10.1016/S2215-0366(14)00091-1

Herrington, J., Miller, J., Pandey, J., & Schultz, R. (2016). Anxiety and social deficits have distinct relationships with amygdala function in autism spectrum disorder. *Social Cognitive and Affective Neuroscience*, 11(6), 907-914. https://doi.org/10.1093/scan/nsw015

Hus, V., Gotham, K., & Lord, C. (2014). Standardizing ADOS Domain Scores: Separating Severity of Social Affect and Restricted and Repetitive Behaviors. *Journal of Autism and Developmental Disorders*, 44(10), 2400–2412. https://doi.org/10.1007/s10803-012-1719-1

Johnson, M.H. (2017). Autism as an adaptive common variant pathway for human brain development. Developmental Cognitive Neuroscience, 25, 5–11. https://doi.org/10.1016/j.dcn.2017.02.004

Johnson, M.H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40(1–2), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

Johnson, M.H., Jones, E.J., & Gliga, T. (2015). Brain Adaptation and Alternative Developmental Trajectories. *Development and Psychopathology*, 27, 425–442.

Jones, E.J.H., Gliga, T., Bedford, R., Charman, T., & Johnson, M.H. (2014). Developmental pathways to autism: A review of prospective studies of infants at risk. *Neuroscience and Biobehavioural Reviews*, 39(100), 1–33. https://doi.org/10.1016/j.neubiorev.2013.12.001.

Kiesel, A., Miller, J., Jolicœur, P., & Brisson, B. (2008). Measurement of ERP latency differences: A comparison of single-participant and jackknife-based scoring methods. *Psychophysiology*, 45(2), 250–274. https://doi.org/10.1111/j.1469-8986.2007.00618.x

Kilpeläinen, R., Luoma, L., Herrgård, E., Sipilä, P., Yppärilä, H., Partanen, J, & Karhu, J. (1999). Distractible children show abnormal orienting to non-attended auditory stimuli. *NeuroReport*, 10(9), 1869–1874. https://doi.org/10.1097/00001756-199906230-00013

Kilpeläinen, R., Luoma, L., Herrgård, E., Yppärilä, H., Partanen, J., Karhu, J. (1999). Persistent frontal P300 brain potential suggests abnormal processing of auditory information in distractible children. *NeuroReport*, 10(16), 3405–3410. https://doi.org/10.1097/00001756-199911080-00027

Kisilevsky, B. S., Hains, S. M. J., Brown, C. A., Lee, C. T., Cowperthwaite, B., Stutzman, S. S., Swansburg, M. L., Lee, K., Xie, X., Huang, H., Ye, H.-H., Zhang, K., & Wang, Z. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior and Development*, 32(1), 59–71. https://doi.org/10.1016/j.infbeh.2008.10.002

Kujala, T., Aho, E., Lepistö, T., Jansson-Verkasalo, E., Nieminen-von Wendt, T., von Wendt, L., & Näätänen, R. (2007). Atypical pattern of discriminating sound features in adults with Asperger syndrome as reflected by the mismatch negativity. *Biological Psychology*, 75(1), 109–114. https://doi.org/10.1016/j.biopsycho.2006.12.007

Lawson, R.P., Rees, G., & Friston, K.J. (2014) An aberrant precision account of autism. *Frontiers in Human Neuroscience*, 8, 302. https://doi.org/10.3389/fnhum.2014.00302

Lepistö, T., Kajander, M., Vanhala, R., Alku, P., Huotilainen, M., Näätänen, R., & Kujala, T. (2008). The perception of invariant speech features in children with autism. *Biological Psychology*, 77(1), 25–31. https://doi.org/10.1016/j.biopsycho.2007.08.010

Lepistö, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., & Näätänen, R. (2005). The discrimination of and orienting to speech and non-speech sounds in children with autism. *Brain Research*, 1066(1–2), 147–157. https://doi.org/10.1016/j.brainres.2005.10.052

Lepistö, T., Nieminen-von Wendt, T., von Wendt, L., Näätänen, R., & Kujala, T. (2007). Auditory cortical change detection in adults with Asperger syndrome. *Neuroscience Letters*, 414(2), 136–140. https://doi.org/10.1016/j.neulet.2006.12.009

Lepistö, T., Silokallio, S., Nieminen-von Wendt, T., Alku, P., Näätänen, R., & Kujala, T. (2006). Auditory perception and attention as reflected by the brain event-related potentials in children with Asperger syndrome. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 117(10), 2161–2171. https://doi.org/10.1016/j.clinph.2006.06.709

Lindström, R., Lepistö-Paisley, T., Makkonen, T., Reinvall, O., Nieminen-von Wendt, T., Alén, R., & Kujala, T. (2018). Atypical perceptual and neural processing of emotional prosodic changes in children with autism spectrum disorders. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 129(11), 2411–2420. https://doi.org/10.1016/j.clinph.2018.08.018

Lindström, R., Lepistö-Paisley, T., Vanhala, R., Alén, R., & Kujala, T. (2016). Impaired neural discrimination of emotional speech prosody in children with autism spectrum disorder and language impairment. *Neuroscience Letters*, 628, 47–51. https://doi.org/10.1016/j.neulet.2016.06.016

Livingston, L.A., & Happé, F. (2017). Conceptualising compensation in neurodevelopmental disorders: Reflections from autism spectrum disorder. *Neuroscience and Biobehavioral Reviews*, 80, 729–742. https://doi.org/10.1016/j.neubiorev.2017.06.005

Lord, C., Rutter, M., Dilavore, P.C., Risi, S., Gotham, K., & Bishop, S. (2012). Autism diagnostic observation schedule. *Second edition*. (ADOS-2). Torrance, CA: Western Psychological Services.

Ludlow, A., Mohr, B., Whitmore, A., Garagnani, M., Pulvermüller, F., & Gutierrez, R. (2014). Auditory processing and sensory behaviours in children with autism spectrum disorders as revealed by mismatch negativity*. Brain and Cognition*, 86, 55–63. https://doi.org/10.1016/j.bandc.2014.01.016

Maris, E. (2012), Statistical testing in electrophysiological studies, *Psychophysiology*, 49(4):549-565. https://doi.org/10.1111/j.1469-8986.2011.01320.x

Määttä, S., Herrgård, E., Saavalainen, P., Pääkkönen, A., Könönen, M., Luoma, L., Laukkanen, E., Yppärilä, H., & Partanen, J. (2005). P3 amplitude and time-on-task effects in distractible adolescents. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 116(9), 2175–2183. https://doi.org/10.1016/j.clinph.2005.06.014

Matsuzaki, J., Kuschner, E. S., Blaskey, L., Bloy, L., Kim, M., Ku, M., Edgar, J., Embick, D., & Roberts, T. P. L. (2019). Abnormal auditory mismatch fields are associated with communication impairment in both verbal and minimally verbal/nonverbal children who have autism spectrum disorder*. Autism Research: Official Journal of the International Society for Autism Research*, 12(8), 1225–1235. https://doi.org/10.1002/aur.2136

Mayes, S. D., & Calhoun, S. L. (2011). Impact of IQ, age, SES, gender, and race on autistic symptoms. *Research in Autism Spectrum Disorders*, 5(2), 749–757. https://doi.org/10.1016/j.rasd.2010.09.002

Messinger, D., Young, G. S., Ozonoff, S., Dobkins, K., Carter, A., Zwaigenbaum, L., Landa, R. J., Charman, T., Stone, W. L., Constantino, J. N., Hutman, T., Carver, L. J., Bryson, S., Iverson, J. M., Strauss, M. S., Rogers, S. J., & Sigman, M. (2013). Beyond Autism: A Baby Siblings Research Consortium Study of High-Risk Children at Three Years of Age. *Journal of the American Academy of Child and Adolescent Psychiatry*, 52(3), 300. https://doi.org/10.1016/j.jaac.2012.12.011

Mullen, E.M. (1995). *Mullen Scales of Early Learning*. AGS ed. Circle Pines, MN: American Guidance Service Inc.

Mundy, P. (2018). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *European Journal of Neuroscience*, 47(6), 497–514. https://doi.org/10.1111/ejn.13720

Norbury, C.F., Brock, J., Cragg, L., Einav, S., Griffiths, H., & Nation, K. (2009). Eye-movement patterns are associated with communicative competence in autistic spectrum disorders. Journal of *Child Psychology and Psychiatry*, 50(7), 834–842. https://doi.org/10.1111/j.1469-7610.2009.02073.x.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*, 2011, 156869, https://doi.org/10.1155/2011/156869

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, 16(10), 504–510. https://doi.org/10.1016/j.tics.2012.08.009

Philip, R., Dauvermann, M., Whalley, H., Baynham, K., Lawrie, S., & Stanfield, A. (2012), A systematic review and meta-analysis of the fMRI investigation of autism spectrum disorders. *Neuroscience and Biobehavioral Reviews*, 36(2), 901-942. https://doi.org/10.1016/j.neubiorev.2011.10.008

Polich, J. (2007). Updating P300: An Integrative Theory of P3a and P3b. Clinical Neurophysiology : *Official Journal of the International Federation of Clinical Neurophysiology*, 118(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Ponton, C. W., Eggermont, J. J., Kwong, B., & Don, M. (2000). Maturation of human central auditory system activity: Evidence from multi-channel evoked potentials. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 111(2), 220–236. https://doi.org/10.1016/s1388-2457(99)00236-9

Rutter, M., Bailey, A., & Lord, C. (2003). *The social communication questionnaire*. Los Angeles, CA: Western Psychological Services.

Schwartz, S., Shinn-Cunningham, B., & Tager-Flusberg, H. (2018). Meta-analysis and systematic review of the literature characterizing auditory mismatch negativity in individuals with autism. *Neuroscience & Biobehavioral Reviews*, 87, 106–117. https://doi.org/10.1016/j.neubiorev.2018.01.008

Sirri, L., Linnert, S., Reid, V., & Parise, E. (2020). Speech Intonation Induces Enhanced Face Perception in Infants. *Scientific Reports* 10(1), 3225. https://doi.org/ 10.1038/s41598-020-60074-7.

The Math Works Inc. (2016). Matlab (2016a) [Computer software]. https://www.mathworks.com/

Van Boxtel, J.J.A., & Lu, H. (2013). A predictive coding perspective on autism spectrum disorders. *Frontiers in Psychology*, 4. https://doi.org/10.3389/fpsyg.2013.00019

van de Cruys, S., Evers, K., van der Hallen, R., van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review*, 121(4), 649–675. https://doi.org/10.1037/a0037665

Vivanti, G., Barbaro, J., Hudry, K., Dissanayake, C., & Prior, M. (2013). Intellectual development in autism spectrum disorders: New insights from longitudinal studies. *Frontiers in Human Neuroscience*, 7, 354. https://doi.org/10.3389/fnhum.2013.00354

Vlaskamp, C., Oranje, B., Madsen, G. F., Møllegaard Jepsen, J. R., Durston, S., Cantio, C., Glenthøj, B., & Bilenberg, N. (2017). Auditory processing in autism spectrum disorder: Mismatch negativity deficits. *Autism Research*, 10(11), 1857–1865. https://doi.org/10.1002/aur.1821

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The Tuning of Human Neonates' Preference for Speech. *Child Development*, 81(2), 517–527. https://doi.org/10.1111/j.1467-8624.2009.01412.x

Whitehouse, A. J. O., & Bishop, D. V. M. (2008). Do children with autism 'switch off' to speech sounds? An investigation using event-related potentials*. Developmental Science*, 11(4), 516–524. https://doi.org/10.1111/j.1467-7687.2008.00697.x

Woodman, G. F. (2010). A Brief Introduction to the Use of Event-Related Potentials (ERPs) in Studies of Perception and Attention. *Attention, Perception & Psychophysics*, 72(8). https://doi.org/10.3758/APP.72.8.2031

World Health Organization. (2018*). International classification of diseases for mortality and morbidity statistics (11th Revision).* https://icd.who.int/browse11/l-m/en

Wronka, E., Kaiser, J., & Coenen, A. M. L. (2008). The auditory P3 from passive and active three-stimulus oddball paradigm. *Acta Neurobiologiae Experimentalis*, 68(3), 362–372.

Wronka, E., Kaiser, J., & Coenen, A. M. L. (2013). Psychometric intelligence and P3 of the event-related potentials studied with a 3-stimulus auditory oddball task. *Neuroscience Letters*, 535, 110–115. https://doi.org/10.1016/j.neulet.2012.12.012

Yu, L., Fan, Y., Deng, Z., Huang, D., Wang, S., & Zhang, Y. (2015). Pitch Processing in Tonal-Language-Speaking Children with Autism: An Event-Related Potential Study. *Journal of Autism and Developmental Disorders*, 45(11), 3656–3667. https://doi.org/10.1007/s10803-015-2510-x

Table 1. Demographic and behavioural characteristics of the tree groups.

	TD	ASD/DD	ASD/noDD
Sample size	17	22	12
Boys	12	20	8
Girls	5	2	4
Age (months)	38.21 (10.99)	34.37 (7.11)	39.62 (7.54)
DQ	n.a.	45.73 (11.67)	88.29 (12.77)
nvDQ		56.98 (12.91)	95.21 (12.88)
vDQ		34.48 (15.11)	81.36 (20.84)
ADOS-2 CSS score	n.a.	7.90 (1.73)	6.58 (1.93)
SA domain		7.20 (2.42)	5.33 (2.05)
RRB domain		8.20 (1.72)	7.75 (1.57)
SCQ total score	6.82 (1.85)	n.a.	n.a.
S+C domain	4.35 (.90)		
R domain	2.35 (1.30)		

DQ: Full-scale ratio-based developmental quotient. nvDQ: non-verbal developmental quotient. vDQ: verbal developmental quotient. ADOS-2: Autism Diagnostic Observation Scale – 2. CSS: Calibrated Severity Score. SA: Social Affect. RRB: Restricted and Repetitive Behaviour. SCQ: Social Communication Questionnaire. C: communication domain; S: reciprocal social interaction; R: restricted, repetitive and stereotyped behaviour.

Table 2. Mean (SD) number of trials per condition.

	TD	ASD/DD	ASD/noDD	Between-group difference
Standard vowel	185.52 (84.52)	184.72 (89.84)	189.66 (82.27)	F(2,48) = .013, p = .987
Standard tone	173.94 (79.70)	196.90 (87.54)	208.00 (91.06)	F(2,48) = .620, p = .542
Within-group difference	t(16) = .599, p = .557	t(21) = -.955, p = .350	t(11) = -.885, p = .395	F(2,48) = .814†, p = .449
Tone-T	27.88 (12.15)	30.18 (14.93)	29.91 (11.19)	F(2,48) = .159, p = .854
Vowel-T	25.64 (11.35)	30.77 (14.46)	30.58 (10.39)	F(2,48) = .911, p = .409
Tone-V	30.23 (15.07)	28.95 (13.21)	30.25 (10.04)	F(2,48) = .059, p = .943
Vowel-V	29.64 (14.19)	29.36 (14.51)	29.41 (8.73)	F(2,48) = .002, p = .998
Within-group difference	F(1,16) = .639, p = .436	F(1,21) = .214, p = .648	F(1,11) = .027, p = .871	F(2,48) = .507†, p = .606

†Condition x Group comparison.

Table 3. Mean (SD) fractional area latency of primary ERP components to standard stimuli (ms).

	P1			N2		
	TD	ASD/DD	ASD/noDD	TD	ASD/DD	ASD/noDD
Vowel	157.57 (13.19)	164.70 (15.98)	163.66 (10.37)	323.80 (17.68)	308.12 (21.64)	317.00 (17.04)
Tone	156.47 (12.12)	152.45 (26.75)	160.62 (13.70)	318.33 (11.64)	298.45 (27.05)	308.55 (20.87)

Table 4. Mean (SD) fractional area latency of difference ERP components to deviant stimuli (ms).

	MMN			P3a		
	TD	ASD/DD	ASD/noDD	TD	ASD/DD	ASD/noDD
Tone-T	226.34 (31.87)	221.29 (24.26)	212.62 (40.92)	358.12 (31.13)	348.63 (30.05)	354.20 (41.47)
Vowel-T	206.18 (30.99)	223.00 (29.69)	214.96 (38.01)	335.11 (37.41)	342.23 (26.45)	354.27 (42.74)
Tone-V	203.26 (35.16)	216.81 (21.53)	225.35 (15.17)	348.33 (33.98)	351.42 (24.41)	366.86 (34.99)
Vowel-V	189.26 (35.88)	215.28 (40.02)	193.68 (65.84)	290.81 (35.22)	307.28 (36.58)	287.04 (25.94)

Table 5. Mixed-design ANOVA results for latency of ERP components elicited by standard stimuli.

	P1			N2		
	F(2,48)	p	η^2	F(2,48)	p	η^2
Stimulus x Group	1.628	.209	.064	.188	.835	.008
Group	.481	.632	.020	4.996	.014*	.172

*Significant at the uncorrected alpha level (.05).

Table 6. Mixed-design ANOVA results for latency of difference ERP components.

	MMN			P3a		
	F(2,48)	p	η^2	F(2,48)	p	η^2
Background x Stimulus x Group	.926	.397	.037	2.156	. 130	.082
Background x Group	2.105	. 138	.081	.793	. 455	.032
Stimulus x Group	2.232	. 112	.085	1.319	. 270	.052
Group	1.800	. 181	.070	.608	. 541	.025

Table 7. Mean (SD) voltages of ERP components elicited by standard stimuli (μV).

	P1			N2		
	TD	ASD/DD	ASD/noDD	TD	ASD/DD	ASD/noDD
Vowel	7.80 (4.19)	5.47 (5.00)	7.21 (4.86)	-.11 (5.77)	-1.93 (7.12)	-0.59 (6.38)
Tone	5.69 (4.73)	4.16 (4.16)	6.46 (3.88)	-2.15 (6.39)	-3.00 (6.39)	-2.03 (8.55)

Table 8. Mean (SD) voltages of difference ERP components (μV).

	MMN			P3a		
	TD	ASD/DD	ASD/noDD	TD	ASD/DD	ASD/noDD
Tone-T	-2.83 (6.47)	-7.94 (5.74)	-2.53 (6.76)	.89 (7.85)	-0.56 (6.97)	1.49 (7.99)
Vowel-T	-2.30 (8.52)	-5.02 (6.88)	-3.76 (6.01)	.55 (10.26)	1.85 (12.13)	.92 (6.84)
Tone-V	-2.33 (7.86)	-7.31 (9.12)	-9.03 (9.74)	6.13 (8.72)	6.46 (8.39)	-3.15 (6.83)
Vowel-V	1.28 (5.99)	-4.45 (6.34)	-1.92 (8.68)	.90 (7.34)	-3.39 (11.00)	-2.59 (8.76)

Table 9. Mixed-design ANOVA results for mean voltages of ERP components to standard stimuli.

	P1			N2		
	F(2,48)	p	η^2	F(2,48)	p	η^2
Stimulus x group	.313	.745	.013	.202	.818	.008
Group	1.761	.182	.064	.307	. 734	.013

Table 10. Mixed-design ANOVA results for mean voltages of difference ERP components.

	MMN			P3a		
	F(2,48)	p	η^2	F(2,48)	p	η^2
Background x stimulus x group	3.182	.055	.050	4.747	.011**	.089
Background x group	1.254	. 307	.052	3.509	.039*	.067
Stimulus x group	.085	. 902	.004	1.639	.212	.052
Group	7.545	.001**	.198	1.654	.209	.048

*Significant at uncorrected alpha (.05). **Significant at Bonferroni-corrected alpha level (α = .0125).

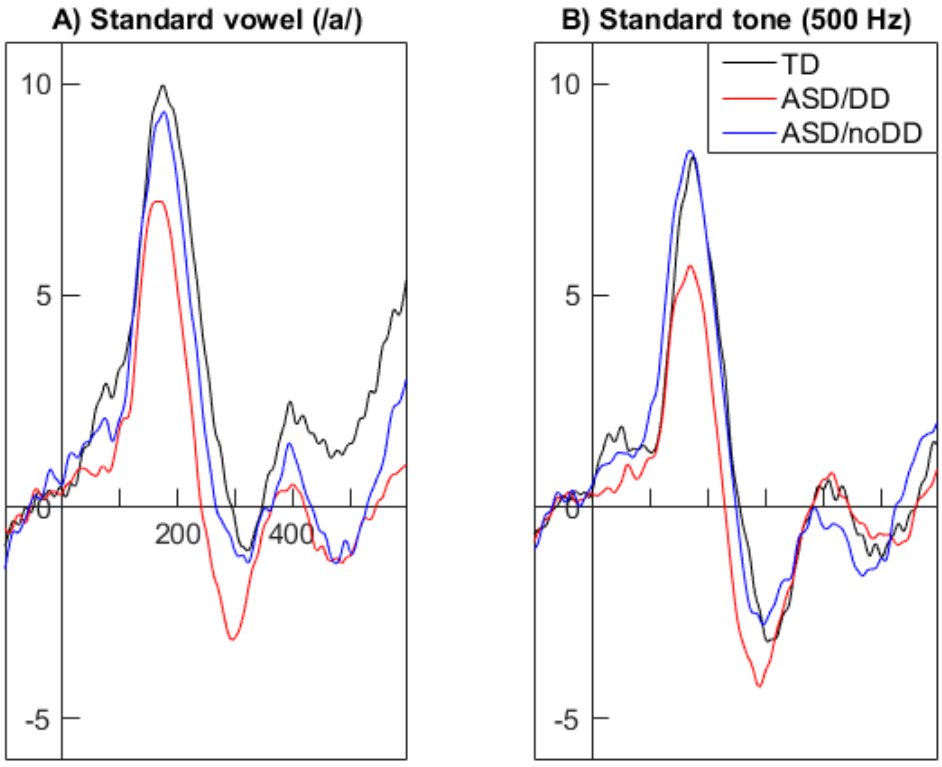

Figure 1. ERP responses to standard vowels and tones for the three groups. y-axis: voltage (µV), x-axis: time (ms). The signal is averaged over Fz and Cz sites.

Figure 2. Difference (deviant minus standard) ERP responses to deviant stimuli in the four conditions for the three groups. y-axis: voltage (µV), x-axis: time (ms). The signal is averaged over Fz and Cz sites.

Figure 2. Difference (deviant minus standard) ERP responses to deviant stimuli in the four conditions for the three groups. y-axis: voltage (µV), x-axis: time (ms). The signal is averaged over Fz and Cz sites.

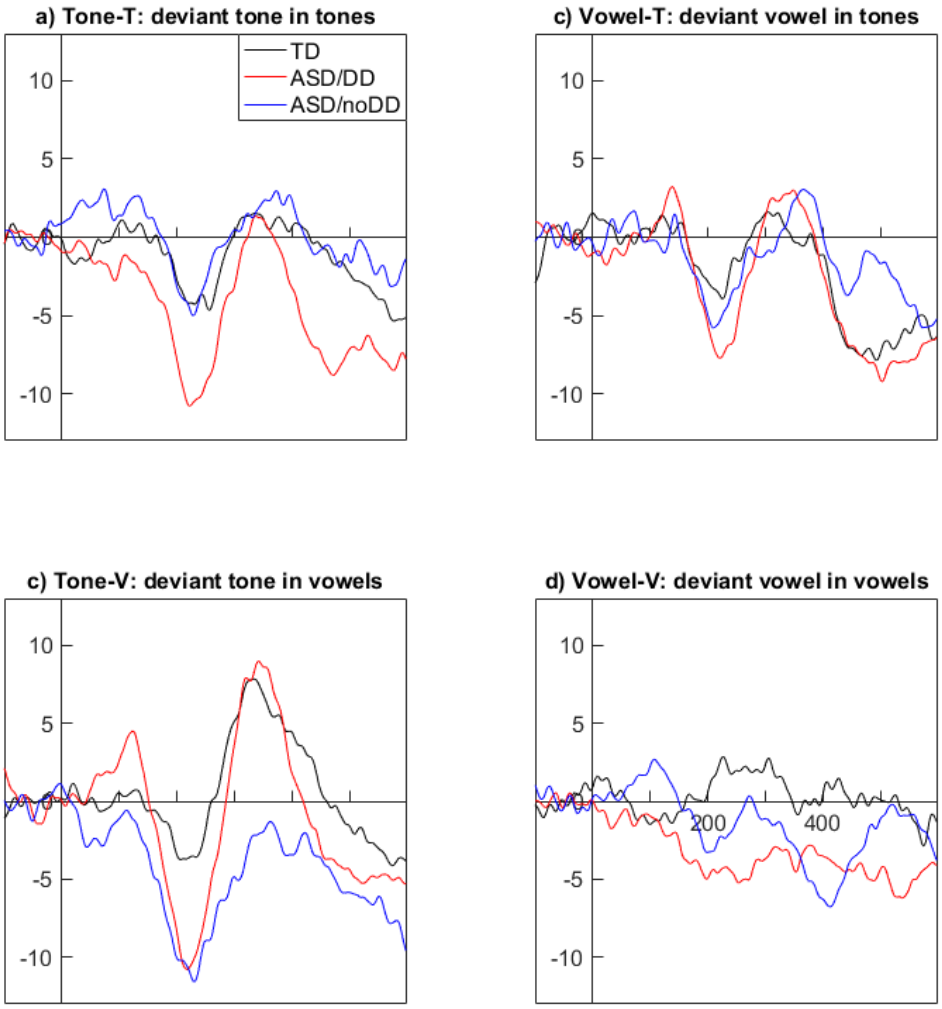

Figure 2. Difference (deviant minus standard) ERP responses to deviant stimuli in the four conditions for the three groups. y-axis: voltage (μV), x-axis: time (ms). The signal is averaged over Fz and Cz sites.

Table S1. Correlations across age, behavioural, and ERP fractional area latency measures in participants with ASD.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.
1. Age	--																		
2. ADOS SA CSS	-.41*	--																	
3. ADOS RRB CSS	.03	.09	--																
4. ADOS CSS	-.44*	.91*	.38*	--															
5. nvDQ	.23	-.22	-.14	-.19	--														
6. vDQ	.32†	-.20	-.03	-.16	.75*	--													
7. DQ	.30†	-.22	-.08	-.18	.92*	.95*	--												
8. P1 standard /a/	.07	-.08	.15	-.11	-.03	-.07	-.06	--											
9. N2 standard /a/	-.29†	.17	-.30†	.04	.21	.15	.19	.04	--										
10. P1 standard tone	.20	-.40*	-.12	-.37*	.22	.13	.18	.36*	.07	--									
11. N2 standard tone	-.41*	-.05	-.26	-.07	.28†	.24	.28†	-.31†	.52*	-.18	--								
12. Tone-T MMN	-.04	.26	.00	.20	-.02	-.15	-.10	.19	-.04	.02	.04	--							
13. Vowel-T MMN	-.12	.07	-.23	.00	-.23	-.31†	-.29†	.20	.11	-.04	.00	.49*	--						
14. Tone-V MMN	.03	.02	-.13	-.08	.08	.15	.12	.46*	.30†	.07	.04	.18	.28†	--					
15. Vowel-V MMN	-.02	.15	-.04	.07	-.12	.01	-.05	.29†	-.02	.33†	-.21	.09	.01	.04	--				
16. Tone-T P3a	-.07	.12	-.01	.02	.03	-.02	.00	.15	.32†	-.19	.28†	.43*	.24	.38*	.02	--			
17. Vowel-V P3a	-.03	-.26	-.23	-.28†	.17	-.12	.01	-.08	.13	.13	.24	.04	.28†	.11	-.02	.09	--		
18. Tone-V P3a	-.02	-.02	.09	-.01	.14	.13	.14	.20	.25	.00	.11	.22	.40*	.47	-.04	.45*	.28	--	
19. Vowel-V P3a	-.39*	.39*	.12	.39*	-.20	-.02	-.11	-.20	.24	-.14	.36*	-.08	-.05	-.18	.12	.02	-.09	-.02	--

* Significant for α = .05. † Significant for α = .1. No correction for multiple comparisons performed. Blue area highlights relations between ERPs and age or behaviour.

Table S2. Correlations across age, behavioural, and ERP fractional area latency measures in TD participants.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.
1. Age	--															
2. SCQ S+C	.42†	--														
3. SCQ RRB	.55*	.18	--													
4. SCQ TOT	.61*	.63*	.84*	--												
5. P1 standard /a/	.29	.10	.26	.28	--											
6. N2 standard /a/	-.31	-.24	-.32	-.45†	-.10	--										
7. P1 standard tone	.24	.16	.29	.27	.16	-.10	--									
8. N2 standard tone	-.31	-.07	-.19	-.18	-.06	.34	-.07	--								
9. Tone-T MMN	-.13	-.04	-.25	-.21	.21	.31	-.06	.17	--							
10. Vowel-T MMN	-.21	-.27	-.01	-.12	.02	.35	.30	.24	.64*	--						
11. Tone-V MMN	-.06	-.08	-.38	-.37	-.13	.64*	.08	.36	.55*	.53	--					
12. Vowel-V MMN	.08	-.28	-.06	-.14	.05	.25	.08	.31	.06	.36	.42†	--				
13. Tone-T P3a	-.17	-.23	-.43†	-.57*	-.14	.64*	.10	-.04	.54	.42†	.65*	-.05	--			
14. Vowel-V P3a	.27	.10	.37	.33	.02	.11	.27	.07	.30	.49†	.08	.19	.12	--		
15. Tone-V P3a	-.27	-.13	-.20	-.20	-.29	.33	.45†	.19	-.17	.19	.34	.38	.11	.09	--	
16. Vowel-V P3a	-.34	-.41†	.05	-.16	-.10	-.13	-.15	-.39	-.22	-.01	-.19	-.10	-.12	-.19	-.01	--

* Significant for α = .05. † Significant for α = .1. No correction for multiple comparisons performed. Blue area highlights relations between ERPs and age or behaviour.

Table S3. Correlations across age, behavioural, and ERP mean voltage measures in participants with ASD.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.
1. Age	--																		
2. ADOS SA CSS	-.41*	--																	
3. ADOS RRB CSS	.03	.09	--																
4. ADOS CSS	-.44*	.91*	.38*	--															
5. nvDQ	.23	-.22	-.14	-.19	--														
6. vDQ	.32†	-.20	-.03	-.16	.75*	--													
7. DQ	.30†	-.22	-.08	-.18	.92*	.95*	--												
8. P1 standard /a/	.07	-.31†	-.11	-.28	.10	.10	.11	--											
9. N2 standard /a/	.26	-.27	.01	-.26	-.06	.06	.00	.73*	--										
10. P1 standard tone	.18	-.33†	-.26	-.42*	.26	.18	.23	.51*	.56*	--									
11. N2 standard tone	.22	-.19	.09	-.19	-.08	.04	-.02	.21	.63*	.66*	--								
12. Tone-T MMN	.14	-.11	-.32†	-.14	.22	.25	.25	-.03	-.12	-.31†	-.36*	--							
13. Vowel-T MMN	.27	-.03	.22	.07	.20	.10	.16	.18	.12	-.02	.10	.10	--						
14. Tone-V MMN	.10	.00	.23	.09	-.08	.04	-.01	-.38*	-.22	-.34*	-.03	.07	.19	--					
15. Vowel-V MMN	-.02	-.29†	-.15	-.32†	.12	-.06	.02	-.39*	-.36*	-.07	-.14	.09	-.36*	-.01	--				
16. Tone-T P3a	-.31†	.19	-.27	.13	.16	.03	.09	.09	-.03	-.21	-.39*	.47*	.06	-.23	-.09	--			
17. Vowel-V P3a	-.04	.08	.09	.09	.11	-.04	.03	.42*	.18	.13	-.10	-.13	.59*	-.23	-.42*	.23	--		
18. Tone-V P3a	-.22	.17	.31†	.25	-.28	-.30†	-.31†	-.16	-.30†	-.43*	-.37*	-.20	.08	.57*	-.06	-.02	.15	--	
19. Vowel-V P3a	-.23	-.09	-.06	-.07	.06	-.01	.02	-.38*	-.50*	-.24	-.24	.10	-.17	.18	.70*	.14	-.32†	.19	--

* Significant for α = .05. † Significant for α = .1. No correction for multiple comparisons performed. Blue area highlights relations between ERPs and age or behaviour.

Table S4. Correlations across age, behavioural, and ERP mean voltage measures in TD participants.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	1.	11.	12.	13.	14.	15.	16.
1. Age	--															
2. SCQ S+C	.42†	--														
3. SCQ RRB	.55*	.18	--													
4. SCQ TOT	.61*	.63*	.84*	--												
5. P1 standard /a/	-.19	.24	-.34	-.22	--											
6. N2 standard /a/	-.10	-.25	-.32	-.46†	.46†	--										
7. P1 standard tone	.24	.27	.00	.12	.41	.24	--									
8. N2 standard tone	.17	-.16	.10	-.01	.04	.63*	.64*	--								
9. Tone-T MMN	-.10	-.09	.12	.09	-.38	-.24	-.51*	-.32	--							
1. Vowel-T MMN	.04	.24	-.03	.13	-.26	-.02	-.59*	-.26	.42*	--						
11. Tone-V MMN	.37	.14	.25	.28	-.60*	-.17	-.18	.06	.36	.29	--					
12. Vowel-V MMN	-.01	.37	-.23	-.05	.04	.16	.27	.24	.00	.22	.20	--				
13. Tone-T P3a	-.14	.06	.29	.24	-.26	-.56*	-.62*	-.66*	.67*	.37	.01	-.09	--			
14. Vowel-V P3a	.04	.15	.10	.16	-.26	-.25	-.37	-.43†	.30	.57*	.02	-.15	.33	--		
15. Tone-V P3a	-.10	.20	.12	.16	-.15	-.33	-.01	-.20	.19	-.04	.53*	.29	.29	-.04	--	
16. Vowel-V P3a	-.18	-.04	.06	.00	.08	.18	.13	.29	.15	.22	-.04	.63*	.24	-.18	.28	--

* Significant for α = .05. † Significant for α = .1. No correction for multiple comparisons performed. Blue area highlights relations between ERPs and age or behaviour.

S5. Analysis of group differences in P3a mean voltages, with p-values calculated while controlling for age and sex.

	P3a	
	F(2,48)	p
Background x stimulus x group	4.747	.0002**
Background x group	3.509	.043*
Stimulus x group	1.639	.380
Group	1.654	.054

Follow-up of Background x stimulus x group interaction while controlling for age:

1) Splitting by stimulus type.
   Tone stimuli: background x group interaction: $F(2, 48) = 5.334$, $p = .0004$, $\alpha = .025$.
   Vowel stimuli: : background x group interaction: $F(2,48) = .203$, $p = .869$, $\alpha = .025$.

2) Tone stimuli. Splitting by background.
   Tones in tones: effect of group: $F(2,48) = .347$, $p = .572$, $\alpha = .025$.
   Tones in vowels: effect of group: $F(2,48) = 6.172$, $p = .0004$, $\alpha = .025$.

Post-hoc tests of pairwise group differences in P3a mean voltages elicited by deviant tones in vowels (tone-V). Univariate ANCOVA with group as factor and age as covariate.

1) ASD/noDD vs TD: $F(1,27) = 9.489$ $p = .0003$, $\alpha = .016$.
2) ASD/noDD vs ASD/DD: $F(1,32) = 11.538$ $p = .0002$, $\alpha = .016$.
   With ADOS SA CSS score added as a covariate: $F(1,32) = 11.538$ $p = .0005$.
3) ASD/DD vs TD: $F(1,37) = .014$ $p = .941$, $\alpha = .016$.

S6. One sample t-test on data from child on anticonvulsant medication.

Because ASD/DD and ASD/noDD differed in P3 mean voltages in response to Tone-V, a one-sample t-test was run to assess the extent to which the child on anticonvulsant medication differed from the rest of the ASD/DD sample, to which she belonged.

The test shows that the P3a mean voltage exhibited by this child to Tone-V stimuli was significantly lower than the ASD/DD group mean ($t(20) = 5.741$, $p = .001$, $\alpha = .05$, Cohen's $d = 1.28$; child's mean P3a voltage: -3.46 µV, ASD/DD group grand-average: 6.93 µV, SD: 8.30). Because overall children with ASD/DD had higher P3a voltages than children ASD/noDD, her data increase within-group variance but reduce between-group variance. Her data were therefore included in the main analysis, as doing so does not enhance the significance of our results.