Social Bayes: Using Bayesian Modeling to Study Autistic Trait–Related Differences in Social Cognition

Meltem Sevgi, Andreea O. Diaconescu, Lara Henco, Marc Tittgemeyer, and Leonhard Schilbach

ABSTRACT

BACKGROUND: The autistic spectrum is characterized by profound impairments of social interaction. The exact subpersonal processes, however, that underlie the observable lack of social reciprocity are still a matter of substantial controversy. Recently, it has been suggested that the autistic spectrum might be characterized by alterations of the brain's inference about the causes of socially relevant sensory signals.

METHODS: We used a novel reward-based learning task that required integration of nonsocial and social cues in conjunction with computational modeling. Thirty-six healthy subjects were selected based on their score on the Autism-Spectrum Quotient (AQ), and AQ scores were assessed for correlations with cue-related model parameters and task scores.

RESULTS: Individual differences in AQ scores were significantly correlated with participants' total task scores, with high AQ scorers performing more poorly in the task ($r = -.39$, 95% confidence interval = -0.68 to -0.13). Computational modeling of the behavioral data unmasked a learning deficit in high AQ scorers, namely, the failure to integrate social context to adapt one's belief precision—the precision afforded to prior beliefs about changing states in the world—particularly in relation to the nonsocial cue.

CONCLUSIONS: More pronounced autistic traits in a group of healthy control subjects were related to lower scores associated with misintegration of the social cue. Computational modeling further demonstrated that these trait-related performance differences are not explained by an inability to process the social stimuli and their causes, but rather by the extent to which participants consider social information to infer the nonsocial cue.

Keywords: Autistic traits, Bayesian modeling, Computational psychiatry, Reward-based learning, Social cognition, Social gaze

https://doi.org/10.1016/j.biopsych.2019.09.032

Autism is characterized by profound impairments of social interaction and communication. These difficulties are thought to be related across the diagnostic divide to autistic trait–associated differences in social perceptual and/or cognitive abilities (1). In recent years, it has become clear that a striking dissociation exists between relatively intact explicit and severely impaired implicit social abilities (2). In other words, high-functioning individuals with autism learn to explicitly think about other persons' mental states and may tend to do so even more often than neurotypical adults, yet they still find it very difficult to engage in real-time social interactions with persons without autism (3,4). Interestingly, exactly which subpersonal processes show autism trait–related differences and could explain everyday life social impairments is still a matter of substantial controversy. Recent studies have provided evidence that many putatively relevant processes, such as action perception, are intact in autism (5). Still, individuals with autism experience striking impairments in everyday life social situations, which raises the question of which processes other than basic perceptual mechanisms may come into play, and how this would occur (6).

A currently prominent theoretical suggestion includes the assumption that the autistic spectrum might be specifically characterized by deficits of predictive coding or Bayesian inference (7,8). Predictive coding formulations of perception propose that expectations in higher brain areas generate top-down predictions that meet bottom-up stimulus-related signals from lower sensory areas. The discrepancy between actual sensory input and predictions of that input is described as a prediction error (PE). The degree to which PEs revise (Bayesian) beliefs depends on the precision afforded to PEs. In other words, the precision of sensory PEs—relative to the precision of prior beliefs (and their PEs)—can have a profound effect on belief updating. In regard to autism, it has been proposed that autistic traits might be related to higher sensory precision, i.e., a stronger reliance on (bottom-up) sensory evidence compared with (top-down) prior beliefs, which can lead to a failure of automatically contextualizing sensory information in an optimal and socially adequate fashion (9,10). Furthermore, the reliance on prior beliefs might be particularly important and relevant in situations of high uncertainty, such as direct social interactions with others, as social agents are arguably the most difficult stimuli to predict (10). Interestingly, this theoretical proposition resonates with clinical descriptions of patients with autism as having a particular dislike for situations of direct social interaction with others, whereas situations of social observation (when other agents are merely observed) are described as less difficult (4).

In light of recent findings, which demonstrate relatively intact perceptual processes across a wide range of tasks in autism, it might be exactly the integration of bottom-up and top-down processes during social interactions and exploitation of social cues provided by others during decision making that could be particularly relevant to understanding the social impairments in autism. In other words, whereas autistic traits may not be associated with disturbances of basic perceptual and learning processes, it is conceivable that such traits may affect social learning processes or the extent to which social information automatically influences decision making and thereby what behavior is actually shown. From a predictive coding perspective, there are two possible pathologies. First, there could be deficits in predicting and inferring the mental states of others owing to an impoverished weight or precision of higher-level prior beliefs; or second, these inferences or representations are unable to influence behavior.

Importantly, recent progress in computational modeling has convincingly demonstrated that Bayesian models can be used to formally investigate perceptual and cognitive mechanisms that underlie social behavior when explicit social advice is provided to study participants (11): in particular, it has been shown that humans employ hierarchical generative models to make inferences about the changing intentions of others when attention is explicitly directed toward them and that they integrate estimates of advice accuracy (i.e., the correctness of the advice, which can be valid or misleading depending on the conflicting interests of the players) with nonsocial sources of information when making decisions. In Bayesian terms, this integration corresponds to an optimal weighting of prosocial and nonsocial cues in terms of their relative precision when making decisions.

In our study, we built on this research by applying hierarchical Bayesian modeling to behavioral data from a novel version of a probabilistic learning paradigm that included a social gaze cue about whose relevance no explicit information was given to investigate autistic trait–related differences in the extent to which healthy individuals integrate and use this piece of social information during task performance. In light of the evidence discussed above, we hypothesized that autistic traits are related to differences in the extent to which individuals are influenced by social cues (i.e., their precision), rather than a general inability to process social cues and putatively underlying mental states. On the behavioral level, this should result in higher total task scores for individuals with lower autistic traits, as they should be more easily able to exploit the additional social information. In terms of the underlying cognitive processes, we hypothesized that this behavioral advantage might be subserved by differences in the effect that social information can have on learning and decision making, which, in turn, would be inversely related to autistic traits. We further predicted that using the social cue should be more difficult under volatile conditions and differentially so for individuals with higher autistic traits.

METHODS AND MATERIALS

Participants

In light of evidence that suggests that autistic traits are distributed as a continuum across the general population and are known to show identical etiology across the diagnostic divide (1), we chose to study healthy participants based on their score on the German translation of the Autism-Spectrum Quotient (AQ) questionnaire (12). This experimental approach of studying autistic traits in neurotypical subjects makes it possible to infer the etiology of autistic traits without potential confounds from a variety of comorbid conditions often noted in patients with autistic spectrum disorder. To capture the extremes of the distribution and have a balanced proportion of participants with high and low AQ scores, 36 subjects were prescreened and invited based on their AQ scores up to 25 (19 men; age range, 20–37 years; mean age, 26.25 years) (Table 1). It has been shown that the AQ has a good discriminative validity at a threshold of 26 (13). Participants did not have any history of neurological or psychiatric disorders and were invited by using a preexisting database of the Max Planck Institute for Metabolic Research comprising healthy native German volunteers. The distribution of AQ scores was as follows: range = 7–23, mean = 15.72, SD = 5.09. All participants gave informed consent before the beginning of the experiment.

Experimental Paradigm

The card game used in our study, which had been originally designed as two cards with associated winning probabilities (14,15), was combined with a face cue presented in the center of the screen (Figure 1A). The eye gaze direction of the face was manipulated to change during each trial and to then be directed toward one of the cards before participants were allowed to make their choice. As a result, two things needed to be learned in the task: first, whether the reward is associated with the green card or the blue card; second, whether the gaze shift is directed toward the card that is rewarded. The probability of whether the face actually looked toward the winning

Table 1. Descriptive Data of Participants

AQ Group	Gender, M/F	Age, Years	AQ	SQ	EQ	IQ (Verbal)
High AQ (*n* = 18)	9/9	25.5 (0.7)	20.4 (0.5)	27.1 (2)	41.1 (2.1)	101.9 (2.1)
Low AQ (*n* = 18)	10/8	27 (1)	11.1 (0.4)	23.9 (2.1)	44.3 (2.6)	103.2 (2.7)

Values are reported as mean (SEM).

AQ, Autism-Spectrum Quotient; EQ, empathy quotient; F, female; M, male; SQ, systemizing quotient.

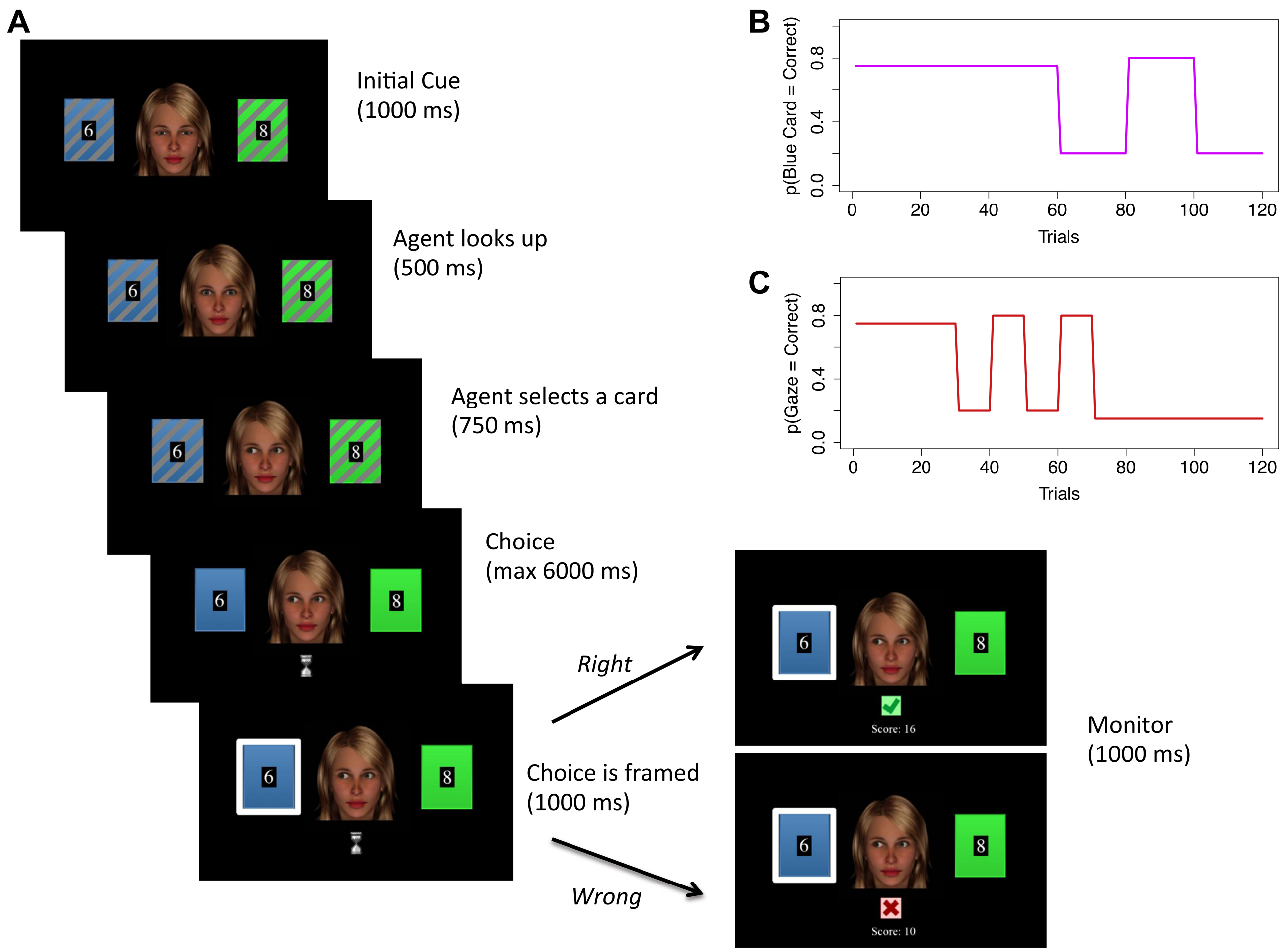

Figure 1. Experimental design. **(A)** Subjects can make a choice once the lines on both cards disappear. If the choice is right on that trial, a green tick is displayed, and the reward value of the right card is added to the total score. If the choice is wrong, a red cross is displayed, and the score remains the same. Probability schedules. **(B)** Probability of the blue card being correct (i.e., card accuracy), and **(C)** probability of the gaze showing the correct card (i.e., gaze accuracy).

card on a given trial (i.e., gaze accuracy) was systematically manipulated in accordance with two probabilistic schedules (see Supplement for more details about the schedules).

Across both conditions, the card and gaze accuracies were varied independently of one another across the experiment (Figure 1B, C). The phases in which the trials have cues with unstable accuracy are referred as volatile phases. In the first half of the experiment (trials 1–60), card accuracy was stable and high, whereas in the second half (trials 60–120), it followed a volatile phase (Figure 1B). For the gaze accuracy, the volatile phase took place during trials 30 to 70. Positions of the cards (left or right) were determined randomly.

Importantly, we manipulated the degree to which the gaze could influence learning about the card probabilities. Although the gaze schedules were matched in terms of overall congruency to the card probabilities, the order of those phases was manipulated, resulting in two different conditions: the congruency first condition began with the gaze as highly congruent to the winning card (with 80% probability of being informative of the winning card), and the incongruency first condition began with the gaze strongly incongruent to the winning card (with only 20% probability of being informative of the winning card) (Figure 1C).

In the instructions, subjects were informed that the cards have winning probabilities, which can change during the experiment and which are independent of the reward magnitude that is displayed on them. On each trial, there was only one correct card, and if subjects chose the correct card, they would receive the score (random numbers between 1 and 9) that had been displayed on it. They were instructed that they would earn an extra amount of money depending on their score at the end of the experiment. Finally, participants were informed about the presence of a face on the screen, which was explained by stating that it was supposed to make the visual display more interesting. Participants did not receive any other information about the face in an attempt to keep the instruction about the gaze cue as implicit as possible. After the experiment, subjects filled out a brief questionnaire.

Perceptual and Response Models

The "observing the observer" approach provides a complete mapping from experimental stimuli to observed

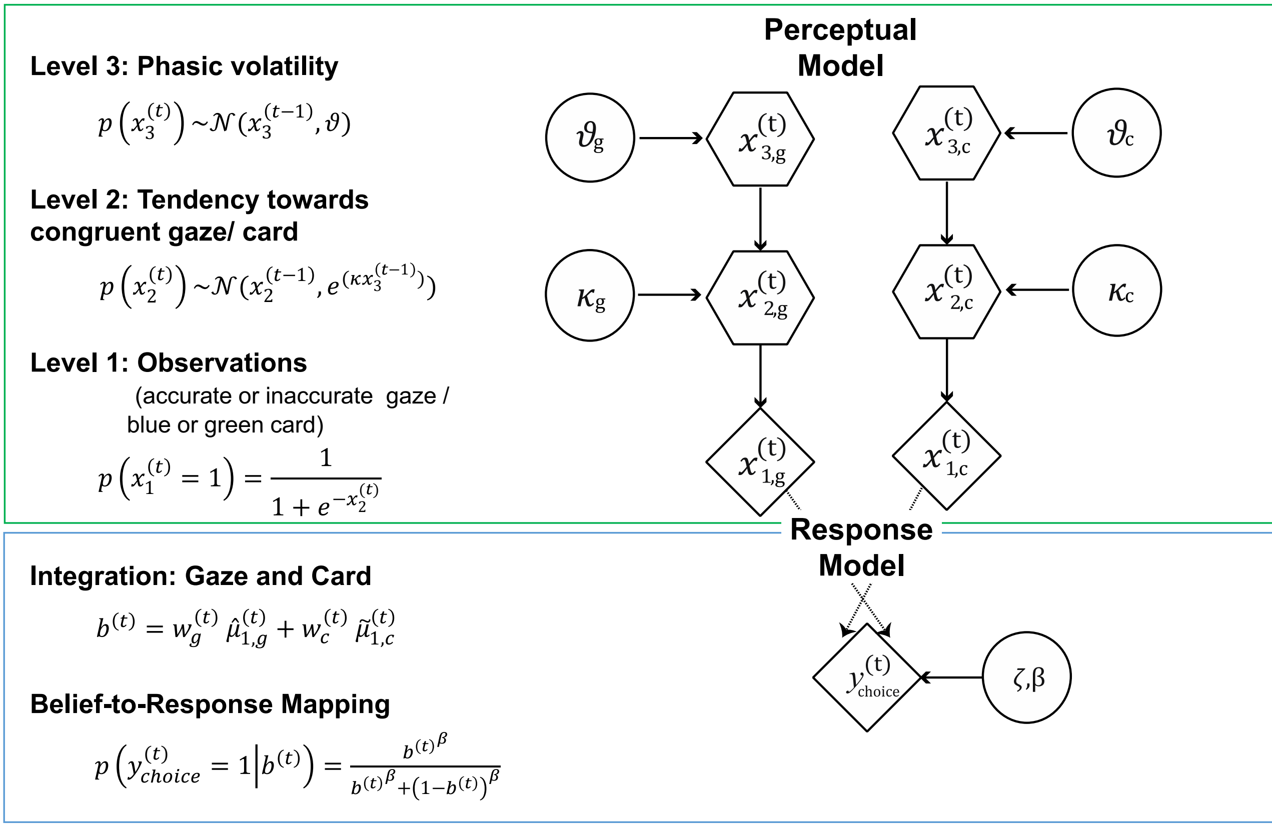

Figure 2. Graphical depiction of two parallel learning systems that were assumed to influence the choice behavior. For any trial t, $x_3^{(t)}$ follows a Gaussian random walk such that $p(x_3^{(t)}) \sim \mathcal{N}(x_3^{(t-1)}, \vartheta)$. The first level state variable $x_1^{(t)}$ is the accuracy at that trial and is a sigmoid transform of the second level state variable $x_2^{(t)}$, which also follows a Gaussian distribution: $\mathcal{N}(x_2^{(t-1)}, \exp(\kappa x_3^{(t-1)}))$, where the variance term depends on κ, which accounts for the coupling between the third level and the second level and the phasic volatility from the previous time step, or $x_2^{(t-1)}$. The response model maps the predicted outcome probabilities to choices via a softmax function.

responses by inversion of the perceptual model and the response model (16). An application of this approach is a generative model called the Hierarchical Gaussian Filter (HGF), which accounts for deterministic and probabilistic relationships between the environment and perceptual states (17). We used a perceptual-response model pair to infer subjects' beliefs about the stimuli. We modeled congruency of response with advice, i.e., the advice given by the social cue (the gaze), using the HGF combined with a response model as implemented in Diaconescu *et al.* (11). This approach allows the estimation of hierarchically coupled hidden states that describe subjects' learning about the environmental statistics, namely, the probability and the volatility of the card and gaze cues, based on their responses. These subjective beliefs are weighted by their precision to form the basis of a response model (of the observed behavior) as explained in detail below. The graphical representation of the perceptual model is shown in Figure 2.

The dynamics of the belief trajectories—the accuracy and volatility estimates as well as their precisions—are determined by 4 learning parameters: κ_g, κ_c, ϑ_g, ϑ_c for gaze and card outcomes, respectively. Across stimulus modalities, parameter κ reflects the coupling between the levels of the model, thus determining the phasic component of the learning rate, and parameter ϑ refers to the meta-volatility, thereby regulating the variance of the cue-outcome volatility. A detailed description of the perceptual models used here is in the Supplement.

Belief Precision. In the HGF, the belief update is proportional to a precision-weighted PE (see Supplement). The belief precision weighting the PE depends on the estimated environmental volatility and the low-level (sensory) precision:

$$\pi_{2,g}^{(t)} = \widehat{\pi}_{2,g}^{(t)} + \widehat{\mu}_{1,g}^{(k)}\left(1 - \widehat{\mu}_{1,g}^{(k)}\right), \quad \pi_{2,c}^{(t)} = \widehat{\pi}_{2,c}^{(t)} + \widehat{\mu}_{1,c}^{(k)}\left(1 - \widehat{\mu}_{1,c}^{(k)}\right) \tag{1}$$

with the precision of the prediction given by:

$$\widehat{\pi}_{2,g}^{(t)} = \frac{1}{1\Big/\pi_{2,g}^{(t-1)} + \exp\left(\kappa_g \mu_{3,g}^{(t-1)}\right)}, \quad \widehat{\pi}_{2,c}^{(t)} = \frac{1}{1\Big/\pi_{2,c}^{(t-1)} + \exp\left(\kappa_c \mu_{3,c}^{(t-1)}\right)} \tag{2}$$

where $\mu_3^{(t-1)}$ is the predicted environmental volatility.

Precision Weighted Response Model. We applied this model to derive subject-specific accuracy and volatility estimates for card and gaze in a parallel manner. On a given trial t, subjects generated a combined belief, $b^{(t)}$, after weighting the posterior expectation of inferred card and gaze accuracies, $\tilde{\mu}_{1,c}^{(t)}$ and $\widehat{\mu}_{1,g}^{(t)}$, to generate actions in the following manner:

$$w_g^{(t)} = \frac{\zeta \widehat{\pi}_{1,g}^{(t)}}{\zeta \widehat{\pi}_{1,g}^{(t)} + \widehat{\pi}_{1,c}^{(t)}},\ w_c^{(t)} = \frac{\widehat{\pi}_{1,c}^{(t)}}{\zeta \widehat{\pi}_{1,g}^{(t)} + \widehat{\pi}_{1,c}^{(t)}} \tag{3}$$

$$b^{(t)} = w_g^{(t)}\ \widehat{\mu}_{1,g}^{(t)} + w_c^{(t)}\ \tilde{\mu}_{1,c}^{(t)} \tag{4}$$

Where $w_g^{(t)}$ and $w_c^{(t)}$ are effective precision ratios of gaze and card cues, $\tilde{\mu}_{1,c}^{(t)}$ is the transformed expected card color probability from the perspective of the gaze (i.e., the estimated card color probability indicated by the gaze), and $\widehat{\mu}_{1,g}^{(t)}$ corresponds to the logistic sigmoid of the current expectation of advisor fidelity:

$$\widehat{\mu}_{1,g}^{(t)} = s\left(\mu_{2,g}^{(t-1)}\right) = \frac{1}{1 + \exp\left(-\mu_{2,g}^{(t-1)}\right)} \tag{5}$$

Response model parameter ζ is the weight on the precision of inferred gaze accuracy or the additional bias toward the social cue; $\widehat{\pi}_{1,g}^{(t)}$ and $\widehat{\pi}_{1,c}^{(t)}$ are precisions (inverse variances) at the first level for gaze and card accuracies, respectively. As the first level estimates are assumed to follow a Bernoulli distribution, one can calculate the precision at each trial by:

$$\widehat{\pi}_{1,g}^{(t)} = \frac{1}{\widehat{\mu}_{1,g}^{(t)}\left(1 - \widehat{\mu}_{1,g}^{(t)}\right)},\ \widehat{\pi}_{1,c}^{(t)} = \frac{1}{\tilde{\mu}_{1,c}^{(t)}\left(1 - \tilde{\mu}_{1,c}^{(t)}\right)} \tag{6}$$

The probability of taking the gaze advice was assumed to be a softmax function:

$$p\left(y^{(t)} = 1 | b^{(t)}\right) = \frac{{b^{(t)}}^{\beta}}{{b^{(t)}}^{\beta} + \left(1 - b^{(t)}\right)^{\beta}} \tag{7}$$

Where $\beta > 0$ is the subject-specific inverse decision temperature parameter.

Hypotheses. Given that we assumed the autistic spectrum to be characterized by differences in sensory precision compared with belief (i.e., prior) precision, we expected differences in two sets of parameters: 1) perceptual model parameter κ, which modulates the influence of sensory compared with belief precision in the belief updating process, and 2) response model parameter ζ, which signals the bias toward the social cue. To this hypothesis, we extracted the parameters of the winning model and subjected them to a two-way analysis of variance with an interaction term (group, condition [congruent first vs. incongruent first gaze schedule], and group × condition). The participant groups were defined by their AQ scores, which were obtained using a median split procedure (median AQ = 15). The models and the routines for all analyses performed here are available as MATLAB (The MathWorks, Inc., Natick, MA) code https://gitlab.ethz.ch/dandreea/mltm. See Table 2 for the prior mean and variance over the parameters.

Other Behavioral Measures. We assessed the relationship between AQ scores and total task scores, as the ability to exploit the additional social information should contribute to task performance. We predicted that the volatility of the input structure may influence subjects' inference about the gaze and subsequent decision to take gaze into account. The influence of probability (high vs. low) and volatility (stable vs. volatile) of the gaze cue on performance was evaluated and compared between two AQ groups.

RESULTS

To investigate group differences in learning about the card probabilities and gaze congruency, we examined group × condition interactions in the parameter maximum a posteriori estimates of the winning model (i.e., M_5, the full HGF with the integrated decision model).

As predicted, subjects with low AQ scores used the gaze schedule to successfully learn about the card probabilities. Whereas they showed superior performance in the task (see Supplement for details), they also displayed significant differences in learning about the card probabilities as a function of the gaze schedule. A significant group × condition interaction was observed for parameter κ_c (group: $F_{1,35}$ = 0.28, $p > .60$; condition: $F_{1,35}$ = 1.02, $p > .30$; interaction: $F_{1,35}$ = 10.74, p = .0025), suggesting that whereas learning of subjects with high AQ scores about the card probabilities was unaffected by the gaze congruency schedule, subjects with low AQ scores adapted their learning rate according to the gaze schedule (Figure 3A). This interaction effect was supported by post hoc t tests, which revealed that participants with low AQ scores showed larger κ_c values in the incongruent compared with the congruent first condition (2-tailed t test; t_{16} = −2.55, p = .04 corrected for multiple comparisons using Bonferroni correction), whereas participants with low AQ scores showed a lack of differences between the 2 conditions (2-tailed t test; t_{16} = 2.08, p = .10 after Bonferroni correction). This result was also related to participants' postexperiment descriptions of their performance. In contrast to participants with high AQ score who showed a lack of task differences, participants with low scores reported having to rely on the gaze information more when the card and gaze input structures started out as incongruent compared with the congruent first condition (group: $F_{1,35}$ = 0.80, $p > .38$; condition: $F_{1,35}$ = 0.80, $p > .37$; interaction: $F_{1,35}$ = 5.89, p = .021) (Supplemental Figure S5A).

Mirroring this effect, we also observed significant group × condition interactions for the precision of the prediction about the card probabilities or $\widehat{\pi}_{2,c}^{(t)}$ (see Equation 5). Subjects with low AQ scores were more confident in their predictions about the card probabilities when the gaze schedule began as congruent than when it began as incongruent to the winning cards compared with the subjects with high AQ scores (group: $F_{1,35}$ = 0.26, $p > .60$; condition: $F_{1,35}$ = 0.97, $p > .30$;

Table 2. Prior Mean and Variance of Perceptual and Response Model Parameters

Model	Parameter	Prior Mean	Prior Variance
Perceptual Models			
Normative HGF	κ_g, κ_c	0.50	0
	ϑ_g, ϑ_c	0.56	0
3-level HGF	κ_g, κ_c	0.50	1
	ϑ_g, ϑ_c	0.56	1
Response Models	β	48	25
Integrated	ζ	1	25
Card Only	ζ	0	0

The prior variances are given in the numeric space in which parameters are estimated. κ and ϑ are estimated in logit space, and the other parameters are estimated in log space. Whereas the prior variances for all parameters are set to be rather broad, we selected a shrinkage prior mean for the decision noise such that behavior is explained more by variations in the rest of the parameters rather than decision noise.

HGF, Hierarchical Gaussian Filter.

interaction: $F_{1,35}$ = 10.82, p = .02). Subjects with high AQ scores showed no differences in belief precision across the two gaze schedules (Figure 3B).

Association Between AQ Scores and Utility of Misleading Advice in a Volatile Environment

Figure 4A illustrates the performance in each phase of the gaze accuracy. As expected, we observed a significant relationship between AQ scores and choosing in accordance with the gaze during volatile low probability phases (R^2 = 22.28%, F_{35} = 9.74, p = .0037), with AQ scores correlating with the number of trials where the subjects took the gaze into account (r = .52, 95% confidence interval = 0.29 to 0.75) (Figure 4C). As there were no AQ group differences when the gaze cue was stable, the results suggest that subjects with high AQ scores do spontaneously take the social cue into account, even when its meaning (i.e., gaze direction) is reversed, thereby implying that lower level aspects of social processing are intact. The group difference emerges when the gaze cue is volatile. In this context, the gaze cue is imprecise (i.e., less reliable) and should be ignored, but subjects with high AQ scores continue to rely on it, leading to poor performance in particular when the gaze cue accuracy is low. In terms of the computational model, this implies that the subjects with high AQ scores do not take into account the precision of the gaze cue when using it to infer the card probabilities.

DISCUSSION

In this study we applied hierarchical Bayesian modeling to investigate autistic trait–related differences in the extent to which healthy individuals integrate and make use of gaze cues in a probabilistic reward learning task. For optimal performance, our task required following both the card and the gaze cues and combining these two sources of information, even though instructions provided very little information about the nature and relevance of the gaze cues, in contrast to other studies using explicit forms of social advice (11,14). As expected, our results demonstrate an inverse relationship between autistic traits (as measured by AQ scores) and total task scores obtained by study participants, such that individuals with higher autistic traits obtained lower total task scores.

We were particularly interested to model perceptual as well as higher-order processing of both card and gaze cues and, in particular, their relationship to action selection, i.e., the extent

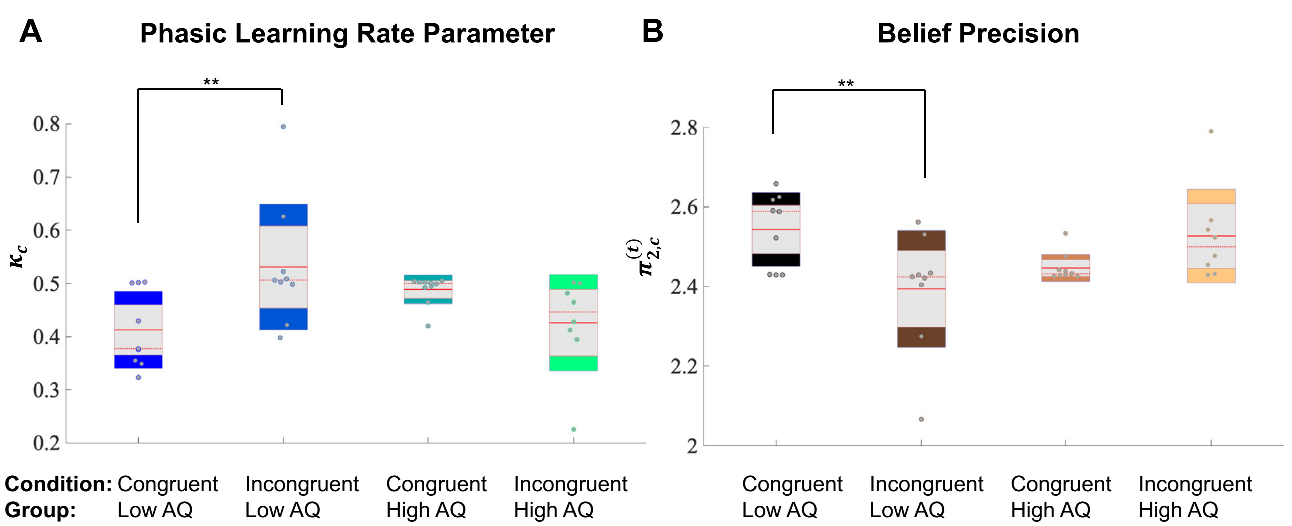

Figure 3. Maximum a posteriori estimates for both groups and conditions. **(A)** κ_c, the parameter representing the coupling between the second level and the third level for the card model, showed a significant group × condition interaction. **(B)** $\pi^{(t)}_{2,c}$, the average belief precision about the card probabilities, also showed a significant group × frame interaction. The interaction suggests significant differences between the two conditions in the low Autism-Spectrum Quotient (AQ) group, which were absent in the high AQ group. See main text for details. Jittered raw data are plotted for each variable. The red line refers to the mean, the interrupted red line refers to the median, the colored background reflects the 95% confidence intervals for the mean, and the gray background refers to 1 SD of the mean. **Significant post hoc t tests after Bonferroni correction.

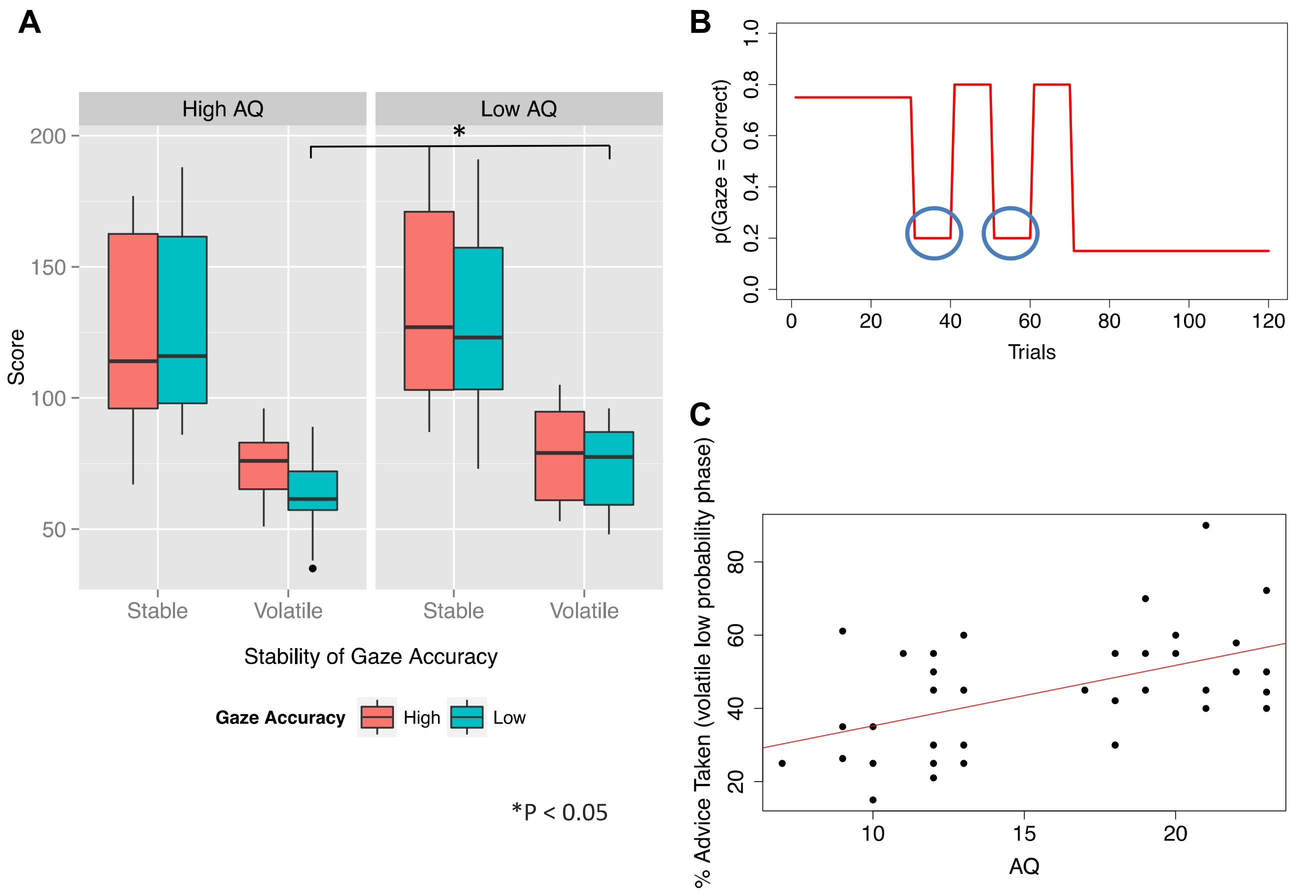

Figure 4. Influence of structure of the environment on the behavior. **(A)** Scores obtained by high and low Autism-Spectrum Quotient (AQ) groups in different phases of the experiment based on the features of the gaze cue (high × low gaze accuracy and stable × volatile periods of gaze accuracy). **(B)** Difference is significant (*p = .034) in the volatile low accuracy phase (circled area). **(C)** During the same phase, the number of trials in which the subjects took the advice, i.e., chose the card that is indicated by the highly misleading gaze, was correlated with AQ traits (r = .52, 95% confidence interval = 0.29 to 0.75).

to which individuals were actually biased by the social information provided on a trial-by-trial basis. Results of our computational analyses provide evidence for AQ-related group differences, such that individuals with lower AQ scores are influenced by the gaze cue more as indicated by an enhanced learning rate modulation as a function of the gaze schedule (Figure 3A). This unmasks a learning deficit in subjects with high AQ scores, namely, the failure to take into account the social context to adapt the learning about the nonsocial stimulus to accurately predict the outcome of the binary lottery. Furthermore, our results indicate that individuals with high AQ scores had particular difficulties integrating the social cue, as they were more likely to rely on the gaze during volatile trials and when the gaze accuracy was low (Supplemental Figure S5). These results jointly suggest that autistic traits are associated with a reduced impact of social information on the precision of higher-level prior beliefs, which has a detrimental effect on probabilistic learning about card outcomes.

By providing these insights into AQ-related differences in social cognition, our study, we believe, is most relevant to current discussions concerned with mechanistic explanations of autistic symptoms: predictive coding theories have reconstructed autism in terms of high-level attenuated precisions relative to sensory precision (9), which results in an enhanced weighting of PEs (10) and a loss of the selective force when processing a context with multiple cues (18). We find reduced modulation of top-down belief precision as a function of social context (Figure 3B), which is another mechanism leading to an enhanced weighting of PEs (9). As stated by Pellicano and Burr (7), Bayesian models provide an important avenue that can help to identify whether autistic trait–associated alterations lie in the reliance on prior knowledge or the optimal update of prior information during learning. In our Bayesian formulation, we addressed this issue by assessing possible relationships of perceptual and response model parameters with AQ scores. We found a relationship between perceptual model and response model parameters with AQ scores, although the latter relationship was less significant. Participants who scored higher on the AQ questionnaire did not take advantage of the gaze schedule to adjust their belief precision about the card outcomes. This finding relates to recent applications of hierarchical Bayesian modeling in the context of autism spectrum disorder (19). Individuals with autism spectrum disorder showed reduced learning rates and thus belief precision about probabilistic (nonsocial) outcomes, while at the same time exhibiting an enhanced learning about volatility. Whereas we observed a reduced modulation of the learning rate by the social cue, we also found a lack of differences in learning about volatility, suggesting the presence of similar but dampened learning deficits in individuals with subclinical autistic traits. These findings appear consistent with a recent suggestion by Palmer *et al.* (20), who proposed that autism may not impair the ability to process social information per se, but rather lead to differences in how the relevant representations are integrated for optimal action selection.

In light of other propositions, which hold that autistic trait–related impairments of social cognition may be particularly relevant in complex and unpredictable situations (18), we further investigated whether subjects' AQ scores were also related to task performance during phases of the experiment, which included volatile and misleading gaze cues. Here our data show that this particularly unstable environment made it more difficult for subjects with higher AQ scores to use the social cue while making decisions. This kind of influence of volatility on behavior parallels results from previous studies, which report that an unpredictable context makes it more difficult for individuals with autism to use social cues in an appropriate way (21). Our finding can therefore be seen as evidence for difficulties of contextualizing social cues in light of high uncertainty.

The modeling approach that we implemented in this study is a promising method for capturing individual differences in the learning and integration of social information. Given the heterogeneity of the population, this could be particularly useful for identifying subgroups that may map onto distinct mechanisms of impaired social interaction in autism. The "observing the observer" approach has indeed been demonstrated to be useful for inference about hidden states and parameters that shape interindividual differences in learning (22). Our results indicate that Bayesian models may be particularly powerful in providing mechanistic explanations of social difficulties, which are particularly relevant to an understanding of psychiatric disorders (4,23,24). Advances in computational psychiatry (25–28) and studies such as this could therefore contribute to mechanistic formulations of psychopathology.

It is important to note that we cannot rule out intact precision-weighted PE processing in patients with autism, as our sample comprised healthy subjects. One can speculate that in a patient sample, impaired inference about the social cue in addition to the reduced integration of social information could be observed. Therefore, future research should include testing patients with a formal diagnosis of autism to explore whether the observed differences hold across the entire autistic spectrum. Furthermore, the experimental paradigm introduced here and our analysis approach could be used together with neuroimaging to investigate which activity and connectivity profiles in brain regions relevant for social cognition underlie the observed autistic trait–related behavioral differences.

Taken together, the results of our study demonstrate autistic trait–related behavioral differences in a task that requires the integration of nonsocial and social information. Using hierarchical Bayesian modeling, we show that these performance differences are subserved by impairments of integrating social information to infer causal structures in the environment, which is consistent with previous findings in autism (5).

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the German Research Foundation in Clinical Research Group 219 (to MT) and as part of Transregional Collaborative Research Center 134 (to LS and MT), the Max Planck Society (to LS and MS), the Swiss National Foundation (Grant No. PZ00P3_167952 [to AOD]), and the Krembil Foundation (to AOD).

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Max Planck Institute for Metabolism Research (MS, MT), Cologne, Germany; Translational Neuromodeling Unit (AOD), University of Zürich and Eidgenössische Technische Hochschule Zürich, Zürich, and Department of Psychiatry (AOD), Universitäre Psychiatrische Kliniken, University of Basel, Basel, Switzerland; Krembil Centre for Neuroinformatics (AOD), Centre for Addiction and Mental Health, University of Toronto, Toronto, Ontario, Canada; Max Planck Institute of Psychiatry (LH, LS); Graduate School for Systemic Neurosciences (LH), Munich; and Department of Psychiatry (LS), University Hospital Cologne, Cologne, Germany.

MS and AOD contributed equally to this work.

Address correspondence to Andreea Oliviana Diaconescu, Ph.D., Krembil Centre for Neuroinformatics, Centre for Addiction and Mental Health, 250 College Street, Toronto, Ontario, Canada; E-mail: andreea.diaconescu@camh.ca.

Received Jul 21, 2019; revised Sep 26, 2019; accepted Sep 30, 2019.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.biopsych.2019.09.032.

REFERENCES

1. Robinson EB, Koenen KC, McCormick MC, Munir K, Hallett V, Happe F, *et al.* (2011): Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). Arch Gen Psychiatry 68:1113–1121.
2. Senju A, Southgate V, White S, Frith U (2009): Mindblind eyes: An absence of spontaneous theory of mind in Asperger syndrome. Science 325:883–885.
3. Schilbach L, Eickhoff SB, Cieslik EC, Kuzmanovic B, Vogeley K (2012): Shall we do this together? Social gaze influences action control in a comparison group, but not in individuals with high-functioning autism. Autism 16:151–162.
4. Schilbach L, Timmermans B, Reddy V, Costall A, Bente G, Schlicht T, *et al.* (2013): Toward a second-person neuroscience. Behav Brain Sci 36:393–414.
5. Cusack JP, Williams JH, Neri P (2015): Action perception is intact in autism spectrum disorder. J Neurosci 35:1849–1857.
6. Hamilton AF (2013): Reflecting on the mirror neuron system in autism: A systematic review of current theories. Dev Cogn Neurosci 3:91–105.
7. Pellicano E, Burr D (2012): When the world becomes 'too real': A Bayesian explanation of autistic perception. Trends Cogn Sci 16:504–510.
8. Sinha P, Kjelgaard MM, Gandhi TK, Tsourides K, Cardinaux AL, Pantazis D, *et al.* (2014): Autism as a disorder of prediction. Proc Natl Acad Sci U S A 111:15220–15225.
9. Friston KJ, Lawson R, Frith CD (2013): On hyperpriors and hypopriors: Comment on Pellicano and Burr. Trends Cogn Sci 17:1.
10. Lawson RP, Rees G, Friston KJ (2014): An aberrant precision account of autism. Front Hum Neurosci 8:302.
11. Diaconescu AO, Mathys C, Weber LA, Daunizeau J, Kasper L, Lomakina EI, *et al.* (2014): Inferring on the intentions of others by hierarchical Bayesian learning. PLoS Comput Biol 10:e1003810.

12. 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.
13. Woodbury-Smith MR, Robinson J, Wheelwright S, Baron-Cohen S (2005): Screening adults for Asperger syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. J Autism Dev Disord 35:331–335.
14. Behrens TE, Hunt LT, Woolrich MW, Rushworth MF (2008): Associative learning of social value. Nature 456:245–249.
15. Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007): Learning the value of information in an uncertain world. Nat Neurosci 10:1214–1221.
16. Daunizeau J, den Ouden HE, Pessiglione M, Kiebel SJ, Stephan KE, Friston KJ (2010): Observing the observer (I): Meta-bayesian models of learning and decision-making. PLoS One 5:e15554.
17. Mathys C, Daunizeau J, Friston KJ, Stephan KE (2011): A bayesian foundation for individual learning under uncertainty. Front Hum Neurosci 5:39.
18. Van de Cruys S, Evers K, Van der Hallen R, Van Eylen L, Boets B, de-Wit L, *et al.* (2014): Precise minds in uncertain worlds: Predictive coding in autism. Psychol Rev 121:649–675.
19. Lawson RP, Mathys C, Rees G (2017): Adults with autism overestimate the volatility of the sensory environment. Nat Neurosci 20:1293–1299.
20. Palmer CJ, Seth AK, Hohwy J (2015): The felt presence o other minds: Predictive processing, counterfactual predictions, and mentalising in autism. Conscious Cogn 36:376–389.
21. Robic S, Sonie S, Fonlupt P, Henaff MA, Touil N, Coricelli G, *et al.* (2015): Decision-making in a changing world: A study in autism spectrum disorders. J Autism Dev Disord 45:1603–1613.
22. Daunizeau J, den Ouden HE, Pessiglione M, Kiebel SJ, Friston KJ, Stephan KE (2010): Observing the observer (II): Deciding when to decide. PLoS One 5:e15555.
23. Schilbach L (2014): On the relationship of online and offline social cognition. Front Hum Neurosci 8:278.
24. Schilbach L (2015): Eye to eye, face to face and brain to brain: Novel approaches to study the behavioral dynamics and neural mechanisms of social interactions. Curr Opin Behav Sci 3:130–135.
25. Montague PR, Dolan RJ, Friston KJ, Dayan P (2012): Computational psychiatry. Trends Cogn Sci 16:72–80.
26. Friston KJ, Stephan KE, Montague R, Dolan RJ (2014): Computational psychiatry: The brain as a phantastic organ. Lancet Psychiatry 1:148–158.
27. Stephan KE, Mathys C (2014): Computational approaches to psychiatry. Curr Opin Neurobiol 25:85–92.
28. Wiecki TV, Poland J, Frank MJ (2015): Model-based cognitive neuroscience approaches to computational psychiatry: Clustering and classification. Clin Psychol Sci 3:378–399.