Title: Reduced perceptual error monitoring is a biomarker of autism

Authors**:** Nathaniel J. Zuk[1,2*], Jaroslav Y. Slobodskoy-Plusnin[1], Yarden Weiss[1], VishnuPriya Sampath[1], Kira M. Düsterwald[3], Athena Akrami[4§], Merav Ahissar[1§]

Affiliations: 1) Edmond & Lily Safra Center for Brain Sciences, Hebrew University, Jerusalem, Israel; 2) NTU Psychology, Nottingham Trent University, Nottingham, United Kingdom; 3) Gatsby Computational Neuroscience Unit, University College London, London, United Kingdom; 4) Sainsbury Wellcome Centre, University College London, London, United Kingdom

§These authors jointly supervised this work

Corresponding author: Nathaniel J. Zuk, NTU Psychology, Nottingham Trent University, Nottingham, United Kingdom (nathaniel.zuk@ntu.ac.uk)

Abstract:

Can reduced flexibility in autistics' behavior be a reliable classifier of autism? We asked this question in the context of pitch discrimination, which is performed with similar accuracy by autistic and non-autistic participants, matched for cognitive skills. We found that three reliable, uncorrelated characteristics quantified reduced behavioral flexibility. Most non-autistic participants showed a clear feedback-related negativity (FRN) in their EEG when incorrect, larger EEG responses during feedback following difficult versus easy trials, and increased perceptual bias following correct and easy trials versus difficult ones. All three measures were weaker in autistic participants. To assess reliability, we replicated all three measures in an experiment approximately a year later. These measures were not correlated within participants, indicating that several mechanisms, associated with anterior cingulate cortex, are atypical in autism. Together, these three measures provide group classification with above 80% accuracy, showing promise for their utility as a neurocognitive biomarker of autism.

Significance statement: A characteristic of autism is cognitive "slowness" in updating expectations in new contexts. Our study shows that this cognitive difference is related to a reduced use of feedback, linked to atypical activity in a brain area involved in error detection and behavioral adjustment. Using EEG, we found that autistic individuals have weaker feedback-related signals and do not adjust their behavior as effectively as non-autistic individuals. Together, these neurocognitive signals could serve as a biomarker for autism. This is the first time such brain responses have been measured in a perceptual task requiring continuous adjustments. Understanding these brain mechanisms can inform strategies to enhance learning and adaptability in autism, providing new insights into the neural basis of behavioral flexibility.

Introduction

People with autism spectrum disorder (ASD) are known to be less flexible in updating their motor and perceptual performance to changes in the environment (Van de Cruys et al., 2014; Lieder et al., 2019). For example, when people perform a perceptual discrimination task they implicitly learn the distribution of the stimuli they experience and form predictions regarding the upcoming stimuli (Ashourian and Loewenstein, 2011). These predictions are updated on each trial (Akrami et al., 2018; Lieder et al., 2019), but in ASD the rate of this update is reduced (Lieder et al., 2019; Sapey-Triomphe et al., 2023; Noel et al., 2024). Other studies have demonstrated slower updating in tasks requiring continuous monitoring of errors (South et al., 2012; D'Cruz et al., 2016; Vishne et al., 2021; Kasten et al., 2023), which could also relate to the slower rate at which they update their internal representation of the external world.

A separate line of evidence using EEG suggests atypical impact of both external and internal feedback. Reduced sensitivity to internal feedback is manifested when participants detect that they have made an erred response in fast motor tasks (such as the flanker task; Gehring et al., 1993, 1995, 2018), which is smaller in autism (Vlamings et al., 2008; Sokhadze et al., 2010; South et al., 2010; Hüpen et al., 2016) (but see Henderson et al., 2006; Groen et al., 2008). Similarly, reduced sensitivity to external feedback, called the feedback-related negativity (FRN) is manifested as an error in the expected outcome (Miltner et al., 1997). The FRN was often studied in the context of success and failure in guessing random rewards (Gehring and Willoughby, 2002; Holroyd and Coles, 2002; Bellebaum and Daum, 2008; Larson et al., 2011) rather than in evaluating the outcome of reward-maximization strategies in perceptual decision making paradigms. In autism, such a study design has yielded mixed results (Groen et al., 2008; Larson et al., 2011; Bellebaum et al., 2014) (for review see Hüpen et al., 2016).

The anterior cingulate cortex (ACC) is important for evaluating both predictions and actions, integrating internal and external information (Falkenstein et al., 1991; Gehring et al., 1993; Dehaene et al., 1994; Carter et al., 1998; Botvinick et al., 1999, 2004; Beckmann et al., 2009; Stevens et al., 2011), including updating predictions (Behrens et al., 2007), although its specific role is not well understood. Indeed, ACC was found to be atypical in ASD (Thakkar et al., 2008; Sapey-Triomphe et al., 2023), in both connectivity (Simms et

al., 2009; Cauda et al., 2011; Postema et al., 2019) and in its pattern of activation (Chan et al., 2011) particularly when processing perceptual errors (Sapey-Triomphe et al., 2021, 2023). This is consistent with recent findings on under-weighting of sensory observation and prediction error in frontal cortices across several mouse models of autism (Noel et al., 2024).

Together, the above observations suggest that online monitoring of both motor and sensory performance is reduced in ASD. We now asked whether all these mechanisms manifest in a single task and whether their measures are reliably reduced in autism. We used a single, simple perceptual task and aimed to measure all three aspects of monitoring: perceptual updates, internal feedback, and external feedback. We found that all forms of performance monitoring are reliably reduced in ASD. Moreover, all these measures, while uncorrelated within participant groups, were replicated in a follow-up study 1-2 years later and were sufficient when combined to classify non-autistic and ASD participants with over 80% accuracy. These results show that reduced performance monitoring, which can be reliably measured behaviorally and neurophysiologically, provides a biomarker for autism.

Methods

All experimental methods were approved by the European Research Council and the Ethics Committee of the Hebrew University.

Study 1

Participants

35 non-autistic and 34 ASD individuals participated in this experiment. One ASD participant was rejected from further analyses because of poor performance which suggested that they did not understand the task, leaving 33 ASD included in our analyses. All autistic individuals had a prior formal diagnosis (by a psychologist, neurologist, or psychiatrist), and participation in our study did not require changes in medication. Except for one participant in each group, all participants had at most four years of musical training. The 35 non-autistics and 33 ASD participants that were included in subsequent analyses were also matched for age, gender, and reasoning scores, measured by Block Design, a standardized test of visuospatial reasoning (Wechsler, 2008) (see Table 1). The Block

Design scores (population average is scaled to 10 and standard deviation to 1.5) indicate high cognitive skills in both groups. By contrast, and as expected, there was a significant difference in the scores of the self-reported autism questionnaire, AQ50 (Autism Quotient questionnaire; (Baron-Cohen et al., 2001)), with an effect size of Cohen's d = 1.10.

Table 1: Demographic information for the participants included in the experiment. For age, Block Design, and AQ50 the mean is specified for each group, with the standard deviation in parentheses. ** $p < 0.01$ based on a Student's t-test.

	Non-autistics	ASD	t-statistic	p-value
Number of participants	35	33		
Number of females	5	6		
Age (years)	25.9 (3.7)	25.4 (5.3)	0.493	0.624
Block design (scaled score)	12.5 (2.2)	12.9 (3.4)	-0.384	0.702
AQ50**	50.6 (15.8)	74.2 (24.6)	-3.269	0.002

Experiment design

Experiments were coded using PsychToolbox in Matlab (Mathworks). In each trial, participants were sequentially presented with two tones 50 ms in duration with 5 ms onset and offset sine ramps. Tone frequencies ranged from 800 to 1250 Hz and were sampled so that both tones in a pair were either above or below 1000 Hz (as in Jaffe-Dax et al., 2017). The ratio of the two tone frequencies ranged from 1.01 to 1.25 (0.17 to 3.86 semitones). Tone onsets were separated by 2 seconds, resulting in a working memory delay (inter-stimulus-interval) of 1.95 seconds. After the second tone, participants had to specify which tone had a higher pitch. Upon response, participants were immediately presented with visual feedback for 400 ms: "Correct" in green if they got the correct answer, or "Incorrect" in red if they were wrong. Participants had 1300 ms to respond after the onset of the second tone, and if they did not respond within this interval, then they saw "Miss" in red on screen for 400 ms. This "time out" interval is not stressful, as the median response time across all

participants was 646 ms with a range of 398 ms to 923 ms. The next trial was always presented four seconds after the first tone onset of the previous trial (see Figure 1A), irrespective of the speed of the participant's response. The inter-trial interval (2 s after the second tone) was set to be equal to the within-trial inter-stimulus interval (2 s) to reduce the effect of within-trial adaptation, which might differ between the groups as it relates to implicit learning (Jaffe-Dax et al., 2017; Regev et al., 2021). The sounds were presented over Sennheiser HD 25 headphones at a comfortable sound level.

Before starting the formal experiment, participants completed a training set of 10-30 trials, presented in blocks of 10 trials with a break in between. To make the task easier, during these training trials the frequency range of the tones was increased to 600 – 1600 Hz, and the tone ratios were randomly selected between 1.35 – 1.4 (equivalent to 5.20 – 5.83 semitones). Participants proceeded to the main experiment upon reaching 8/10 trials correct in a block, or upon completing the three 10-trial training blocks, even if this threshold performance was not attained (this was the case for 7/33 ASD participants and 5/35 non-autistics). In the main experiment, participants performed 121 trials without breaks.

Behavioral analysis

Recent trials biased the decision of participants. This bias can be viewed as a contraction of the perceptual memory of the first tone in the trial (whose representation is noisier by the time the second tone is presented) towards the expected frequency, namely the geometric mean of the tone frequencies presented first in the previous trial. Consequently, in trials where the first tone's frequency is closer than the second tone's frequency to geometric mean of the previous trial performance improves (as representation of the first tone becomes more distant from that of the second tone in the trial, "bias+" trials). By contrast, in trials where the first tone is farther from the previous trial's frequencies, the representation of the distance between the tones decreases, hampering performance ("bias-" trials). Previous work has shown that people with autism are less biased by the recent trial (Lieder et al., 2019).

Performance was assessed by calculating the proportion of correct trials, ignoring misses. To quantify the effect of recent history on choice bias, we split trials into "bias+"

157 and "bias-" trials and calculated the difference in performance between these sets of trials.
158 To understand the effects of external feedback ("Correct" or "Incorrect" trials) and trial
159 difficulty on subsequent behavior, we further grouped trials based on the participant's
160 accuracy (Correct/Incorrect) and the trial difficulty (Easy/Hard trials). For trial difficulty,
161 easy versus hard trials were identified with respect to the median frequency difference
162 which was 1.242 semitones. Trials with smaller frequency difference were tagged "hard"
163 whereas trials with larger frequency difference were tagged "easy". We asked how implicit
164 strategies differed between trials that followed easy ones and trials that followed difficult
165 ones.
166
167 *EEG recording and preprocessing*
168 Biosemi ActiveTwo with a 32-channel layout was used to record EEG while
169 participants performed the task. Additionally, external channels were recorded bilaterally
170 on the mastoids, to the outside of each eye, above each eye, and on the nose (7 external
171 electrodes total). The EEG data was digitized at 512 Hz.
172 All preprocessing was done in Matlab. After recording, the EEG data was filtered
173 using a 4th order zero-phase Butterworth filter with -3 dB cutoff frequencies of 1-30 Hz.
174 Eyeblinks and eye motion artifacts were isolated using fastICA (Hyvärinen and Oja, 2000)
175 applied to the 32 EEG channels as well as the externals and then removed. The data was
176 re-referenced to the average of the mastoid channels. If one of the mastoid channel's
177 variance exceeded the median plus two times the interquartile range of the main 32 EEG
178 channels (not including the externals), the mastoid channel was deemed too noisy and only
179 the other mastoid channel was used for reference. In two non-autistic participants, both
180 mastoid channels exceeded this range, and we used the average of neighboring P7 and P8
181 electrodes instead.
182
183 *Calculating evoked responses*
184 All analyses of the EEG were done using channel Cz referenced to the mastoids.
185 Event-related potentials (ERP) to the stimuli and the button presses (the participant's
186 response in the task) were first calculated individually for each participant. ERPs to the
187 stimuli were baselined to the interval of -100 to 0 ms before the stimulus, and ERPs to the

188 button presses were baselined -400 to -200 ms; we chose a different interval before the
189 button press to avoid a negativity that appeared between -200 to 0 ms in some participants.
190 To avoid the influence of outlier EEG signals on the evoked responses, all ERPs were
191 calculated using a spline-based median of the evoked responses (see Zuk et al., 2021 for
192 information on the spline transform). First, each EEG segment was transformed into a basis
193 of 173 cubic splines which enforced a smooth representation of the EEG segments with a
194 -3 dB cutoff frequency around 40 Hz. Second, the median of the basis spline weights across
195 trials was calculated. Then, the basis spline weights were transformed back into the time
196 domain for interpretation, plotting, and further analysis. These steps were also used when
197 calculating the median ERP across participants. For the shaded regions in the plots of the
198 ERP (e.g. Figure 1B), a bootstrap resampling procedure with replacement was done across
199 participants and repeated 200 times to calculate the 95% confidence interval of the median
200 ERP. Analyses of individual participant's ERP magnitudes are based on the minima or
201 maxima of the median ERPs in the time domain.
202
203 Study 2
204 Study 2 was a replication of Study 1, and the same software and hardware were
205 used for administering this experiment. 51 participants (26 non-autistics and 25 ASD) took
206 part in this replication. Of these, 18 non-autistics and 18 ASD had previously participated
207 in Study 1. The time interval between Study 1 and Study 2 was 1-2 years.
208
209 *Group classification*
210 To explore group-level classification, we trained a linear support vector machine
211 (SVM) with second-degree polynomial features using scikit-learn (Python). We chose this
212 model architecture because it performed consistently well across a range of data splits (80%
213 for training, 20% for testing) and performed at least as well as alternative approaches,
214 including random forest, logistic regression, decision trees, and more complex SVMs with
215 radial basis function (RBF) kernels. The feature set included behavioural recency bias after
216 correct and easy trials, the ERP peak-to-peak amplitude difference between easy and hard
217 trials, and FRN magnitude. The preprocessing pipeline involved z-scoring and polynomial
218 feature expansion (degree = 2, excluding the bias term). The regularisation parameter (C),

which controls the trade-off between margin size and classification error, was tuned using a grid search over logarithmically spaced values between 0.01 and 100, with 5-fold cross-validation on the training set. Model performance was evaluated using accuracy and area under the ROC curve (AUC) on a held-out test set, and statistical significance was assessed with 1000-iteration permutation testing. For the main analysis, we performed 20 stratified random 80/20 train–test splits and reported average test accuracy and AUC across seeds. We then examined a representative model using the most common hyperparameter (C=0.1) with model performance (accuracy) closest to the median performance, and we used this model to assess the contribution of each feature. Additional evaluations tested generalization across participants (training on unique individuals and testing on non-unique, or overlapping, individuals).

Feature contribution was assessed by inspecting the weights associated with individual features and calculating the amount of impact each weight had on the data. In general, linear SVM classifiers are trained to choose the weights w associated with each feature or interaction feature so that the decision function $f(x)=w^{\top}x+b$ best classifies positive values of $f(x)$ to one class and negative values to the other, where x is a vector for each sample containing feature values and b is a (constant) bias term (here, non-autistic and ASD are the positive and negative classes respectively). By isolating just $w^{\top}x$ for each feature, we can compare the feature's contribution to the classifier. A feature that is a good discriminator (and has low class overlap) should transform data from a non-autistic participant to a positive value and an ASD participant to a negative value. We plot these feature contributions using a kernel density estimator (KDE) in Figure 7B-D.

Results

Event-related potential (ERP) and behavioral responses to stimuli are similar for people with and without ASD

Participants performed two-tone pitch discrimination, during which we recorded EEG. On each trial, two pure tones with different frequencies were presented serially. Participants were asked to determine which of the two tones had a higher pitch (Figure 1A, in the example the second tone frequency is "higher"). Upon responding, they were immediately presented with visual feedback: "Correct" in green text or "Incorrect" in red

text. Since we were interested in the ability to track performance, which might depend on performance level, we aimed to attain similar success rates for the two groups. We used a simple temporal protocol, with two-second intervals both between stimuli and between trials, in order to minimize possible effects of within-trial neural adaptation that may differ in neurodivergent groups (Jaffe-Dax et al., 2017). With these intervals, the peaks of the three main EEG response components (N1 between 70-130 ms, P2 between 150-230 ms, and P3 between 250-350 ms) to the two tones (f1 and f2) were similar in the two groups (Rank-sum test: 1st N1, $z = -0.52$, $p = 0.606$; 2nd N1, $z = -0.11$, $p = 0.912$; 1st P2, $z = 0.18$, $p = 0.854$; 2nd P2, $z = -0.43$, $p = 0.668$; 1st P3, $z = -1.88$, $p = 0.061$; 2nd P3, $z = -0.87$, $p = 0.384$). We found difference in N1 magnitudes between the tones, indicating no significant within-trial adaptation (signed-rank test: Non-autistics, $z = -0.87$, $p = 0.385$; ASD, $z = -0.85$, $p = 0.396$). P2 showed a weak, though significant, adaptation in non-autistics (Non-autistics, $z = -2.23$, $p = 0.023$), which was not significantly different from the absence of adaptation in ASD (ASD, $z = -0.96$, $p = 0.339$; rank-sum test for the difference between groups: $z = -0.92$, $p = 0.357$). As expected, both groups had a P3 component following the second tone (Figure 1B), which characterizes time points where a categorical decision is made ("which tone was higher?"; Nahum et al., 2010; Jaffe-Dax et al., 2017). The P3 magnitude was similar for the two groups (rank-sum test: $z = 0.209$, $p = 0.835$). Behaviorally, both groups had similar median success rates in the task (Figure 1C; Non-autistics = 71.9%, ASD = 70.0%; rank-sum test: $z = 0.76$, $p = 0.447$), though the cross-participant distribution was broader in the ASD group.

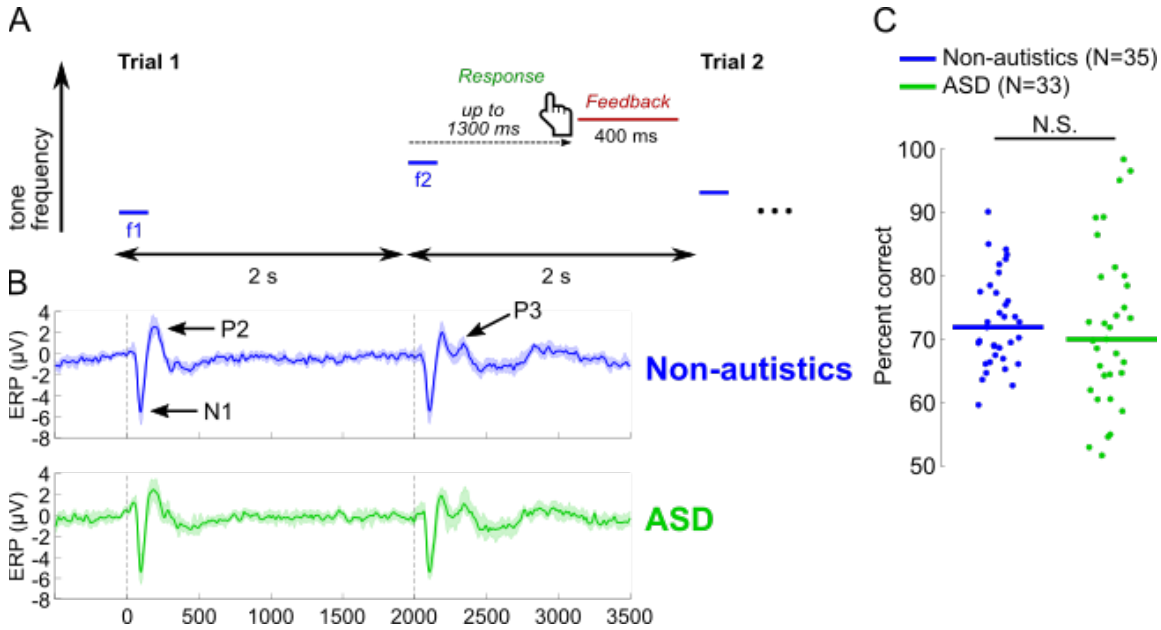

Figure 1: Non-autistics and ASD show similar performance and similar ERP responses to the tones in each trial. (A) A schematic diagram of a trial's temporal structure. Participants performed a two-tone discrimination task where they had to determine if the pitch of the second tone was higher than that of the first. Visual feedback ("Correct" or "Incorrect") was given immediately upon response in every trial. (B) Grand average response (median of participants' median ERPs, see Methods) for electrode Cz. Time zero denotes the onset of the first tone on each trial. Thick lines show the median across participants' median ERPs. Shaded regions show the 95% confidence interval of the median using a bootstrap resampling across participants. Non-autistics are plotted in blue (top) and ASD in green (bottom). Dashed vertical lines mark the onset of each tone. The three main response components, N1, P2 and P3 are marked on the top plot. (C) Distributions of participants' performance accuracy, in each group. Each dot denotes the percent correct of an individual participant, the bars indicate the median in each group. Median performance was similar in the two groups (no significant difference in a rank-sum test).

The Feedback-Related Negativity (FRN) is significantly weaker in ASD

Under similar task performance and EEG responses to tones in the two groups, we next asked whether the response to external feedback is also similar. The visual feedback was provided immediately (within 20 ms) after the participant's button press. To quantify the EEG response to the feedback, we aligned our analyses to participants' button press. We rejected three ASD participants who performed extremely well in the experiment and thus had too few incorrect answers to get a reliable FRN response (< 6 trials incorrect; 3 best performers, see Figure 1C where these participants performed better than all non-

294 autistics). The EEG signal following "incorrect" feedback is known as the "feedback-
295 related negativity" (FRN) (Miltner et al., 1997). The topography of activity at the time of
296 the negativity (e.g. inset in Figure 2B) has been source-localized in previous studies to
297 ACC (Dehaene et al., 1994; Miltner et al., 1997). Figure 2A plots samples of 12 single
298 participants, six with ASD (right) and six non-autistics (left), matched for the rank order
299 of their FRN with respect to their group. It illustrates that the majority of non-autistic
300 participants (sampled here by 1st to 26th in descending FRN magnitude) have a clear and
301 large FRN, whereas most ASD (also sampled here by 1st to 26th) do not.

302 To quantify the magnitude of the FRN, we determined, for each participant, the
303 minimal response between 200-350 ms after the participant's button press (Figure 2C).
304 This time window captured the span of the negativity for all participants (dashed vertical
305 lines in Figure 2A, median ERP in Figure 2B). FRN magnitude was then quantified as the
306 difference between the minimum in incorrect and the minimum in correct trials. Figure 2C
307 indicates that in both groups the minimum for the evoked response to incorrect trials was
308 more negative than to positive trials (signed-rank test: Non-autistic, $z = 5.16$, $p < 0.001$;
309 ASD, $z = 3.16$, $p = 0.002$). However, this was significantly smaller in ASD participants
310 than non-autistics (Figure 2D, rank-sum test: $z = -3.56$, $p < 0.001$). Importantly, the effect
311 size of this group difference was large (Cohen's $d = 0.90$). For reference, the effect size
312 (Cohen's d) of AQ, a questionnaire specifically designed to detect ASD traits (Baron-
313 Cohen et al., 2001), was 1.10 for our participants, comparable to the effect size of the FRN
314 measure. We did not find a correlation within each group between AQ score and FRN
315 magnitude (Non-autistics, $\rho = -0.274$, $p = 0.117$; ASD, $\rho = -0.062$, $p = 0.733$), perhaps
316 because FRN magnitude also depends on performance levels (in non-autistics it was shown
317 that "incorrect" feedback yields a larger error- or feedback-related negativity when the task
318 is easier (Miltner et al., 1997) and when errors are more unexpected (Gehring et al., 1993;
319 Holroyd and Coles, 2002)).

320 To highlight this difference at the single participant level, figure 2D on the right
321 shows a histogram of the differences between incorrect and correct evoked responses in
322 each group. It shows that by using discriminant boundary of -2.5 μV (close to the median
323 magnitude for ASD participants, plotted in Figure 2D), only a minority of the ASD

participants (13/30) show an FRN magnitude lying below this boundary, whereas nearly all (33/35) of the non-autistics have magnitudes below this boundary.

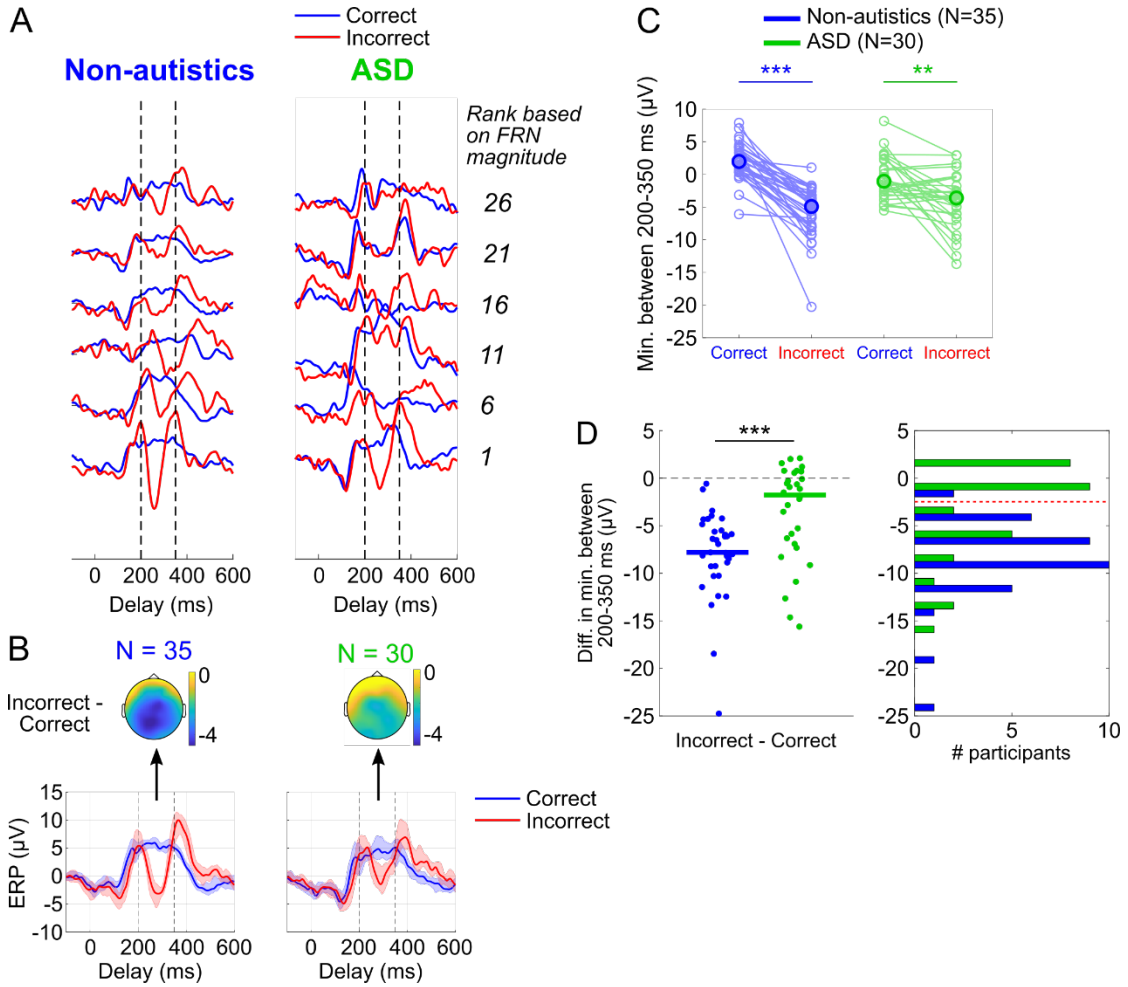

Figure 2: ASD participants' EEG activity in response to external feedback (Feedback-Related Negativity, FRN) is weaker than that of non-autistics'. (A) Single participants' FRN, defined as the difference between (minimal) responses following "Correct" (blue) and "Incorrect" (red) feedback within 200-350 ms after their behavioral response (range marked by the dashed lines). FRNs were ranked according to these differences. We plotted responses of six participants from each group, whose ranked magnitude of FRNs were matched with respect to their group (non-autistics on the left, ASD on the right). These ranks are denoted on the right (1st being largest). While all six non-autistics had an FRN, only two of the six plotted autistics' responses manifest a clear FRN. (B) Median ERP across participants, shaded error bars indicate the bootstrapped 95% confidence intervals of the median. The topography of the difference in voltage between Incorrect and Correct trials averaged within this time interval is shown above each plot. (C) Quantification of the FRN magnitude as the difference between the ERP minimum for incorrect and correct trials, within the range 200-350 ms as indicated in A and B, plotted separately for each participant. At the group level both

groups show a significant FRN. (D) Individual dots show the difference in magnitude between correct and incorrect, and the horizontal bars denote the median of each group. The group difference in FRN is highly significant (*** = $p < 0.001$). On the left, each symbol denotes a single participant. Note that more than half of the ASD participants (above the green line) are distributed symmetrically around zero difference, meaning that they have no clear FRN. By contrast, only 2/35 neurotypicals do not show a clear FRN (upper-most blue dots in the plot). The histogram on the right also shows the distribution of FRN in each group with 2.5 μV bins. It illustrates the efficiency of FRN magnitude as an indicator of group; a plausible discrimination boundary of -2.5 μV is shown as a red dashed line.

Only non-autistics' EEG responses are modulated by task difficulty

External feedback is often processed in tandem with one's confidence in their decision (Fleming and Daw, 2017; Grogan et al., 2023). When participants are more confident in their knowledge of task-relevant information they are more likely to use that information (Samaha et al., 2019). This is a form of internal feedback that can also modulate the participants' behavior in the task, but does it affect their EEG responses to feedback?

As a proxy for participants' assurance of their perceptual decision, we separated the trials into "Hard" and "Easy" (Figure 3A) using a median of the frequency difference in the trial as the threshold (the threshold of 1.242 semitones was the median frequency difference across all trials and participants). We reasoned that in easy trials, participants are more confident in their decision. Indeed, non-autistics' EEG responses were modulated by task difficulty (Figure 3B), such that the amplitude was larger following hard trials. We used the peak-to-peak difference in the evoked response, calculated as the difference between the minimum in 0-150 ms following the participant's button press and the maximum in 300-450 ms following the button press (see the dashed lines in Figure 3B) to quantify the magnitude of this difference (Figure 3C). We found that, in non-autistics, this difference was significantly larger in hard trials compared to easy ones (signed-rank: $z = -2.69$, $p = 0.007$), suggesting that participants' effort following difficult discriminations modulated the EEG response to feedback. ASD participants, however, did not show any difference ($z = 1.51$, $p = 0.131$). The group difference was significant (rank-sum: $z = -2.98$, $p = 0.028$; Cohen's $d = 0.574$).

Our results indicate that both internal feedback and external feedback significantly modulate the EEG response in non-autistics but not in ASD, showing that ASD participants have weaker processing of both types of feedback.

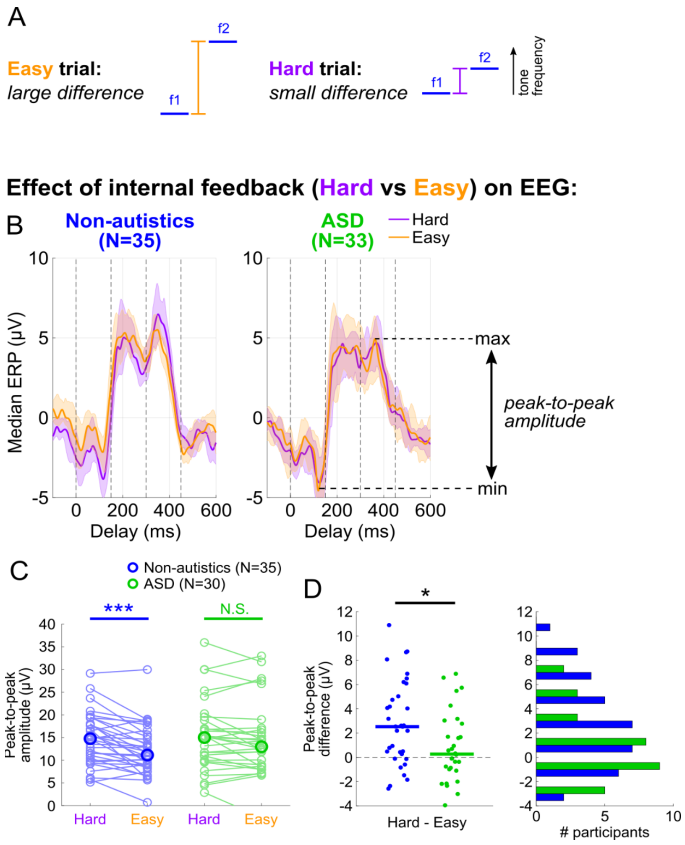

Figure 3: Weaker effects of internal feedback were observed in ASD participants compared to non-autistics. (A) Diagram showing the distinction between easy and hard trials. Easy trials have a larger frequency difference than the median difference across all possible trials (1.242 semitones), while hard trials have frequency differences smaller than this difference. (B) Median evoked responses across participants, time-aligned to the behavioral response, plotted separately for hard (purple) and easy (orange) trials. (C) The peak-to-peak difference in the evoked EEG responses to hard compared to easy trials is significantly larger in non-autistic participants. Note that, here, ASD participants with excellent performance (also excluded from Figure 2) were also rejected to remain consistent with the other measures used later for autism/non-autism classification. (D) Individual dots show the difference in peak-to-peak magnitude between easy and hard trials. The group difference is significant (rank-sum $z = -2.98$, $p = 0.028$, Cohen's $d = 0.574$). * $p < 0.05$, *** $p < 0.001$.

Only non-autistics' perceptual bias is modulated online by task difficulty and success

The modulation of non-autistics' EEG response by task difficulty suggests that it may also have a behavioral impact. We therefore examined how external and internal feedback affect participants' response tendencies in subsequent trials. Previous work had shown that participants' behavioral responses in a two-interval forced choice task are substantially biased by previous trials' statistics (Ashourian and Loewenstein, 2011; Raviv et al., 2012, 2014; Akrami et al., 2018; Lieder et al., 2019). Within a Bayesian framework, it was proposed that since the mental representation of the first tone is noisier, it is contracted more towards an implicit prior (predicted frequency), which is formed by previous trials' frequencies (Ashourian and Loewenstein, 2011) (but see Boboeva et al., 2024). Because of this, when the first tone is closer to this prior, contraction improves performance (consequently termed bias+ trials), whereas when it is farther, performance deteriorates (bias-) (Raviv et al., 2012; Akrami et al., 2018) (Figure 4A).

We calculated the recency bias, namely the bias of the first tone towards the most recent trial's average frequency (Lieder et al., 2019). The bias is the difference in success rate between bias+ trials (where this contraction improves performance) and bias- trials (where performance decreases) (Ashourian and Loewenstein, 2011; Raviv et al., 2012). In line with previous work (Lieder et al., 2019), the bias for ASD participants was on average smaller than non-autistics (rank-sum test: $z = 2.39$, $p = 0.017$). Next, we divided these trials based on the difficulty of the previous trial (easy versus hard, based on the threshold of 1.242 semitones) and based on the participant's accuracy in the previous trial (correct or incorrect) (Figure 4B). All participants were included for previously correct trials. For incorrect trials, we excluded participants who did not have sufficient trials in at least one of the conditions; 32 non-autistics and 28 ASD participants were included.

We first looked at the effect of the previous trial's difficulty (internal feedback) on performance, separately for correct and incorrect trials (external feedback). Importantly, we found that non-autistic participants showed a larger recency bias when the previous trial was easy compared to when it was hard (signed-rank test: $z = -3.01$, $p = 0.003$). In contrast, ASD participants did not show this effect on bias in correct trials ($z = 1.02$, $p = 0.309$), and the bias difference between easy and hard trials when the previous trial was easy was significantly weaker than non-autistics' ($z = 2.88$, $p = 0.004$). The group difference was

420 particularly driven by a difference in bias when participants were previously correct in an
421 easy trial, where non-autistics showed a larger bias than ASD ($z = 2.72$, $p = 0.006$, Cohen's
422 $d = 0.752$).

423 Together, the behavioral group differences suggest that the magnitude of the
424 recency bias and the participant's subsequent decisions are affected by the difficulty and
425 accuracy in the previous trial. When non-autistic participants are confident in the stimulus
426 information from the previous trial (after easy) and their confidence has been externally
427 validated (after correct), the effect on the perceptual prior is stronger on their subsequent
428 decisions. ASD participants have weaker processing of internal feedback, and their
429 decisions were unaffected by the accuracy or internal reliability of their response.
430 Importantly, this observation is in line with previous studies reporting reduced online
431 flexibility in updating perceptual priors based on recent events (Sapey-Triomphe et al.,
432 2021). But while studies assessed stimulus statistics, here we assessed the impact of
433 certainty and show that the effect on recency bias is reliable and effective.

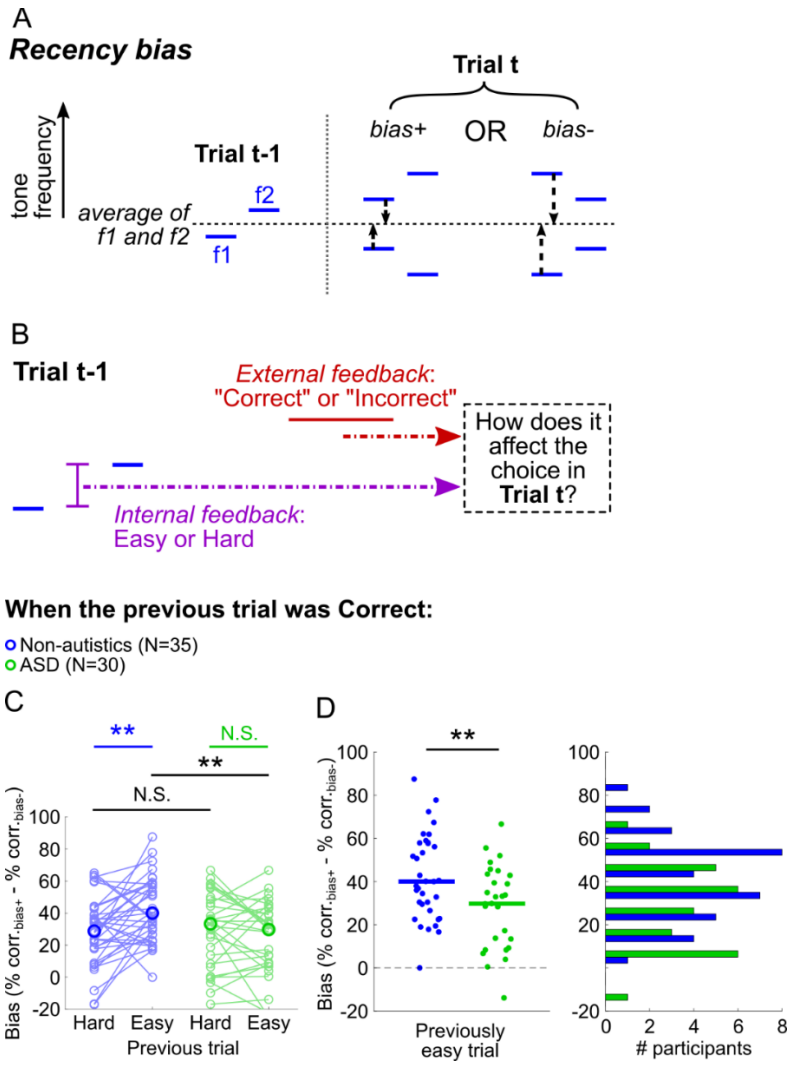

Figure 4: Only non-autistics show a larger recency bias following correct and easy trials. (A) A schematic illustration of recency bias towards recent trial's frequencies in perceptual decisions. In bias+ trials the first tone's frequency is closer to the mean frequency of the previous trial,, whereas in bias- the second tone is closer. Better performance occurs in "bias+" trials versus "bias-" trials though they differ only in their preceding trial's frequency. (B) Diagram illustrating the main question of how internal feedback (Hard / Easy) and external feedback (Correct / Incorrect) in the previous trial affect the participant's choices in the current trial. (C) When non-autistic participants were previously correct yet more confident (easy trials), there was significantly larger recency bias in the subsequent trial. This effect was not found for the ASD group. (D) A closer comparison of the bias difference between groups when the previous trial was both easy and correct, plotted identically to figures 2D and 3D. The group difference was significant ($p = 0.006$, Cohen's d $= 0.752$). ** $p < 0.01$

A mini control experiment – Real feedback is important

Since an FRN response to perceptual performance has not been studied before to the best of our knowledge, and to better understand its nature and the temporal dynamics of the response, we administered a small control experiment with three non-autistic participants and three ASD participants. The non-autistic participants were selected because they had a large FRN in the previous experiment and were available for the follow up, and the ASD participants were selected based on willingness to participate in a follow-up session. All participants did three blocks in a random order, where each block had a different type of feedback: True feedback was identical to the previous experiment (as in Figure 1A); Delayed feedback showed true feedback, but the feedback was delayed by 200 ms relative to the participant's response; and Random feedback showed "Correct" or "Incorrect" randomly after the participant's response. In the latter condition, participants were told that the feedback would be random and uninformative, but they were still instructed to look at the feedback during the task.

Our first test compared true and delayed feedback in non-autistics; for participants who showed a reliable FRN, was the ERP timed to the button press or the feedback? When the feedback was delayed, both the correct response and the incorrect response were also delayed by 200 ms (Figure 5A), indicating that the ERP and FRN was timed to the external feedback rather than to the response itself. Next, we compared the FRN to true and random feedback, asking whether, in the absence of informative feedback, the FRN response disappears. In non-autistics, randomizing the feedback causes the FRN to disappear (Figure 5B). Additionally, non-autistics' positivity 200 ms after a correct response also disappeared, indicating that the signal was related to information in the feedback rather than the visual input. Interestingly, the ASD participants had EEG responses to both correct and incorrect feedback that did not appear to differ between the true feedback and random feedback conditions, unlike the non-autistics (Figure 5B). While preliminary, this result suggests that ASD participants may process feedback the same way irrespective of its informativeness.

Taken together, the mini follow-up study indicates that the responses we observed in non-autistics were time-locked to the feedback. Importantly, both the positivity in the

response to Correct feedback after 200 ms (Figure 5B) and the FRN reflected the processing of errors in the task, rather than visual processing of the feedback.

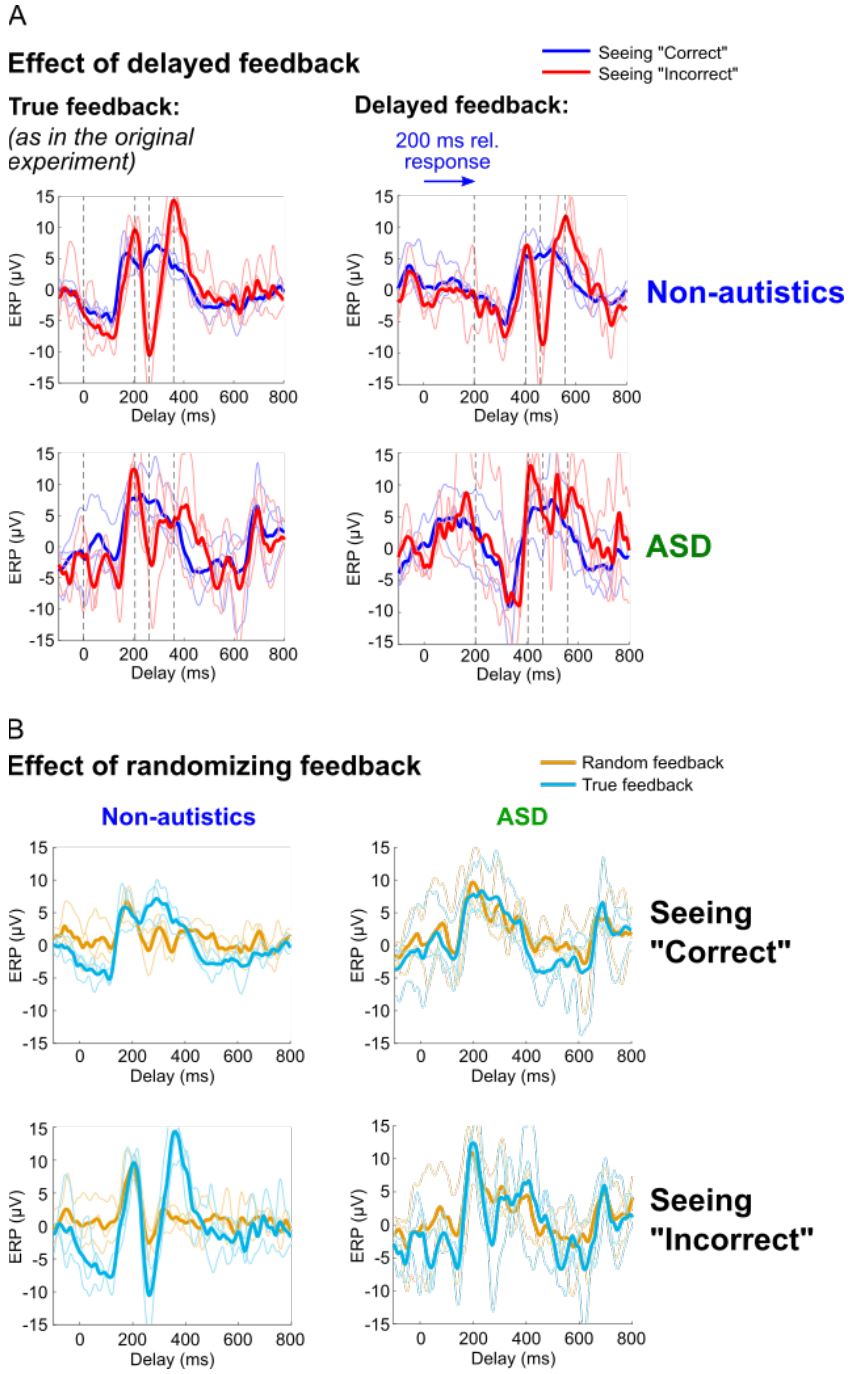

Figure 5: A (mini) follow-up control experiment (3 participants from each group) for feedback timing and validity. (A) Evoked responses to correct (blue) and incorrect (red) feedback presented immediately after the participant's response (left), and delayed by 200ms (right), shown separately for the non-autistic (top) and ASD participants (bottom). For non-ASD participants the FRN response was delayed by 200ms. For autistic participants there is still little to no FRN. Interestingly, at a shorter time window than the FRN

there is a large negative response, which does not differ between correct and incorrect response, and suggests some enhanced anticipation that should be studied further. The thin colored lines are evoked responses for individual subjects, and the thick line is the average. Dashed lines label the peaks and troughs of the evoked response to true feedback (on the left). The dashed lines on the right mark the same time points in A shifted by 200 ms, to make clear the shift in the response. (B) Non-autistics (left) and ASD (right) evoked responses after seeing "Correct" or "Incorrect" on screen when the feedback was true (blue, as in A) or randomized (orange).

Follow-up study replicates prior results

In Study 1 we got impressive results for three aspects of online performance monitoring. Given that the two-tone discrimination protocol is composed of 121 trials only, we wanted to verify that these parameters are replicable and that a reliable classification can be done based on these parameters. We therefore invited participants to return for a follow-up assessment with the same protocol, approximately 1-2 years after participating in Study 1. Overall, we had 51 participants in this study (25 non-autistics and 23 ASD), and 36 of them were people who also participated in study 1, allowing assessment of within participants' replicability. Two ASD participants had above 95% trials correct in discrimination (as they did in Study 1) and hence were excluded from FRN calculations (in both Study 1 and Study 2), leaving 18 non-autistics and 16 ASD who participated in both studies. As in Study 1, overall performance was similar for the two groups (mean percent correct in non-autistics = 79%, in ASD = 78%).

All three measures related to performance monitoring that were significantly different between groups in Study 1 were also significantly different in Study 2 (Figure 6A-C): FRN magnitude was larger in non-autistics (Figure 6A, $z = -2.87$, $p = 0.004$, Cohen's $d = 0.852$); ASDs' difference in ERP magnitude following easy and difficult trials was smaller than non-autistics' (Figure 6B, $z = 2.18$, $p = 0.028$, Cohen's $d = 0.680$) and ASDs' reduced flexibility, as manifested in their reduced effect of prior's magnitude following easy trials, was smaller (Figure 6C, $z = 2.70$, $p = 0.006$, Cohen's $d = 0.883$).

Importantly, within people who participated in both studies, each of these measures were correlated, even though more than a year passed between the two measures (Figure 6D-F; Pearson's correlations – FRN: Non-autistics, $r = 0.484$, $p = 0.042$; ASD, $r = 0.760$, $p < 0.001$; Hard-Easy Peak-to-peak: Non-autistics, $r = 0.437$, $p = 0.070$; ASD, $r = 0.512$, p

= 0.042; Bias following easy and correct: Non-autistics, r = 0.540, p = 0.021; ASD, r = 0.767, p < 0.001). The replicability and correlation of three measures indicate that each of these three measures reliably characterize the participants.

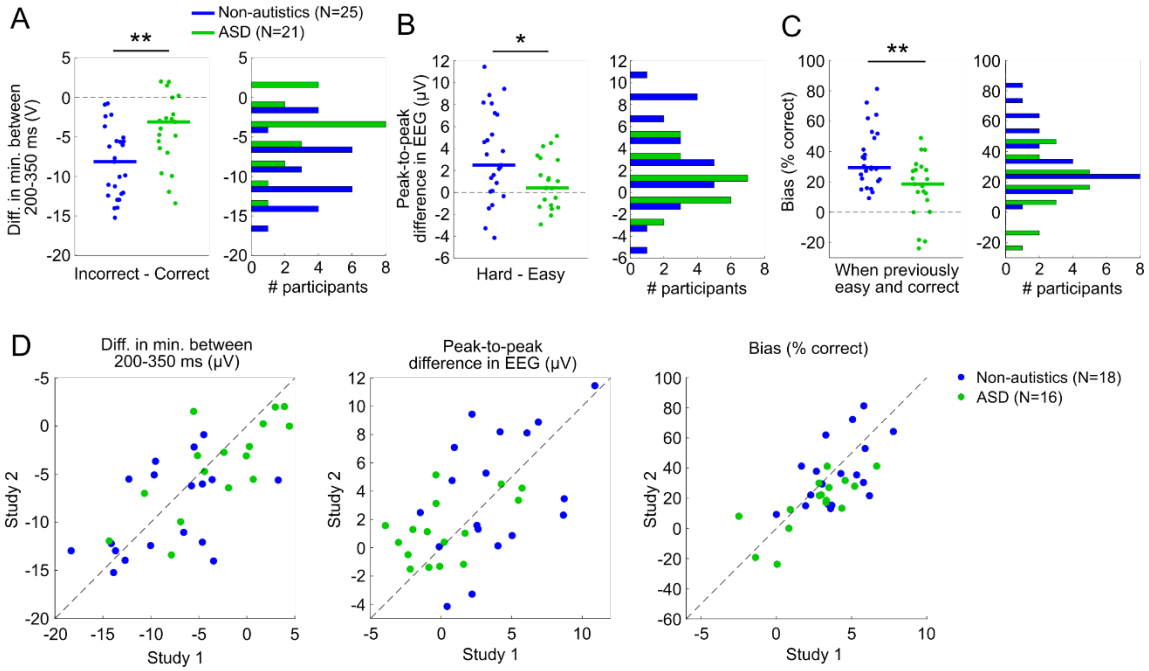

Figure 6 – In Study 2, all parameters were replicated and correlated within participants across sessions. (A) Single participant distribution of FRN magnitude, plotted identically to Figure 2D (Cohen's d = 0.852). (B) Single participant distribution of the difference between Hard and Easy peak-to-peak magnitudes of the EEG following the participant's button press, plotted identically to Figure 3D (Cohen's d = 0.680). (C) Single participants' distribution of magnitude of online bias updating (following easy and correct trials), plotted identically to Figure 4D (Cohen's d = 0.883). (D) Within participant correlation of these parameters between study 1 and study 2, for the subgroup that participated in both studies.

Additionally, we combined the participants for study 1 and study 2 (taking only their first session; 42 non-autistics, 35 ASD) and examined if there were correlations between these measures. Interestingly, we found no correlations between measures within each group (Spearman's correlations, two-tailed, FRN vs peak-to-peak difference: non-autistics, ρ = 0.06, p = 0.727; ASD, ρ = -0.19, p = 0.262; FRN vs bias after easy and correct: non-autistics, ρ = 0.12, p = 0.442; ASD, ρ = -0.12, p = 0.483; peak-to-peak difference vs bias after easy and correct: non-autistics, ρ = 0.21, p = 0.176; ASD, ρ = 0.23, p = 0.197).

542 This implies that the measures may represent distinct cognitive processes which are all
543 generally weaker in autistics.
544
545 *Combining these neurocognitive measures classifies ASD participants*
546 Given that the three performance-monitoring measures were both replicable across
547 sessions and were uncorrelated with each other, we reasoned that a classification model
548 based on all three could be both effective and informative, as it can point to the specific
549 performance monitoring characteristics that are atypical in each participant.
550 We trained a linear support vector machine (SVM) classifier using second-degree
551 polynomial features (for details see Methods). This model choice allows for linear decision
552 boundaries while capturing both main effects and pairwise interactions among features.
553 Over random train-test splits of all participants, the model achieved a median test accuracy
554 of 80.4% (±7.0%) and a mode (for 6/20 train-test splits) of 82.6%. Similarly, the median
555 receiver operating characteristic (ROC) area under the curve (AUC) was 0.86 (±0.06).
556 AUC represents the probability that a randomly chosen non-autistic participant's data will
557 be ranked higher by the model than a randomly chosen autistic participant's data, where an
558 AUC of 0.5 is chance level. An AUC above 0.8 is indicative of a good model classifier
559 (Hanley and McNeil, 1982).
560 Next, a representative model using the most common hyperparameter setting with
561 performance close to the median accuracy across all models was selected to examine
562 feature contributions to classification. This specific model achieved 82.6% test accuracy
563 and an AUC of 0.88 (Fig. 7B-D; see Methods). FRN and behavioral recency bias following
564 easy and correct trials contributed more to the separation of classes than the peak-to-peak
565 ERP magnitude, where separation can be understood based on the amount of overlap
566 between the density distributions for non-autistic and ASD groups. Interaction effects
567 between the peak-to-peak amplitude and FRN magnitude as well as between FRN
568 magnitude and bias were also contributive (data not shown).
569 Performance remained robust when training on participants' initial experience with
570 the experiment (i.e. all participants in study 1 and non-repeating participants in study 2)
571 and testing on individuals who had already participated once before (participants in study
572 2 who also completed study 1) with an accuracy of 77% and AUC of 0.85. Permutation

tests confirmed that classification performance was significantly above chance across all data splits ($p < 0.001$).

Overall, these results demonstrate that these three behavioral and neurophysiological measures alone can identify an autistic participant with good accuracy, implying that they are a promising biomarker of autism.

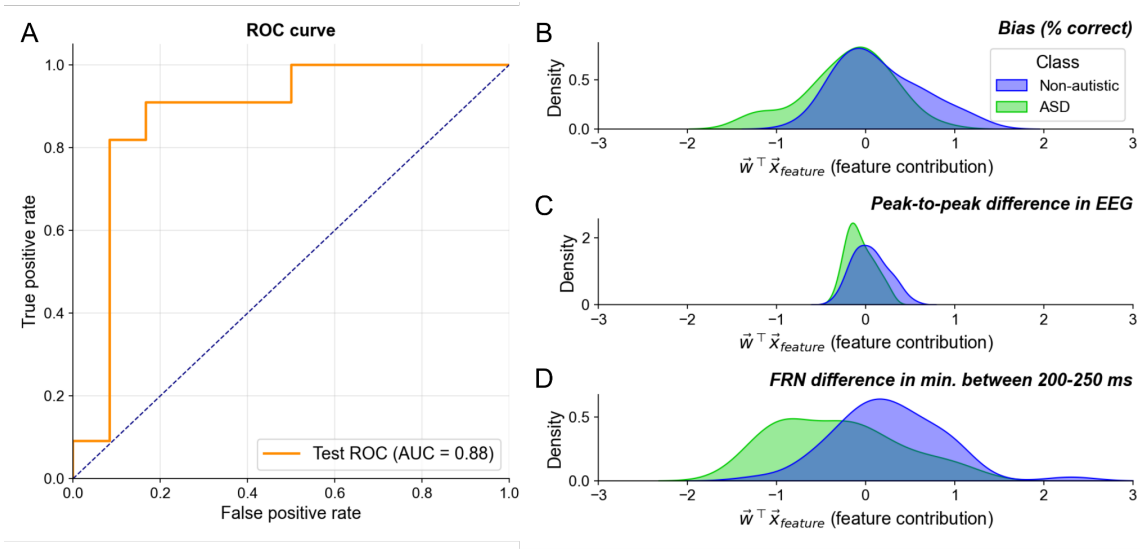

Figure 7: Group classification based on the three parameters of performance monitoring. (A) Receiver operator characteristic (ROC) curve for the representative model (AUC is 0.88). (B-D) shows the contribution of each main parameter to this classification model, with negative values favoring ASD classification, and positive values favoring NT. The FRN magnitude difference (D) has the largest impact (i.e. largest effect size) since there are fewer areas of overlap between the true classes and each is pushed away, while behavioral bias (B) provides a smaller but still significant contribution. The ERP peak-to-peak amplitude (C) contributes much less.

Discussion

In this study, non-autistics and autistic individuals performed a two-tone discrimination task, exhibiting similar levels of performance. Yet autistics' flexibility during performance was much reduced; while non-autistic individuals adapted their reliance on perceptual predictions and their EEG responses to feedback and task difficulty, people with autism did not. These neurocognitive measures were highly robust, but interestingly they were not correlated with each other. Together, however, they provided a reliable, theory-based, classifier of autism.

First, we found that error processing of external feedback, manifested by a large EEG component termed the feedback-related negativity (FRN) is produced by almost all non-autistics. By contrast, FRN was weaker or absent among most ASD participants. Groups also differed in their processing of internal feedback: non-autistics had a larger EEG response to hard trials (small frequency difference) compared with easy trials, while no such difference was found in autistic participants. Behaviorally, non-autistics showed a larger perceptual bias after correct and easy trials, indicating that estimates of reliability and confidence affect one's subsequent perceptual prior (Samaha et al., 2019). ASD participants did not show this trend, demonstrating less flexibility in updating their perceptual prior that stems from reduced monitoring of their internal evaluation of performance.

Since ACC is involved in both error monitoring and evaluating perceptual priors, and since both characterize autism, we initially thought they may stem from a common mechanism. However, ACC's function in cognition is multifaceted, responsible for monitoring many pieces of information that inform subsequent actions. Indeed, ACC has a complex parcellation (Bush et al., 2000), with different subregions functionally connected to various areas of frontal, parietal, and limbic regions of the brain (Devinsky et al., 1995; Margulies et al., 2007). Though all three of our measures have been associated with ACC, the lack of correlation between them may not be surprising, as perceptual and response bias may stem from different processing levels (Margulies et al., 2007; Fritsche et al., 2017; Oesch et al., 2025). Taken together, our results suggest that different atypicalities associated with ACC may contribute to autism spectrum.

The reduction in serial effects in perceptual bias that we found here links two findings in the literature. One indicates that ASD participants use recent prior information less than non-autistics in perceptual tasks. This is the case when inter-trial intervals are short and require fast trial-to-trial updating (Lieder et al., 2019; Boboeva et al., 2024) or when external statistics are volatile such that optimal inference yields large trial-to-trial prediction updating (Lawson et al., 2017; Noel et al., 2024). Similarly, other studies have indicated poor online error correction in motor tasks (Noel et al., 2021; Vishne et al., 2021; Kasten et al., 2023) and perhaps reduced conflict resolution when fast responses are

required (South et al., 2010). Both are suggested to be mediated by an ACC pattern of activation (Botvinick et al., 2001; Kolling et al., 2016a).

The weaker FRN response in the ASD group aligns with other human studies, using functional imaging, showing weaker activation in autistics' ACC (Sapey-Triomphe et al., 2023). ACC encodes the expectation of errors or conflicting information (Carter et al., 1998; Botvinick et al., 1999; Kennerley et al., 2006; Brown and Braver, 2007) (for review see also (Botvinick et al., 2001)) which is subsequently important for identifying changes in task rules and the value of future choices (Kolling et al., 2016b; Takeuchi et al., 2022). We now show that these ACC signals (as measured with the FRN and peak-to-peak differences in the EEG response) are produced during error monitoring in perceptual discrimination. This also affected the majority of the ASD participants (17/30 ASD participants had magnitudes above -2.5 μV) and its effect size (Cohen's d = 0.90, see Figure 2) was comparable to that of the AQ self-report questionnaire (Baron-Cohen et al., 2001) (d = 1.10, see Table 1), designed specifically to detect self-report social difficulties in autism.

Previous EEG work had also shown that the error-related negativity in the absence of external feedback (hence reflecting internal feedback) is weaker in ASD participants compared to non-autistics (Vlamings et al., 2008; Sokhadze et al., 2010; South et al., 2010; Hüpen et al., 2016). However, evidence for group differences in the FRN were mixed (Hüpen et al., 2016). Importantly, much of this work had focused on the effect of feedback on reward expectations (Groen et al., 2008; Larson et al., 2011; Bellebaum et al., 2014). Specifically, these studies varied the probability that the participant could get a reward after making a decision. Our results imply that group differences in error monitoring require the formation of perceptual expectations rather than processing of the reward outcome itself. In prior work, the actual percept generating the error could be dismissed in subsequent trials because the probability of a reward was randomized (Larson et al., 2011), while in our study it is well-established that participants use previous information to both learn about the task rules (Lak et al., 2020; Pedrosa et al., 2023) and update their perceptual bias in subsequent trials (Raviv et al., 2012; Akrami et al., 2018; Lieder et al., 2019).

By combining all three measures of reduced flexibility in updating implications of outcomes, accurate classification of ASD versus non-autistic individuals (above 80%) was

possible, highlighting the potential utility of this multi-dimensional approach for identifying neurocognitive signatures of autism. Importantly, the measures were also reliably reproduced within participants though the studies were conducted more than a year apart. Taken together, these results demonstrate the importance of several independent components in error monitoring during a perceptual task and merge two lines of observations in the ASD literature. It is notable that despite the variability within groups, each of these three measures reliably differed between groups.

Several studies have addressed the question of biomarkers for ASD. They compared motor characteristics during development (Wu et al., 2018), based on atypical brain connectivity (Pagnozzi et al., 2018), altered event-related potentials (Sysoeva et al., 2018; Mason et al., 2022), differences in neurochemical (Yizhar et al., 2011; Robertson et al., 2016; Marotta et al., 2020) or immunological profiles (Xu et al., 2015; Masi et al., 2017). However, the heterogeneity of ASD complicates biomarker development (Charman et al., 2017; Frye et al., 2019; Jensen et al., 2022). Our study suggests that slower updating of priors is a general trait. It is in line with a recent rodent study, which assessed three genetic mouse models of ASD, and found that all underutilize adaptive statistical regularities during decision-making, and frontal cortices underweight sensory observation and prediction errors (Noel et al., 2024). Similarly, reduced online error monitoring and use of perceptual priors was found before in several different paradigms (South et al., 2010; Noel et al., 2021, 2022; Sapey-Triomphe et al., 2021, 2023), further supporting the hypothesis that this is a prevailing characteristic.

Taken together, our findings open a promising avenue for developing objective, reliable, neurocognitive markers for better understanding of autism and for measuring the specific impact of intervention on behavioral monitoring in autism.

Acknowledgements

This work was funded by the Israel Science Foundation (grant 1731/24) and by the European Research Council (grant 833694). NZ was also funded by the ELSC-SWC Postdoctoral Fellowship, co-advised by MA and AA. YW was also funded by the ELSC PhD Fellowship. KD was funded by the Gatsby Charitable Foundation (GAT3850). AA was funded by the Gatsby Charitable Foundation (GAT3755) and the Wellcome Trust (219627/Z/19/Z).

Author contributions

NZ, AA, and MA designed the experiment. NZ, JS, YW, VS collected the data. NZ, JS, and KD wrote the analysis scripts and analyzed the data. NZ, JS, KD, AA, and MA drafted the manuscript, and all authors contributed to editing the manuscript.

Data availability

All EEG and behavioral data from the experiment are available here: <OSF link>. The code used to run the experiment and analyze the data can be found here: https://gitlab.com/auditory-predictive-coding/erp-adaptation. The code used to run the classification analysis can be found here: https://github.com/kiradust/asd-cat.

Declaration of interests

The authors report no competing interests.

References

Akrami A, Kopec CD, Diamond ME, Brody CD (2018) Posterior parietal cortex represents sensory history and mediates its effects on behaviour. Nature 554:368–372.

Ashourian P, Loewenstein Y (2011) Bayesian Inference Underlies the Contraction Bias in Delayed Comparison Tasks. PLOS ONE 6:e19551.

Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E (2001) The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians. J Autism Dev Disord 31:5–17.

Beckmann M, Johansen-Berg H, Rushworth MFS (2009) Connectivity-Based Parcellation of Human Cingulate Cortex and Its Relation to Functional Specialization. J Neurosci 29:1175–1190.

Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS (2007) Learning the value of information in an uncertain world. Nat Neurosci 10:1214–1221.

Bellebaum C, Brodmann K, Thoma P (2014) Active and observational reward learning in adults with autism spectrum disorder: relationship with empathy in an atypical sample. Cognitive Neuropsychiatry 19:205–225.

Bellebaum C, Daum I (2008) Learning-related changes in reward expectancy are reflected in the feedback-related negativity. European Journal of Neuroscience 27:1823–1835.

Boboeva V, Pezzotta A, Clopath C, Akrami A (2024) Unifying network model links recency and central tendency biases in working memory Sharpee TO, Colgin LL, eds. eLife 12:RP86725.

Botvinick M, Nystrom LE, Fissell K, Carter CS, Cohen JD (1999) Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature 402:179–181.

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. Psychological Review 108:624–652.

Botvinick MM, Cohen JD, Carter CS (2004) Conflict monitoring and anterior cingulate cortex: an update. Trends in Cognitive Sciences 8:539–546.

Brown JW, Braver TS (2007) Risk prediction and aversion by anterior cingulate cortex. Cognitive, Affective, & Behavioral Neuroscience 7:266–277.

Bush G, Luu P, Posner MI (2000) Cognitive and emotional influences in anterior cingulate cortex. Trends in Cognitive Sciences 4:215–222.

Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD (1998) Anterior Cingulate Cortex, Error Detection, and the Online Monitoring of Performance. Science 280:747–749.

Cauda F, Geda E, Sacco K, D'Agata F, Duca S, Geminiani G, Keller R (2011) Grey matter abnormality in autism spectrum disorder: an activation likelihood estimation meta-analysis study. J Neurol Neurosurg Psychiatry 82:1304–1313.

Chan AS, Han YMY, Leung WW, Leung C, Wong VCN, Cheung M (2011) Abnormalities in the anterior cingulate cortex associated with attentional and inhibitory control deficits: A neurophysiological study on children with autism spectrum disorders. Research in Autism Spectrum Disorders 5:254–266.

749 Charman T et al. (2017) The EU-AIMS Longitudinal European Autism Project (LEAP):
750 clinical characterisation. Molecular Autism 8:27.

751 D'Cruz A-M, Mosconi MW, Ragozzino ME, Cook EH, Sweeney JA (2016) Alterations
752 in the functional neural circuitry supporting flexible choice behavior in autism
753 spectrum disorders. Transl Psychiatry 6:e916–e916.

754 Dehaene S, Posner MI, Tucker DM (1994) Localization of a Neural System for Error
755 Detection and Compensation. Psychol Sci 5:303–305.

756 Devinsky O, Morrell MJ, Vogt BA (1995) Contributions of anterior cingulate cortex to
757 behaviour. Brain 118:279–306.

758 Falkenstein M, Hohnsbein J, Hoormann J, Blanke L (1991) Effects of crossmodal divided
759 attention on late ERP components. II. Error processing in choice reaction tasks.
760 Electroencephalography and Clinical Neurophysiology 78:447–455.

761 Fleming SM, Daw ND (2017) Self-evaluation of decision-making: A general Bayesian
762 framework for metacognitive computation. Psychological Review 124:91–114.

763 Fritsche M, Mostert P, de Lange FP (2017) Opposite Effects of Recent History on
764 Perception and Decision. Current Biology 27:590–595.

765 Frye RE, Vassall S, Kaur G, Lewis C, Karim M, Rossignol D (2019) Emerging
766 biomarkers in autism spectrum disorder: a systematic review. Annals of
767 Translational Medicine 7:792–792.

768 Gehring W, Coles M, Meyer D, Donchin E (1995) A brain potential manifestation of
769 error-related processing. Electroencephalogr Clin Neurophysiol Suppl 44:261–
770 272.

771 Gehring WJ, Goss B, Coles MGH, Meyer DE, Donchin E (1993) A Neural System for
772 Error Detection and Compensation. Psychol Sci 4:385–390.

773 Gehring WJ, Goss B, Coles MGH, Meyer DE, Donchin E (2018) The Error-Related
774 Negativity. Perspect Psychol Sci 13:200–204.

775 Gehring WJ, Willoughby AR (2002) The Medial Frontal Cortex and the Rapid
776 Processing of Monetary Gains and Losses. Science 295:2279–2282.

777 Groen Y, Wijers AA, Mulder LJM, Waggeveld B, Minderaa RB, Althaus M (2008) Error
778 and feedback processing in children with ADHD and children with Autistic
779 Spectrum Disorder: An EEG event-related potential study. Clinical
780 Neurophysiology 119:2476–2493.

781 Grogan JP, Rys W, Kelly SP, O'Connell RG (2023) Confidence is predicted by pre- and
782 post-choice decision signal dynamics. Imaging Neuroscience Available at:
783 https://doi.org/10.1162/imag_a_00005 [Accessed August 10, 2023].

Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29–36.

Henderson H, Schwartz C, Mundy P, Burnette C, Sutton S, Zahka N, Pradella A (2006) Response monitoring, the error-related negativity, and differences in social behavior in autism. Brain and Cognition 61:96–109.

Holroyd CB, Coles MGH (2002) The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. Psychological Review 109:679–709.

Hüpen P, Groen Y, Gaastra GF, Tucha L, Tucha O (2016) Performance monitoring in autism spectrum disorders: A systematic literature review of event-related potential studies. International Journal of Psychophysiology 102:33–46.

Hyvärinen A, Oja E (2000) Independent component analysis: algorithms and applications. Neural Networks 13:411–430.

Jaffe-Dax S, Frenkel O, Ahissar M (2017) Dyslexics' faster decay of implicit memory for sounds and words is manifested in their shorter neural adaptation King AJ, ed. eLife 6:e20557.

Jensen AR, Lane AL, Werner BA, McLees SE, Fletcher TS, Frye RE (2022) Modern Biomarkers for Autism Spectrum Disorder: Future Directions. Mol Diagn Ther 26:483–495.

Kasten K, Jacoby N, Ahissar M (2023) Poor synchronization yet adequate tempo-keeping in adults with autism. Autism Research 16:1161–1173.

Kennerley SW, Walton ME, Behrens TEJ, Buckley MJ, Rushworth MFS (2006) Optimal decision making and the anterior cingulate cortex. Nat Neurosci 9:940–947.

Kolling N, Behrens T, Wittmann M, Rushworth M (2016a) Multiple signals in anterior cingulate cortex. Current Opinion in Neurobiology 37:36–43.

Kolling N, Wittmann MK, Behrens TEJ, Boorman ED, Mars RB, Rushworth MFS (2016b) Value, search, persistence and model updating in anterior cingulate cortex. Nat Neurosci 19:1280–1285.

Lak A, Hueske E, Hirokawa J, Masset P, Ott T, Urai AE, Donner TH, Carandini M, Tonegawa S, Uchida N, Kepecs A (2020) Reinforcement biases subsequent perceptual decisions when confidence is low, a widespread behavioral phenomenon Salinas E, Frank MJ, Salinas E, Brody CD, Ding L, eds. eLife 9:e49834.

Larson MJ, South M, Krauskopf E, Clawson A, Crowley MJ (2011) Feedback and reward processing in high-functioning autism. Psychiatry Research 187:198–203.

819 Lawson RP, Mathys C, Rees G (2017) Adults with autism overestimate the volatility of
820 the sensory environment. Nat Neurosci 20:1293–1299.

821 Lieder I, Adam V, Frenkel O, Jaffe-Dax S, Sahani M, Ahissar M (2019) Perceptual bias
822 reveals slow-updating in autism and fast-forgetting in dyslexia. Nat Neurosci
823 22:256–264.

824 Margulies DS, Kelly AMC, Uddin LQ, Biswal BB, Castellanos FX, Milham MP (2007)
825 Mapping the functional connectivity of anterior cingulate cortex. NeuroImage
826 37:579–588.

827 Marotta R, Risoleo MC, Messina G, Parisi L, Carotenuto M, Vetri L, Roccella M (2020)
828 The Neurochemistry of Autism. Brain Sciences 10:163.

829 Masi A, Glozier N, Dale R, Guastella AJ (2017) The Immune System, Cytokines, and
830 Biomarkers in Autism Spectrum Disorder. Neurosci Bull 33:194–204.

831 Mason L et al. (2022) Stratifying the autistic phenotype using electrophysiological
832 indices of social perception. Science Translational Medicine 14:eabf8987.

833 Miltner WHR, Braun CH, Coles MGH (1997) Event-Related Brain Potentials Following
834 Incorrect Feedback in a Time-Estimation Task: Evidence for a "Generic" Neural
835 System for Error Detection. Journal of Cognitive Neuroscience 9:788–798.

836 Nahum M, Daikhin L, Lubin Y, Cohen Y, Ahissar M (2010) From Comparison to
837 Classification: A Cortical Tool for Boosting Perception. J Neurosci 30:1128–
838 1136.

839 Noel J-P, Balzani E, Acerbi L, Benson J, Laboratory TIB, Savin C, Angelaki DE (2024)
840 A common computational and neural anomaly across mouse models of
841 autism. :2024.05.08.593232 Available at:
842 https://www.biorxiv.org/content/10.1101/2024.05.08.593232v1 [Accessed July
843 16, 2024].

844 Noel J-P, Shivkumar S, Dokka K, Haefner RM, Angelaki DE (2022) Aberrant causal
845 inference and presence of a compensatory mechanism in autism spectrum disorder
846 Yu X, Shinn-Cunningham BG, Beierholm U, eds. eLife 11:e71866.

847 Noel J-P, Zhang L-Q, Stocker AA, Angelaki DE (2021) Individuals with autism spectrum
848 disorder have altered visual encoding capacity. PLOS Biology 19:e3001215.

849 Oesch LT, Thomas MC, Sandberg D, Couto J, Churchland AK (2025) Anterior cingulate
850 cortex mixes retrospective cognitive signals and ongoing movement signatures
851 during decision-making. Available at:
852 https://www.biorxiv.org/content/10.1101/2025.04.11.648398v3 [Accessed August
853 12, 2025].

854 Pagnozzi AM, Conti E, Calderoni S, Fripp J, Rose SE (2018) A systematic review of
855 structural MRI biomarkers in autism spectrum disorder: A machine learning
856 perspective. International Journal of Developmental Neuroscience 71:68–82.

857 Pedrosa V, Menichini E, Pajot-Moric Q, Vincent P, Zhou L, Teachen L, Latham P,
858 Akrami A (2023) Humans, rats and mice show species-specific adaptations to
859 sensory statistics in categorisation behaviour. :2023.01.30.526119 Available at:
860 https://www.biorxiv.org/content/10.1101/2023.01.30.526119v1 [Accessed July
861 16, 2024].

862 Postema MC et al. (2019) Altered structural brain asymmetry in autism spectrum disorder
863 in a study of 54 datasets. Nat Commun 10:4958.

864 Raviv O, Ahissar M, Loewenstein Y (2012) How Recent History Affects Perception: The
865 Normative Approach and Its Heuristic Approximation. PLOS Computational
866 Biology 8:e1002731.

867 Raviv O, Lieder I, Loewenstein Y, Ahissar M (2014) Contradictory Behavioral Biases
868 Result from the Influence of Past Stimuli on Perception. PLOS Computational
869 Biology 10:e1003948.

870 Regev TI, Markusfeld G, Deouell LY, Nelken I (2021) Context Sensitivity across
871 Multiple Time scales with a Flexible Frequency Bandwidth. Cerebral Cortex
872 Available at: https://doi.org/10.1093/cercor/bhab200 [Accessed October 7, 2021].

873 Robertson CE, Ratai E-M, Kanwisher N (2016) Reduced GABAergic Action in the
874 Autistic Brain. Current Biology 26:80–85.

875 Samaha J, Switzky M, Postle BR (2019) Confidence boosts serial dependence in
876 orientation estimation. Journal of Vision 19:25.

877 Sapey-Triomphe L-A, Pattyn L, Weilnhammer V, Sterzer P, Wagemans J (2023) Neural
878 correlates of hierarchical predictive processes in autistic adults. Nat Commun
879 14:3640.

880 Sapey-Triomphe L-A, Timmermans L, Wagemans J (2021) Priors Bias Perceptual
881 Decisions in Autism, But Are Less Flexibly Adjusted to the Context. Autism
882 Research 14:1134–1146.

883 Simms ML, Kemper TL, Timbie CM, Bauman ML, Blatt GJ (2009) The anterior
884 cingulate cortex in autism: heterogeneity of qualitative and quantitative
885 cytoarchitectonic features suggests possible subgroups. Acta Neuropathol
886 118:673–684.

887 Sokhadze E, Baruth J, El-Baz A, Horrell T, Sokhadze G, Carroll T, MD AT, PhD LS,
888 MD MFC (2010) Impaired Error Monitoring and Correction Function in Autism.
889 Journal of Neurotherapy Available at:

890 https://www.tandfonline.com/doi/abs/10.1080/10874201003771561 [Accessed
891 April 25, 2024].

892 South M, Larson MJ, Krauskopf E, Clawson A (2010) Error processing in high-
893 functioning Autism Spectrum Disorders. Biological Psychology 85:242–251.

894 South M, Newton T, Chamberlain PD (2012) Delayed reversal learning and association
895 with repetitive behavior in autism spectrum disorders. Autism Research 5:398–
896 406.

897 Stevens FL, Hurley RA, Taber KH (2011) Anterior Cingulate Cortex: Unique Role in
898 Cognition and Emotion. JNP 23:121–125.

899 Sysoeva OV, Constantino JN, Anokhin AP (2018) Event-related potential (ERP)
900 correlates of face processing in verbal children with autism spectrum disorders
901 (ASD) and their first-degree relatives: a family study. Molecular Autism 9:41.

902 Takeuchi D, Roy D, Muralidhar S, Kawai T, Bari A, Lovett C, Sullivan HA, Wickersham
903 IR, Tonegawa S (2022) Cingulate-motor circuits update rule representations for
904 sequential choice decisions. Nat Commun 13:4545.

905 Thakkar KN, Polli FE, Joseph RM, Tuch DS, Hadjikhani N, Barton JJS, Manoach DS
906 (2008) Response monitoring, repetitive behaviour and anterior cingulate
907 abnormalities in autism spectrum disorders (ASD). Brain 131:2464–2478.

908 Van de Cruys S, Evers K, Van der Hallen R, Van Eylen L, Boets B, de-Wit L, Wagemans
909 J (2014) Precise minds in uncertain worlds: Predictive coding in autism.
910 Psychological Review 121:649–675.

911 Vishne G, Jacoby N, Malinovitch T, Epstein T, Frenkel O, Ahissar M (2021) Slow
912 update of internal representations impedes synchronization in autism. Nat
913 Commun 12:5439.

914 Vlamings PHJM, Jonkman LM, Hoeksma MR, Van Engeland H, Kemner C (2008)
915 Reduced error monitoring in children with autism spectrum disorder: an ERP
916 study. European Journal of Neuroscience 28:399–406.

917 Wechsler D (2008) Wechsler Adult Intelligence Scale--Fourth Edition.

918 Wu D, José JV, Nurnberger JI, Torres EB (2018) A Biomarker Characterizing
919 Neurodevelopment with applications in Autism. Sci Rep 8:614.

920 Xu N, Li X, Zhong Y (2015) Inflammatory Cytokines: Potential Biomarkers of
921 Immunologic Dysfunction in Autism Spectrum Disorders. Mediators of
922 Inflammation 2015:531518.

923 Yizhar O, Fenno LE, Prigge M, Schneider F, Davidson TJ, O'Shea DJ, Sohal VS,
924 Goshen I, Finkelstein J, Paz JT, Stehfest K, Fudim R, Ramakrishnan C,

Huguenard JR, Hegemann P, Deisseroth K (2011) Neocortical excitation/inhibition balance in information processing and social dysfunction. Nature 477:171–178.

Zuk NJ, Murphy JW, Reilly RB, Lalor EC (2021) Envelope reconstruction of speech and music highlights stronger tracking of speech at low frequencies. PLOS Computational Biology 17:e1009358.