Atypical visual-auditory predictive coding in autism spectrum disorder: Electrophysiological evidence from stimulus omissions

Thijs van Laarhoven[1] [iD]**, Jeroen J Stekelenburg[1],
Mart LJM Eussen[2] and Jean Vroomen[1]**

Abstract

Autism spectrum disorder is a pervasive neurodevelopmental disorder that has been linked to a range of perceptual processing alterations, including hypo- and hyperresponsiveness to sensory stimulation. A recently proposed theory that attempts to account for these symptoms, states that autistic individuals have a decreased ability to anticipate upcoming sensory stimulation due to overly precise internal prediction models. Here, we tested this hypothesis by comparing the electrophysiological markers of prediction errors in auditory prediction by vision between a group of autistic individuals and a group of age-matched individuals with typical development. Between-group differences in prediction error signaling were assessed by comparing event-related potentials evoked by unexpected auditory omissions in a sequence of audiovisual recordings of a handclap in which the visual motion reliably predicted the onset and content of the sound. Unexpected auditory omissions induced an *increased* early negative omission response in the autism spectrum disorder group, indicating that violations of the prediction model produced larger prediction errors in the autism spectrum disorder group compared to the typical development group. The current results show that autistic individuals have alterations in visual-auditory predictive coding, and support the notion of impaired predictive coding as a core deficit underlying atypical sensory perception in autism spectrum disorder.

Lay abstract

Many autistic individuals experience difficulties in processing sensory information (e.g. increased sensitivity to sound). Here we show that these difficulties may be related to an inability to process unexpected sensory stimulation. In this study, 29 older adolescents and young adults with autism and 29 age-matched individuals with typical development participated in an electroencephalography study. The electroencephalography study measured the participants' brain activity during unexpected silences in a sequence of videos of a handclap. The results showed that the brain activity of autistic individuals during these silences was increased compared to individuals with typical development. This increased activity indicates that autistic individuals may have difficulties in processing unexpected incoming sensory information, and might explain why autistic individuals are often overwhelmed by sensory stimulation. Our findings contribute to a better understanding of the neural mechanisms underlying the different sensory perception experienced by autistic individuals.

Keywords

autism spectrum disorder, event-related potentials, predictive coding, visual-auditory

Introduction

Autism spectrum disorder (ASD) is a pervasive neurodevelopmental disorder characterized by impairments in social communication and social interaction, and restricted, repetitive patterns of behavior, interests, or activities (American Psychiatric Association, 2013; Robertson &

[1]Tilburg University, The Netherlands
[2]Yulius Mental Health Organization, The Netherlands

Corresponding author:
Thijs van Laarhoven, Department of Cognitive Neuropsychology, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.
Email: T.J.T.M.vanLaarhoven@TilburgUniversity.edu

Baron-Cohen, 2017). In addition, ASD has been linked to a range of perceptual processing alterations, including atypical processing of facial emotions (Eussen et al., 2015; Harms et al., 2010; Pellicano et al., 2007; Uljarevic & Hamilton, 2013) and hypo- and hyperresponsiveness to sensory stimulation (Baranek et al., 2013; Robertson & Baron-Cohen, 2017).

A recently proposed theory that attempts to account for these symptoms, states that autistic individuals have a decreased ability to anticipate upcoming sensory stimulation (Lawson et al., 2014; Pellicano & Burr, 2012; van de Cruys et al., 2014). A key element of the predictive coding framework is the assumption that incoming sensory information is continuously contrasted with internal predictions about the current state of our environment based on previous experiences (Friston, 2005). Any discrepancy between the sensory input and prior expectations results in the computation of an error signal. These prediction errors are crucial to adequately contextualize sensory information. They inform our perception about the current state of the world, and indicate that our current internal predictive model is not able to adequately predict upcoming sensory stimulation, and, thus, needs to be updated to resolve similar prediction errors in the future. Given that the world is not static (i.e. two perceptual experiences are never completely alike), prediction errors are always present to some extent. Although prediction errors are typically evoked by unexpected and "newsworthy" sensory stimulation that ought to increase our attention (e.g. a car ignoring a crosswalk), they may sometimes be spurious and uninformative (e.g. someone dropping a glass at a party). Thus, in order to adequately adjust the impact of prediction violations on updates of the predictive model, prediction errors need to be processed with a certain degree of flexibility: some prediction errors should be processed with "high priority," while others should be ignored and suppressed. Recently, it has been proposed that an inability to flexibly process prediction errors may be the core deficit underlying the socio-communicative impairments in ASD (van de Cruys et al., 2014). Others have posited that an imbalance in the importance ascribed to sensory input—relative to prior expectations—may cause autistic individuals to overweigh the significance of prediction errors (Lawson et al., 2014). It has also been argued that autistic individuals have a decreased ability to infer the probabilistic structure of their environment—resulting in imprecise or attenuated prior expectations or (in Bayesian terms) "hypo-priors" (Pellicano & Burr, 2012). Although conceptually distinct (for an overview, see Brock, 2012; Friston et al., 2013; Lawson et al., 2014; van de Cruys et al., 2013), all these theoretical accounts may result in an overreliance on sensory input. Given that perceptual cues are often noisy and ambiguous, a predictive model that is biased toward sensory input—rather than modulated by prior experience—may generate predictions that are "overfitted" to specific contexts, but do not generalize well to new experiences in which the sensory environment is often volatile. Following this reasoning, new experiences may generate large prediction errors in autistic individuals, since their overfitted prior expectations are likely to be violated by novel sensory input. Failing to contextualize and generalize sensory information in an optimal fashion—based on both current sensory input and prior expectations—may lead to atypical sensitivity to sensory stimulation (including hypo- and hyperresponsiveness), which could ultimately affect sensory processing, perception, and social interaction. Understanding the neural basis of the potential impairments in predictive coding in ASD may thus very well be a fundamental part of the explanation of why autistic individuals often struggle with social communication and interaction with their environment.

Recent evidence suggests that predictive coding might indeed be impaired in autistic individuals (van Laarhoven et al., 2019). In this study, the neural response to self- versus externally initiated tones was examined in a group of autistic individuals and a group of age-matched individuals with typical development (TD). The amplitude of the auditory N1 component of the event-related potential (ERP) is typically attenuated for self-initiated sounds, compared to sounds with identical acoustic and temporal features that are triggered externally (Baess et al., 2011; Baess et al., 2008; Bendixen et al., 2012; Martikainen et al., 2005). This attenuation effect has been ascribed to internal prediction models predicting the sensory consequences of one's own motor actions. The results of this study showed that (unlike in the TD group), self-initiation of the tones did not attenuate the auditory N1 in the ASD group, indicating that they may be unable to fully anticipate the (auditory) sensory consequences of their own motor actions. This raises the question of whether the ability to predict the actions of other individuals is altered as well in ASD. Given that the behavior of other individuals is arguably more difficult to predict than self-initiated actions, and the fact that autistic individuals have great difficulty with understanding the thoughts and emotions of their own and those of others (Robertson & Baron-Cohen, 2017), this seems plausible.

A growing area of interest and relevance in the study of predictive coding focuses on the electrophysiological responses to unexpected stimulus omissions of predictable sounds (SanMiguel, Widmann, et al., 2013; Stekelenburg & Vroomen, 2015). Auditory stimulation can be made predictable either by a motor act or anticipatory visual information (such as in a handclap, in which the movement of the hands precedes the sound) that reliably predicts the timing and content of the sound. Unexpected omissions of predictable sounds typically evoke an early negative omission response (oN1) that peaks between 45 and 100 ms in the electroencephalography (EEG) during the period of silence where the sound was expected (SanMiguel, Saupe,

& Schröger, 2013; SanMiguel, Widmann, et al., 2013; Stekelenburg & Vroomen, 2015; van Laarhoven et al., 2017). The amplitude of the auditory oN1 is hypothesized to be modulated by the prediction and prediction error (Arnal & Giraud, 2012; Friston, 2005). For sounds that are highly predictable, precise auditory predictions can be formed about the content and timing of the sound. If incoming auditory stimulation does not match (but violates) this prior expectation, such as during unexpected auditory omissions, the prediction error is large, and thus the oN1 is enlarged. If no clear predictions can be formed about an upcoming sound, the prediction is less likely to be violated, and so the oN1 is attenuated or absent during auditory omissions. Several studies have indeed shown that the oN1 is only elicited by unexpected omissions of sounds of which both the timing and content is predictable either by a motor act or anticipatory visual information, and not by omissions of unpredictable sounds or auditory omissions per se (Bendixen et al., 2012; SanMiguel, Saupe, & Schröger, 2013; SanMiguel, Widmann, et al., 2013; van Laarhoven et al., 2017). Hence, the oN1 can be considered as an early marker of auditory prediction error.

In the current study, we used a stimulus omission approach to examine the electrophysiological markers of prediction errors in auditory prediction by vision in autistic individuals to assess their ability to anticipate the sensory consequences of others' actions. An experimental paradigm was applied that was similar to those used in previous studies showing robust and consistent visual-auditory oN1 effects in TD individuals (Stekelenburg & Vroomen, 2015; van Laarhoven et al., 2017). EEG was recorded in a group of older adolescents and young adults with a clinical diagnosis of ASD, and in an age-matched group of individuals with TD. Between-group differences in visual-auditory predictive coding were assessed by comparing ERPs evoked by unexpected auditory omissions in a sequence of audiovisual recordings of a handclap, in which the visual motion reliably predicted the timing and content of the sound (Stekelenburg & Vroomen, 2007, 2015). Atypical enlargement of the oN1 response, a neural marker of prediction error, was considered as evidence for altered visual-auditory predictive coding and a potential indication of overreliance on sensory input.

Methods

Participants

Twenty-nine autistic individuals (eight female), and 29 age-matched individuals with TD (six female) participated in this study (ASD: $M = 18.64$ years, $SD = 2.11$; TD: $M = 18.93$ years, $SD = 1.22$). Inclusion criteria for participants in both groups were normal or corrected-to-normal vision and hearing, Full Scale Intelligence Quotient (FSIQ) ≥ 80, and no active use of sedatives 2 days prior to the experiment. Additional inclusion criteria for the ASD group were a clinical classification of ASD according to the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; *DSM*-IV-TR; American Psychiatric Association, 2000), and absence of severe comorbid neurological disorders (e.g. epilepsy). Additional inclusion criteria for the TD group were absence of any neurological or neuropsychiatric disorder (e.g. ASD, attention deficit hyperactivity disorder (ADHD), epilepsy).

Participants with ASD were recruited at a mental health institution for ASD (de Steiger, Yulius Mental Health, Dordrecht, the Netherlands). At the time of the experiment, all participants in the ASD group were receiving clinical treatment at this mental health institution due to severe mental problems and impaired functioning in activities of daily living linked to ASD. Participants with TD were recruited at Tilburg University and a high school located in the city of Tilburg.

For all participants in the ASD group the clinical *DSM*-IV TR classification of ASD was confirmed by two independent clinicians. Additional diagnostic information was retrieved when available, including Autism Diagnostic Observation Schedule (ADOS) scores (Lord et al., 2012), and Social Responsiveness Scale (SRS) scores (Constantino & Gruber, 2013). FSIQ was measured with the Dutch versions of the Wechsler Adult Intelligence Scale (WAIS-IV-NL) in participants ≥ 18 years, and the Wechsler Intelligence Scale for Children (WISC-III-NL) in participants <18 years. Demographic details of the ASD group and TD group are shown in Table 1. Specific data on socioeconomic status and educational attainment levels were not recorded. There were no significant differences in age, $t(56) = -0.64$, $p = 0.53$, and gender, $t(56) = 0.61$, $p = 0.55$), but the average FSIQ score was higher for the TD group (mean FSIQ 112.07, $SD = 11.68$) compared to the ASD group (mean FSIQ 103.03, $SD = 16.76$), $t(56) = 2.38$, $p = 0.02$.

All procedures were undertaken with the understanding and written consent of each participant and—for participants below the age of 18—a parent or another legally authorized representative. Participants with ASD and TD participants who were recruited at the high school were reimbursed with €25 for their participation. TD participants recruited at Tilburg University received course credits as part of a curricular requirement. All experimental procedures were approved by the local medical ethical review board (METC Brabant, protocol ID: NL52250.028.15) and performed in accordance with the ethical standards of the Declaration of Helsinki.

Stimuli and procedure

Participants were individually tested in a sound attenuated and dimly lit room, and were seated in front of a 19-in. CRT monitor (Iiyama Vision Master Pro 454) positioned at eye-level at a viewing distance of approximately 70 cm.

Table 1. Participant demographics for the autism spectrum disorder (ASD) and typically developing (TD) group.

	ASD	TD
Gender[n.s.]	21 male, 8 female	23 male, 6 female
Age[n.s.]	$M = 18.64$, $SD = 2.11$, range: 15–24	$M = 18.93$, $SD = 1.22$, range: 15–20
Full Scale IQ*	$M = 103.03$, $SD = 16.76$, range: 80–134	$M = 112.07$, $SD = 11.68$, range: 88–136
ADOS $n = 17$	$M = 10.06$, $SD = 5.19$, range: 4–22	–
SRS $n = 22$	$M = 72.91$, $SD = 9.68$, range: 55–92	–

M: mean; *SD*: standard deviation; IQ: intelligence quotient; ADOS: Autism Diagnostic Observation Schedule; SRS: Social Responsiveness Scale.
[n.s.]nonsignificant.
*$p < 0.05$.
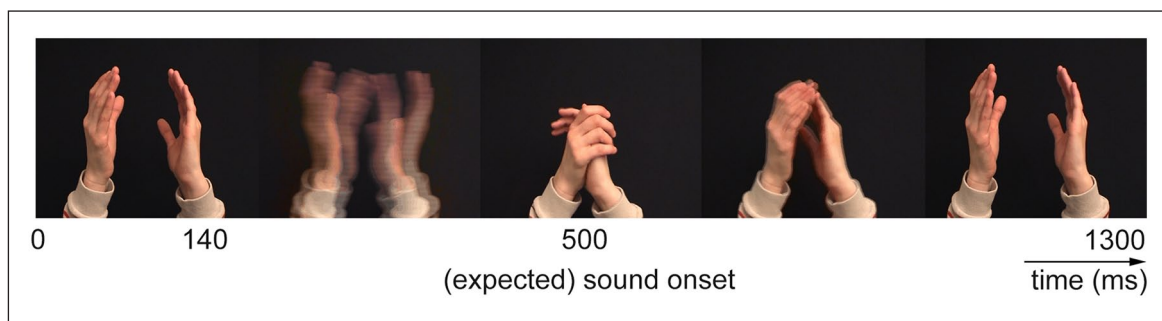
Figure 1. Time-course of the video used in the visual-auditory (VA) and visual (V) condition.

Stimulus materials were adapted from previous work on visual-auditory predictive coding in TD individuals (Stekelenburg & Vroomen, 2015; van Laarhoven et al., 2017). Visual stimuli consisted of a video-recording portraying the visual motion of a single handclap (Figure 1). The video started with the hands separated. Subsequently, the hands moved to each other, and after collision returned to their original starting position. The total duration of the video was 1300 ms. The video was presented at a frame rate of 25 frames/s, at a refresh rate of 100 Hz, and a resolution of 640 × 480 pixels (14° horizontal and 12° vertical visual angle). Auditory stimuli consisted of an audio recording (sampling rate 44.1 kHz) of the handclap portrayed in the video, and were presented at approximately 61 dB (A) sound pressure level over JAMO S100 stereo speakers, located directly on the left and right sides of the monitor. Stimulus presentation was controlled using E-Prime 1.2 (Psychology Software Tools Inc., Sharpsburg, PA, USA).

Three conditions were included in the experiment: visual-auditory (VA), visual (V), and auditory (A). In the VA condition, the video of the handclap was presented synchronously with the audio recording of the handclap. The handclap sound occurred 360 ms after the start of the hand movement. The auditory interstimulus interval was 1300 ms. Standard VA trials were interspersed with unpredictable omissions of the handclap sound in 12% of the trials (cf. SanMiguel, Saupe, & Schröger, 2013; Stekelenburg & Vroomen, 2015). These omission trials were randomly intermixed with standard VA trials with the restrictions that the first five trials of each block, and the two trials immediately following an omission trial were always standard VA trials. The VA condition was presented in seven blocks of 200 trials, resulting in a total of 1400 stimulus presentations in the VA condition (1232 standard VA trials and 168 auditory stimulus omissions). In the V and A condition, only the video-recording or the sound of the handclap was presented, respectively. The V and A conditions were presented in two blocks of 100 trials, resulting in a total of 200 stimulus presentations in the V and A condition. Block order was quasi-randomized across participants such that V and A blocks were never presented successively.

The V condition was included to correct for visual activity in the auditory omission trials of the VA condition (see "EEG recording"). The auditory oN1 is assumed to be correlated to the amplitude of the N1 that the expected sound would normally elicit (SanMiguel, Widmann, et al., 2013). The A condition was therefore included to test whether potential between-group differences in omission responses could be attributed to differences in sensory processing of the handclap sound itself.

To ensure that participants watched the visual stimuli and remained vigilant, 8% of all VA, V, and A trials consisted of catch trials. Participants were required to respond with a button press after onset of a catch stimulus (i.e. a small white square superimposed on the handclap video, presented at the center of the screen, measuring 1° horizontal and 1° vertical visual angle). To prevent possible interference of (delayed) motor responses, these catch trials never preceded an omission trial. Average percentage of detected catch trials across conditions was high ($M = 98.30$, $SD = 2.81$) and did not differ between conditions or groups, and there was no condition × group interaction effect (all p values > 0.08), indicating that participants in both groups attentively participated in all conditions. Total duration of the experiment was approximately 45 min.

EEG acquisition and processing

The EEG was sampled at 512 Hz from 64 locations using active Ag–AgCl electrodes (BioSemi, Amsterdam, the Netherlands) mounted in an elastic cap and two mastoid electrodes. Electrodes were placed in accordance with the extended International 10–20 system. Two additional electrodes served as reference (Common Mode Sense active electrode) and ground (Driven Right Leg passive electrode). Horizontal electrooculogram (EOG) was recorded using two electrodes placed at the outer canthi of the left and right eyes. Vertical EOG was recorded from two electrodes placed above and below the right eye. BrainVision Analyzer 2.0 (Brain Products, Gilching, Germany) software was used for ERP analyses. EEG was referenced offline to an average of left and right mastoids and band-pass filtered (0.01–30 Hz, 24 dB/octave). The (residual) 50-Hz interference was removed by a 50-Hz notch filter. Raw data were segmented into epochs of 1000 ms, including a 200-ms pre-stimulus baseline period. Epochs were time-locked to the expected sound onset of auditory omission trials in the VA condition, and to the corresponding timestamp of trials in the V condition and to sound onset in the A condition. After EOG correction (Gratton et al., 1983), epochs with an amplitude change exceeding ±150 μV at any EEG channel were rejected, averaged, and baseline corrected for each condition separately. All participants were included in the final analysis. On average, 13.45 ($SD = 17.02$) of the presented 168 auditory omission trials were rejected, corresponding to 7.96% ($SD = 10.13$). Percentages of rejected trials were similar for the standard trials in the VA condition ($M = 8.22\%$, $SD = 10.01$), visual trials in the V condition ($M = 7.37\%$, $SD = 8.98$), and auditory trials in the A condition ($M = 11.00\%$, $SD = 15.65$). Across all conditions, 8.64% ($SD = 9.85$) of the trials were rejected. There were no significant differences in percentages of rejected trials between groups or conditions, and there was no condition × group interaction effect (all p values > 0.10). The ERP of the V condition was subtracted from the

auditory omission ERPs in the VA condition to nullify the contribution of visual activity to the omission ERPs. Consequently, the VA–V difference waves reflect prediction-related activity—induced by unexpected auditory omissions —devoid of visual activity (Stekelenburg & Vroomen, 2015; van Laarhoven et al., 2017).

Results

The group-averaged auditory omission ERPs (Figure 2) showed two distinct negative deflections in both groups: oN1 (45–100 ms), oN2 (100–200 ms). In accordance with previous research on auditory omission responses (SanMiguel, Widmann, et al., 2013), maximal amplitude of the oN1 and oN2 was measured at electrode FT7. The two negative omission responses were followed by late positive potentials oP3 (300–550 ms), showing maximal amplitudes measured at electrodes Cz.

Visual inspection of the ERPs showed that the oN1 for the ASD group was more pronounced compared to the oN1 for the TD group, while the oN2 and oP3 deflections appeared to be similar for both groups. The oN1 and oN2 deflections showed a bilateral scalp distribution in both groups, while the oP3 components had a central scalp distribution (Figure 3). Based on these scalp distributions, a left-temporal (F7, F5, F3, FT7, FC5, FC3, T7, C5, C3) and right-temporal (F4, F6, F8, FC4, FC6, FT8, C4, C6, T8) region of interest (ROI) were selected for the oN1 and oN2 time windows. A central–parietal (C1, Cz, C2, CP1, CPz, CP2) ROI was selected for the oP3 time window. The presence of statistically significant omission responses was tested by conducting separate repeated measures multivariate analysis of variance (MANOVAs) on the mean activity for each time window, with the within-subjects variables Electrode and ROI for the oN1 and oN2 time windows, and Electrode for the oP3 time window, and between-subjects factor Group (ASD, TD) for all time windows.

oN1 time window (45–100 ms)

The overall mean activity in the oN1 time window differed from pre-stimulus baseline levels, $F(1, 57) = 5.73$, $p = 0.02$, $\eta_p^2 = 0.09$. There was a main effect of Group, $F(1, 56) = 4.32$, $p = 0.04$, $\eta_p^2 = 0.07$, indicating that the mean activity in the oN1 time window (averaged across ROIs and electrodes) was 0.52 μV more negative in the ASD group compared to the TD group (see Figure 4(a) for group medians and interquartile ranges). There were no main effects of ROI, $F(1, 56) = 0.40$, $p = 0.53$, $\eta_p^2 = 0.01$, and Electrode $F(8, 49) = 1.95$, $p = 0.07$, $\eta_p^2 = 0.24$, and no significant interaction effects between the factors Electrode, ROI, and Group (all p values > 0.31).

To examine if the between-group difference in oN1 mean activity could be attributed to differences in sensory
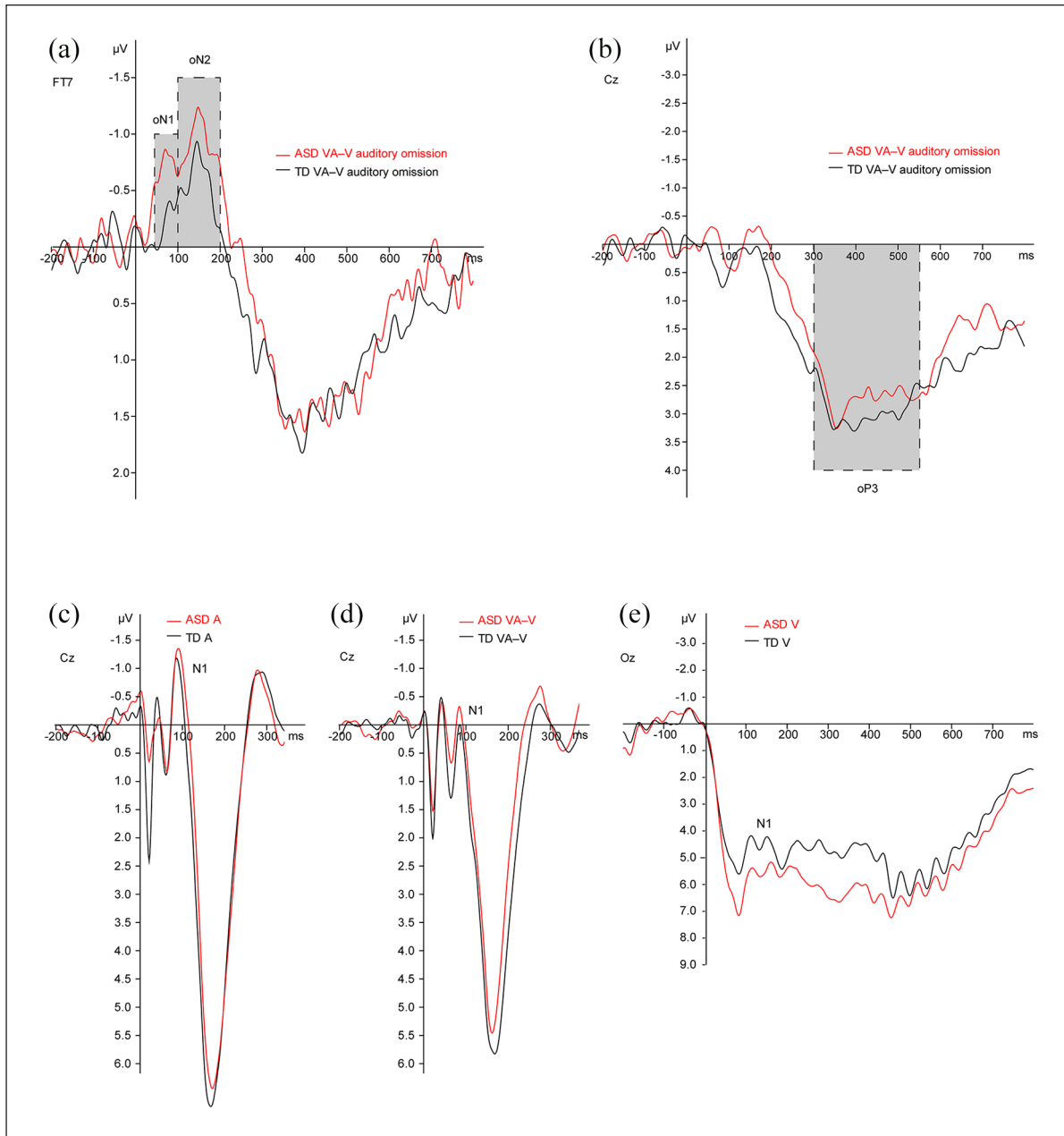
Figure 2. Direct comparison of the group-averaged ERPs. Auditory omission ERPs and visual-auditory (VA) ERPs were corrected for visual activity via subtraction of the visual (V) waveform. (a) The first negative component of the auditory omission ERPs peaked in a time window of 45–100 ms (oN1). A second negative component reached its maximum in 100–200 ms (oN2). Maximal amplitude of the oN1 and oN2 was measured at electrode FT7. (b) The two negative omission responses were followed by late positive potentials showing maximal amplitudes measured at electrodes Cz in a time window of 300–550 ms (oP3). (c–e) Group-averaged ERPs for auditory (A), standard visual-auditory (VA–V), and visual (V) stimulation showing maximal amplitudes measured at electrodes Cz (A, standard VA–V), and Oz (V).

processing of the sound and video of the handclap stimulus itself, three separate repeated measures MANOVAs were conducted on the peak amplitude of the N1 evoked by auditory trials in the A condition in a time window of 50–150 ms, the peak amplitude of the N1 evoked by standard trials in the VA–V condition in a time window of 50–150 ms, and the mean activity evoked by visual trials in the V condition in a time window of 75–175 ms. All analyses were conducted on the electrodes showing maximal activity (A: Cz, CPz; standard VA–V: Cz, CPz, V: O1, Oz, O2). The MANOVA on the peak amplitude of the auditory N1 in the A condition revealed no main effect of Group, $F(1, 56) = 0.19$, $p = 0.66$, $\eta_p^2 = 0.003$, and Electrode, $F(1, 56) = 0.05$, $p = 0.83$, $\eta_p^2 = 0.001$, and no
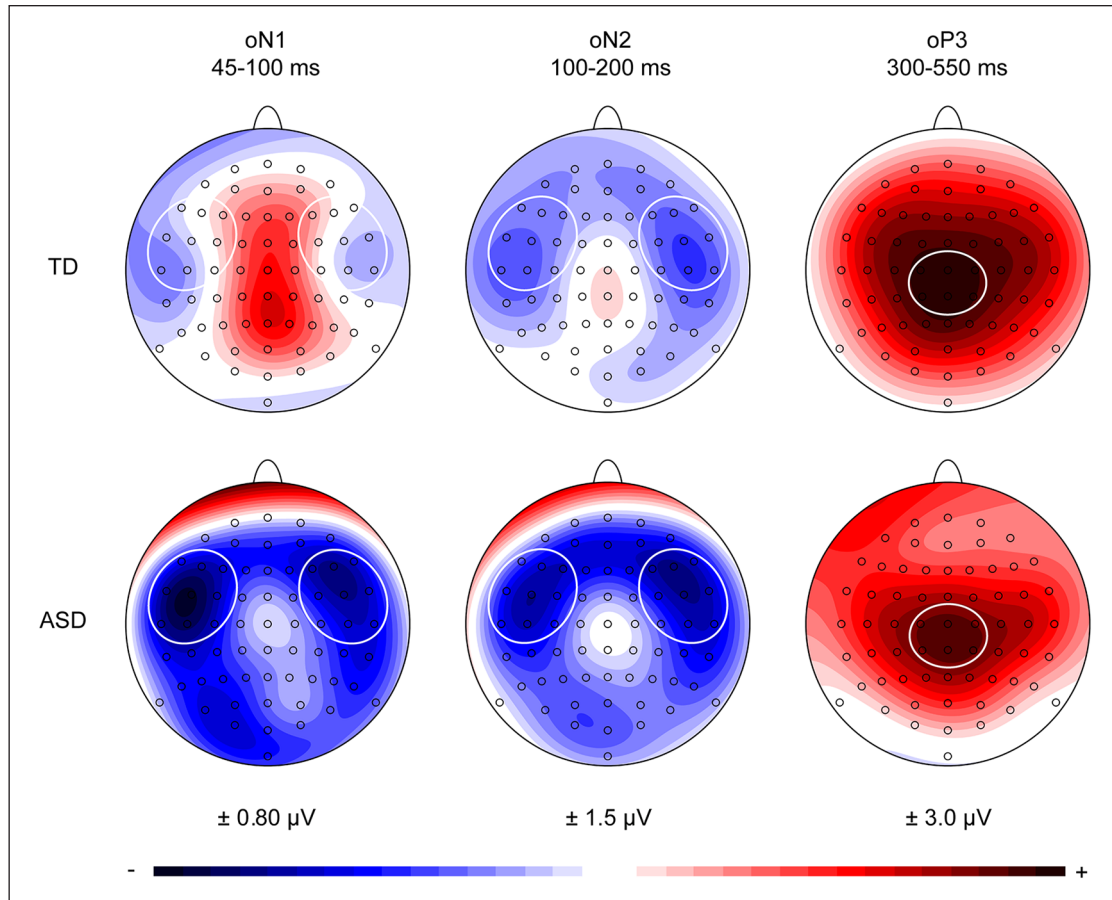
Figure 3. Scalp potential maps of the group-averaged visual-corrected auditory omission responses in the denoted oN1 (45–100 ms), oN2 (100–200 ms), and oP3 (300–550 ms) time windows. Based on these scalp distributions, a left-temporal (F7, F5, F3, FT7, FC5, FC3, T7, C5, C3) and right-temporal (F4, F6, F8, FC4, FC6, FT8, C4, C6, T8) region of interest were selected for the oN1 and oN2 time windows. A central-parietal (C1, Cz, C2, CP1, CPz, CP2) region of interest was selected for the oP3 time window.

interaction effect between the factors Group and Electrode, $F(1, 56) = 2.17, p = 0.15, \eta_p^2 = 0.04$. Similarly, the MANOVA on the peak amplitude of the auditory N1 evoked by standard trials in the VA–V condition revealed no main effect of Group, $F(1, 56) = 0.04, p = 0.85, \eta_p^2 = 0.001$, and Electrode, $F(1, 56) = 1.33, p = 0.25, \eta_p^2 = 0.02$, and no interaction effect between the factors Group and Electrode, $F(1, 56) = 0.01, p = 0.91, \eta_p^2 < 0.001$, indicating that the N1 evoked by the handclap sound was similar for both groups.

The MANOVA on the mean activity of the visual N1 in the V condition revealed a main effect of Electrode, $F(2, 55) = 20.83, p < 0.001, \eta_p^2 = 0.43$. Post hoc paired samples t-tests (Bonferroni corrected) showed that the overall mean activity in the visual N1 time window significantly differed between all three electrodes (all p values < 0.05), such that activity was most negative at Oz, and least negative at O2. More importantly, there was no main effect of Group, $F(1, 56) = 1.73, p = 0.19, \eta_p^2 = 0.03$, and no significant Group × Electrode interaction, $F(2, 55) = 2.71, p = 0.08, \eta_p^2 = 0.09$. Hence, the between-group difference in oN1 response could not be attributed to differences in

auditory or visual stimulus processing per se, but instead, more likely reflects a difference in prediction error signaling.

To ensure that the difference in FSIQ between the ASD and TD group was not a confounding factor for the difference in oN1 mean activity, a post hoc partial correlation analysis controlling for group membership was conducted correlating individual oN1 mean activity averaged across the left- and right-temporal ROI to FSIQ. This analysis revealed that the oN1 mean activity was not affected by FSIQ ($r = -0.03, p = 0.85$), thereby ruling out FSIQ as a confounding factor for the difference in oN1 mean activity between the ASD and TD group.

oN2 time window (100–200 ms)

The overall mean activity in the oN2 time window differed from pre-stimulus baseline levels, $F(1, 57) = 21.27$, $p < 0.001, \eta_p^2 = 0.28$. There was no main effect of Group, $F(1, 56) = 2.07, p = 0.16, \eta_p^2 = 0.04$, ROI, $F(1, 56) = 0.72$, $p = 0.40, \eta_p^2 = 0.01$, and Electrode, $F(8, 49) = 1.55$, $p = 0.20, \eta_p^2 = 0.03$, and no significant interaction effects between the factors Group, ROI, and Electrode (all p values
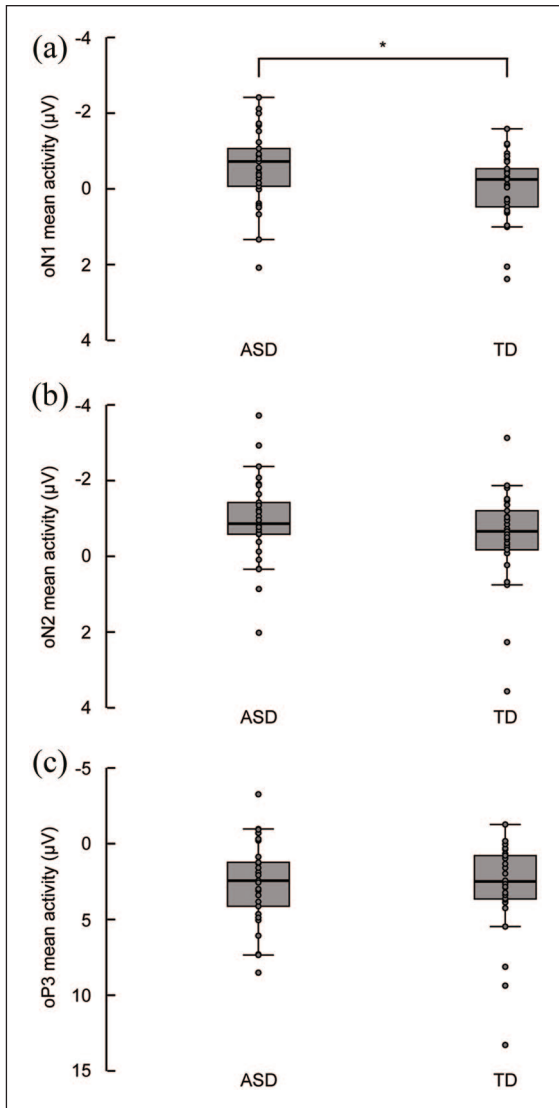
Figure 4. Boxplots displaying the group medians and interquartile ranges overlaid with individual data points of the visual-corrected auditory omission responses for the ASD and TD group in the denoted oN1 (45–100 ms), oN2 (100–200 ms), and oP3 (300–550 ms) time windows averaged across regions of interest and electrodes. (a) The mean activity in the oN1 time window was significantly more negative in the ASD group compared to the TD group. (b and c) The mean activity in the oN2 and oP3 time windows was similar in both groups.

> 0.19), indicating that the mean activity in the oN2 time window was similar in both groups (see Figure 4(b) for group medians and interquartile ranges).

oP3 time window (300–550 ms)

The overall mean activity in the oP3 time window differed from pre-stimulus baseline levels, $F(1, 57) = 49.61$, $p < 0.001$, $\eta_p^2 = 0.47$. There was no main effect of Group, $F(1, 56) = 6.88$, $p = 0.72$, $\eta_p^2 = 0.002$, and Electrode, $F(5, 52) = 2.29$, $p = 0.06$, $\eta_p^2 = 0.18$, and no significant

Group × Electrode interaction, $F(5, 52) = 0.15$, $p = 0.98$, $\eta_p^2 = 0.01$, indicating that the mean activity in the oP3 time window was similar in both groups (see Figure 4(c) for group medians and interquartile ranges).

Discussion

The current study tested the hypothesis that predictive coding is impaired in ASD due to overly precise internal prediction models by comparing the neural correlates of visual-auditory prediction errors between autistic individuals and individuals with TD using a stimulus omission paradigm. The data revealed clear group differences in the early electrophysiological indicators of visual-auditory predictive coding. The oN1, a neural marker of prediction error, was significantly more pronounced in the ASD group, indicating that violations of the visual-auditory predictive model—induced by unexpected auditory omissions—produced larger prediction errors in the ASD group compared to the TD group. Importantly, the increased prediction error signaling in the ASD group could not be explained by between-group differences in the processing of the physical characteristics of the applied stimuli. The current results could thus be indicative of altered visual-auditory predictive coding in ASD.

Previous studies have shown that increasing attention toward an auditory stimulus may increase the amplitude of the N1 response (Lange et al., 2003), whereas drawing attention away may result in N1 attenuation (Horváth & Winkler, 2010). Whether attention can affect the oN1 remains to be investigated. But if so, it might be argued that increased attention to the handclap sounds may have resulted in an amplitude increase of the oN1 in the ASD group. An argument against this view is that the N1 for auditory and audiovisual stimulation during standard trials was similar in the ASD and TD group, indicating that sustained attentional differences between groups are an unlikely account for the increased oN1 response in the ASD group.

In both the TD and ASD groups, the oN1 was followed by an oN2 and oP3 response. The current results mirror those of previous studies applying motor- and visual-auditory omission paradigms (SanMiguel, Saupe, & Schröger, 2013; SanMiguel, Widmann, et al., 2013; Stekelenburg & Vroomen, 2015; van Laarhoven et al., 2017), in which the oN1 was also followed by an oN2 and oP3 response. The oN2 is assumed to reflect higher order error evaluation associated with stimulus deviance or the presence of conflict in the context of action monitoring (SanMiguel, Saupe, & Schröger, 2013; SanMiguel, Widmann, et al., 2013; Stekelenburg & Vroomen, 2015; van Laarhoven et al., 2017); in this case, a conflict between the visually anticipated sound and the omitted sound. The oP3 likely reflects attention-orienting triggered by the unexpected omission of the sound, and the subsequent updating of the internal forward model to minimize future error (Baldi & Itti, 2010; Polich, 2007). Previous research

has shown that the oN1 response and oN2–oP3 complex is only elicited by unexpected omissions of sounds of which both the timing and content is predictable (SanMiguel, Saupe, & Schröger, 2013; SanMiguel, Widmann, et al., 2013; Stekelenburg & Vroomen, 2015; van Laarhoven et al., 2017). The enlarged oN1 response and typical oN2 and oP3 suggest that individuals in the ASD group were able to use the visual motion to predict the upcoming sound during audiovisual stimulation in the standard trials. The current results thus seem to argue against the imprecise or attenuated priors account of ASD (Pellicano & Burr, 2012). When the visual-auditory prediction was not fulfilled, but disrupted by an auditory omission, the ASD group showed an increased error response—as indicated by the atypically large oN1. Given that the amplitude of the oN1 is assumed to be modulated by the precision of the prediction (Arnal et al., 2011; Friston, 2005), the current results suggest that sensory prediction might be overly precise in ASD, as previously hypothesized (van de Cruys et al., 2014). An overly precise predictive model may generate predictions that are overfitted to specific contexts. This overfitting significantly impairs the generalizability of prior expectations to new sensory experiences, which in turn leads to disproportionately large prediction errors in response to unexpected variations in sensory input. The continuous signaling of prediction errors and overfitting of prediction models likely requires an excessive amount of attentional resources—which might explain why autistic individuals are often overwhelmed by sensory stimulation.

In relatively rigid, unambiguous situations, autistic individuals can successfully learn and apply new contingencies (Dawson et al., 2008), and they often excel in detail-focused tasks in which overfitted predictions are advantageous (Robertson & Baron-Cohen, 2017). The experimental paradigm applied in the current study provided a relatively unambiguous context (especially when compared with complex and social interactions). One might therefore expect that, even though the auditory omissions were infrequent and unpredictable, an overly precise predictive model would incorporate the occasional occurrence of an auditory omission after a certain number of iterations to minimize prediction errors in the future. Still, the prediction error—reflected in the oN1—remained atypically large, which suggests that there was little to no habituation to the auditory omissions in the ASD group. A persistent bias toward sensory input impedes the influence of prior expectations on perception and may cause each unexpected sensory experience to be handled as an error. The current findings may thus be in line with the notion that autistic individuals show alterations in habituation to (unexpected) sensory stimulation because they systematically overweigh the significance of sensory input over prior expectations (Lawson et al., 2014). It should be noted, however, that the signal-to-noise ratio of the current data does not allow for an analysis of oN1 amplitude over time; so whether habituation to the auditory omissions was

indeed absent in the ASD group remains to be elucidated. Future studies should therefore address if the increased prediction error response in the ASD group can be attributed to overly precise sensory predictions or a lack of habituation to unexpected sensory stimulation. Nevertheless, the current results imply that even in a relatively stable context with little noise, autistic individuals may experience difficulties in anticipating upcoming auditory stimulation.

Recent evidence has shown that self-initiation of tones does not attenuate the auditory N1 in autistic individuals, indicating that autistic individuals may have alterations in anticipating the auditory consequences of their own motor actions (van Laarhoven et al., 2019). The current study extends these findings by demonstrating that the ability to anticipate the sensory consequences of others' actions may be altered in ASD as well. While different predictive mechanisms may underlie N1 attenuation (as a marker of fulfilled prediction) and elicitation of the oN1 (as a marker of prediction error), both the absence of N1 attenuation and increased prediction error signaling may indicate that autistic individuals experience difficulties in anticipating upcoming sensory events and seemingly process every new experience afresh rather than mediated by prior expectations. Interaction with the environment becomes especially challenging in social situations, which are inherently noisy and volatile—and thus require flexible and fine-tuned processing of prior expectations, sensory input, and prediction errors. A potential consequence of this failure to contextualize sensory information and suppress prediction errors is a constant state of vigilance or sensory alertness—symptoms associated with sensory overload and hyperresponsiveness to sensory stimulation. Indeed, there is evidence that autistic individuals systematically overestimate the volatility of their environment (Lawson et al., 2017). Over time, this may lead to frustration, (social) anxiety, repetitive behaviors (e.g. insistence on sameness and stimming as an adaptive coping strategy to control sensory stimulation and attempt to minimize prediction errors), and ultimately, avoidance or hyporesponsiveness to sensory stimulation.

Future studies should focus on when the currently observed alterations in prediction error signaling first emerge throughout development, as the neural response to prediction disruptions may serve as an early marker of autistic symptomatology and potential target for intervention. Ultimately, future work may reveal if and how these alterations in predictive coding can be remediated through clinical applications to improve sensory-perceptual and social functioning of autistic individuals.

Conclusion

The current results confirm our hypothesis that autistic individuals show alterations in visual-auditory predictive coding. Specifically, unexpected auditory omissions in a

sequence of audiovisual recordings in which the visual motion reliably predicted the timing and content of the sound elicited an increased prediction error response in our sample of autistic individuals. The current data suggest that autistic individuals may have impairments in the ability to anticipate the sensory consequences of others' actions, and support the notion of impaired predictive coding as a core deficit underlying atypical sensory perception in ASD.

ORCID iD

Thijs van Laarhoven https://orcid.org/0000-0002-4034-7178

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.).

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Publishing.

Arnal, L. H., & Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends in Cognitive Sciences*, *16*(7), 390–398. https://doi.org/10.1016/j.tics.2012.05.003

Arnal, L. H., Wyart, V., & Giraud, A. L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, *14*(6), 797–801. https://doi.org/10.1038/nn.2810

Baess, P., Horváth, J., Jacobsen, T., & Schröger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology*, *48*(9), 1276–1283. https://doi.org/10.1111/j.1469-8986.2011.01196.x

Baess, P., Jacobsen, T., & Schröger, E. (2008). Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. *International Journal of Psychophysiology*, *70*(2), 137–143. https://doi.org/10.1016/j.ijpsycho.2008.06.005

Baldi, P., & Itti, L. (2010). Of bits and wows: A Bayesian theory of surprise with applications to attention. *Neural Networks*, *23*(5), 649–666. https://doi.org/10.1016/j.neunet.2009.12.007

Baranek, G. T., Watson, L. R., Boyd, B. A., Poe, M. D., David, F. J., & McGuire, L. (2013). Hyporesponsiveness to social and nonsocial sensory stimuli in children with autism, children with developmental delays, and typically developing children. *Development and Psychopathology*, *25*(2), 307–320. https://doi.org/10.1017/S0954579412001071

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of Psychophysiology*, *83*(2), 120–131. https://doi.org/10.1016/j.ijpsycho.2011.08.003

Brock, J. (2012). Alternative Bayesian accounts of autistic perception: Comment on Pellicano and Burr. *Trends in Cognitive Sciences*, *16*(12), 573–574. https://doi.org/10.1016/j.tics.2012.10.005

Constantino, J. N., & Gruber, C. P. (2013). *Social Responsiveness Scale (SRS-2)*. Western Psychological Services.

Dawson, M., Mottron, L., & Gernsbacher, M. A. (2008). Learning in Autism. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (pp. 759–772). Academic Press. https://doi.org/10.1016/b978-012370509-9.00152-2

Eussen, M. L. J. M., Louwerse, A., Herba, C. M., Van Gool, A. R., Verheij, F., Verhulst, F. C., & Greaves-Lord, K. (2015). Childhood facial recognition predicts adolescent symptom severity in autism spectrum disorder. *Autism Research*, *8*(3), 261–271. https://doi.org/10.1002/aur.1443

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

Friston, K., Lawson, R., & Frith, C. D. (2013). On hyperpriors and hypopriors: Comment on Pellicano and Burr. *Trends in Cognitive Sciences*, *17*(1), 1. https://doi.org/10.1016/j.tics.2012.11.003

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484. https://doi.org/10.1016/0013-4694(83)90135-9

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in autism spectrum disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, *20*(3), 290–322. https://doi.org/10.1007/s11065-010-9138-6

Horváth, J., & Winkler, I. (2010). Distraction in a continuous-stimulation detection task. *Biological Psychology*, *83*(3), 229–238. https://doi.org/10.1016/j.biopsycho.2010.01.004

Lange, K., Rösler, F., & Röder, B. (2003). Early processing stages are modulated when auditory stimuli are presented at an attended moment in time: An event-related potential study. *Psychophysiology*, *40*(5), 806–817. https://doi.org/10.1111/1469-8986.00081

Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nature Neuroscience*, *20*(9), 1293–1299. https://doi.org/10.1038/nn.4615

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

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule-2 manual*. Western Psychological Services.

Martikainen, M. H., Kaneko, K. I., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, *15*(3), 299–302. https://doi.org/10.1093/cercor/bhh131

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

Pellicano, E., Jeffery, L., Burr, D., & Rhodes, G. (2007). Abnormal adaptive face-coding mechanisms in children with autism spectrum disorder. *Current Biology*, *17*(17), 1508–1512. https://doi.org/10.1016/j.cub.2007.07.065

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Robertson, C. E., & Baron-Cohen, S. (2017). Sensory perception in autism. *Nature Reviews Neuroscience*, *18*(11), 671–684. https://doi.org/10.1038/nrn.2017.112

SanMiguel, I., Saupe, K., & Schröger, E. (2013). I know what is missing here: Electrophysiological prediction error signals elicited by omissions of predicted "what" but not "when." *Frontiers in Human Neuroscience*, *7*, Article 407. https://doi.org/10.3389/fnhum.2013.00407

SanMiguel, I., Widmann, A., Bendixen, A., Trujillo-Barreto, N., & Schröger, E. (2013). Hearing silences: Human auditory processing relies on preactivation of sound-specific brain activity patterns. *Journal of Neuroscience*, *33*(20), 8633–8639. https://doi.org/10.1523/JNEUROSCI.5821-12.2013

Stekelenburg, J. J., & Vroomen, J. (2007). Neural correlates of multisensory integration of ecologically valid audiovisual events. *Journal of Cognitive Neuroscience*, *19*(12), 1964–1973. https://doi.org/10.1162/jocn.2007.19.12.1964

Stekelenburg, J. J., & Vroomen, J. (2015). Predictive coding of visual-auditory and motor-auditory events: An electrophysiological study. *Brain Research*, *1626*, 88–96. https://doi.org/10.1016/j.brainres.2015.01.036

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders*, *43*(7), 1517–1526. https://doi.org/10.1007/s10803-012-1695-5

van de Cruys, S., De-Wit, L., Evers, K., Boets, B., & Wagemans, J. (2013). Weak priors versus overfitting of predictions in autism: Reply to Pellicano and Burr (TICS, 2012). *I-Perception*, *4*(2), 95–97. https://doi.org/10.1068/i0580ic

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

van Laarhoven, T., Stekelenburg, J. J., Eussen, M. L. J. M., & Vroomen, J. (2019). Electrophysiological alterations in motor-auditory predictive coding in autism spectrum disorder. *Autism Research*, *12*(4), 589–599. https://doi.org/10.1002/aur.2087

van Laarhoven, T., Stekelenburg, J. J., & Vroomen, J. (2017). Temporal and identity prediction in visual-auditory events: Electrophysiological evidence from stimulus omissions. *Brain Research*, *1661*, 79–87. https://doi.org/10.1016/j.brainres.2017.02.014