Autistic traits lower the flexibility of information sampling: insights from a two-stage decision model

Haoyang Lu[1,2] Li Yi[3]* Hang Zhang[2,3,4]*

1. Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China
2. Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China
3. School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China
4. PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China

*Corresponding authors:
Li Yi, email: yilipku@pku.edu.cn
Hang Zhang, email: hang.zhang@pku.edu.cn

Keywords: autistic traits, information sampling, probabilistic inference, Bayesian theory, computational psychiatry

Short title: Autistic traits and information sampling

Counts:

- Abstract:	277 words
- Main text:	3414 words
- Methods:	1653 words
- Figures	6
- References	56
- Supplementary figures	3
- Supplementary tables	1

Abstract

Information sampling can reduce uncertainty in future decisions but is often costly. To maximize expected gain, people need to balance sampling cost and information gain. Suboptimal information sampling is observed in many core symptoms of autism spectrum disorder (ASD) and may be a general impairment in ASD. Here we aimed to understand how autistic traits influence the optimality of information sampling and to identify the particularly affected cognitive processes. Healthy human adults of varying autistic traits performed a probabilistic inference task, where they drew samples sequentially to accumulate evidence for their inference and could stop sampling at any time. Each additional sample would increase the probability of correct inference but also incur a fixed cost. We manipulated the cost and evidence associated with each sample and compared participants' performance to optimal choices. We found that, compared to their peers, participants with high autistic traits had a lower efficiency of winning rewards in the high cost and low evidence condition due to oversampling. Computational modeling of participants' sampling choices revealed two decision stages: When the first stage does not reach the choice of stopping sampling, a second stage is probabilistically involved, which provides a second chance to stop sampling. We found that autistic traits did not influence the choices in either stage but specifically modulated the participant's probability to recruit the second decision stage ("second-thought probability"). In particular, participants with high autistic traits were less flexible in adjusting their second-thought probabilities with the increase of cost. To conclude, high-levels of autistic traits are associated with suboptimal information sampling and such

suboptimality implies a failure of adjusting the strategic complexity (with or without second thought) appropriately with sampling cost.

Author Summary

Children with autism can spend hours practicing lining up toys or learning all about cars or lighthouses. This kind of behaviors, we think, may reflect suboptimal information sampling strategies, that is, a failure to balance the gain of information with the cost (time, energy, or money) of information sampling. We hypothesized that suboptimal information sampling is a general characteristic of people with autism or high level of autistic traits. In our experiment, we tested how participants may adjust their sampling strategies with the change of sampling cost and information gain in the environment. Though all participants were healthy young adults who had similar IQs, those with high level of autistic traits had a lower efficiency of winning rewards due to oversampling under high sampling cost. Further computational modeling revealed a two-stage decision process, with the second stage being optional and serving as a second chance to stop sampling. Autistic traits did not influence decisions in either stage but lowered the strategic flexibility of using the second stage to reduce costly sampling. In other words, people should increase their “second thoughts” with the increase of sampling cost but those with high autistic traits fail to do it well.

Introduction

Information helps to reduce uncertainty in decision making but is often costly to collect. For example, to confirm whether a specific tumor is benign or malignant may require highly invasive surgery procedures. In such cases, it can be more beneficial to tolerate some degree of uncertainty and take actions first. To maximize survival, humans and animals need to balance the cost and benefit of information sampling and sample the environment optimally [1,2].

However, autism spectrum disorder (ASD)—a neurodevelopmental disorder characterized by social impairments and repetitive behaviors [3]—seem to be accompanied by suboptimal information sampling, though in various guises. For example, individuals with repetitive behaviors tend to spend time on redundant information that helps little to reduce uncertainty [4]. Eye-tracking studies reveal that people with ASD have atypical gaze patterns in ambiguous or social scenes, that is, they sample the visual environment in an inefficient way [5,6]. According to the recently developed Bayesian theories of ASD that explain a variety of perceptual, motor, and cognitive symptoms [7–13], deviation from Bayesian optimality in information processing is primary to ASD [4,14–17]. In this Bayesian framework, information sampling is referred as “disambiguatory active inference” [4] and plays an important role in guiding the subsequent inferences or decisions. We hereby conjectured that ASD symptoms such as repetitive behaviors and ineffecient gaze patterns reflect general impairments in information sampling.

The present study is aimed to understand how autistic traits in typical people may influence their optimality of information sampling. Autistic traits of the whole population form a continuum, with ASD diagnosis situated on the high end [18–24]. Moreover, autistic traits share genetic and biological etiology with ASD [25]. Thus, quantifying autistic-trait-related differences in healthy people can provide unique perspectives as well as a useful surrogate for understanding the symptoms of ASD [23,26]. In particular, we tested healthy adults of different levels of autistic traits in an information sampling task adapted from Jänsch and Hare [27]: On each trial of the experiment, participants could draw samples sequentially to accumulate evidence for a probabilistic inference and would receive monetary rewards for correct inferences. Each additional sample may increase their probability of correct inference but also incur a fixed monetary cost. We manipulated the cost and evidence per sample and compared participants' performance to the sampling strategy that maximizes expected gain. As expected, high-levels of autistic traits were accompanied by more deviation from optimality. Compared to their peers, participants with high-level autistic traits received lower rewards in the high-cost and low-evidence condition due to oversampling, where the optimal strategy would sacrifice accuracy to save cost.

What cognitive processes in information sampling are particularly affected by autistic traits? Through computational modeling, we further decomposed participants' sampling choices into multiple sub-processes and found that the influence of autistic traits was surprisingly selective and subtle. In particular, participants' sampling choices resulted from a two-stage decision process: When the first decision stage does not reach the choice of stopping sampling, a second decision stage is probabilistically

involved to arbitrate, which offers a second chance to consider stopping sampling. The first and second stages were respectively controlled by cost and evidence and neither stage showed autistic-trait-related differences. What varied with levels of autistic traits was the probability of recruiting the second stage: When sampling cost went from low to high, participants with high autistic traits were less flexible in adjusting their probability of re-considering stopping sampling so that they oversampled.

Results

The computerized experimental task is illustrated in Fig 1a. On each trial, participants first saw two jars filled with opposite ratios of pink and blue beads and were told that one jar had been secretly selected by the experimenter. They could sample up to 20 beads sequentially with replacement from the selected jar to infer which jar had been selected. Each key press would randomly sample one bead and participants could decide to stop sampling at any moment. For each correct inference, participants would receive 10 points minus the total sampling cost. Their goal was to earn as many points as possible, which would be redeemed into monetary bonus in the end. The cost of sampling one bead could be 0, 0.1, or 0.4 points, referred below as zero-, low-, and high-cost conditions respectively. The pink-to-blue ratios of the two jars could be 60%:40% vs. 40%:60%, or 80%:20% vs. 20%:80%, which corresponded to lower (60/40) or higher (80/20) evidence per sample favoring one jar against another. The sample size that maximizes expected gain would change with the cost and evidence conditions (Fig 1b, see Methods).

Forty-five healthy young adults participated in our experiment, whose autistic traits were measured by the self-reported Autism Spectrum Quotient (AQ) questionnaire [18]. We divided all participants evenly into three groups of low, middle, and high AQ scores, with the level of autistic traits of the high-AQ group well above the average of the general population and overlapping with the ASD population [28,29]. This tripartite division enabled us to identify possible non-linear as well as linear effects of autistic traits.

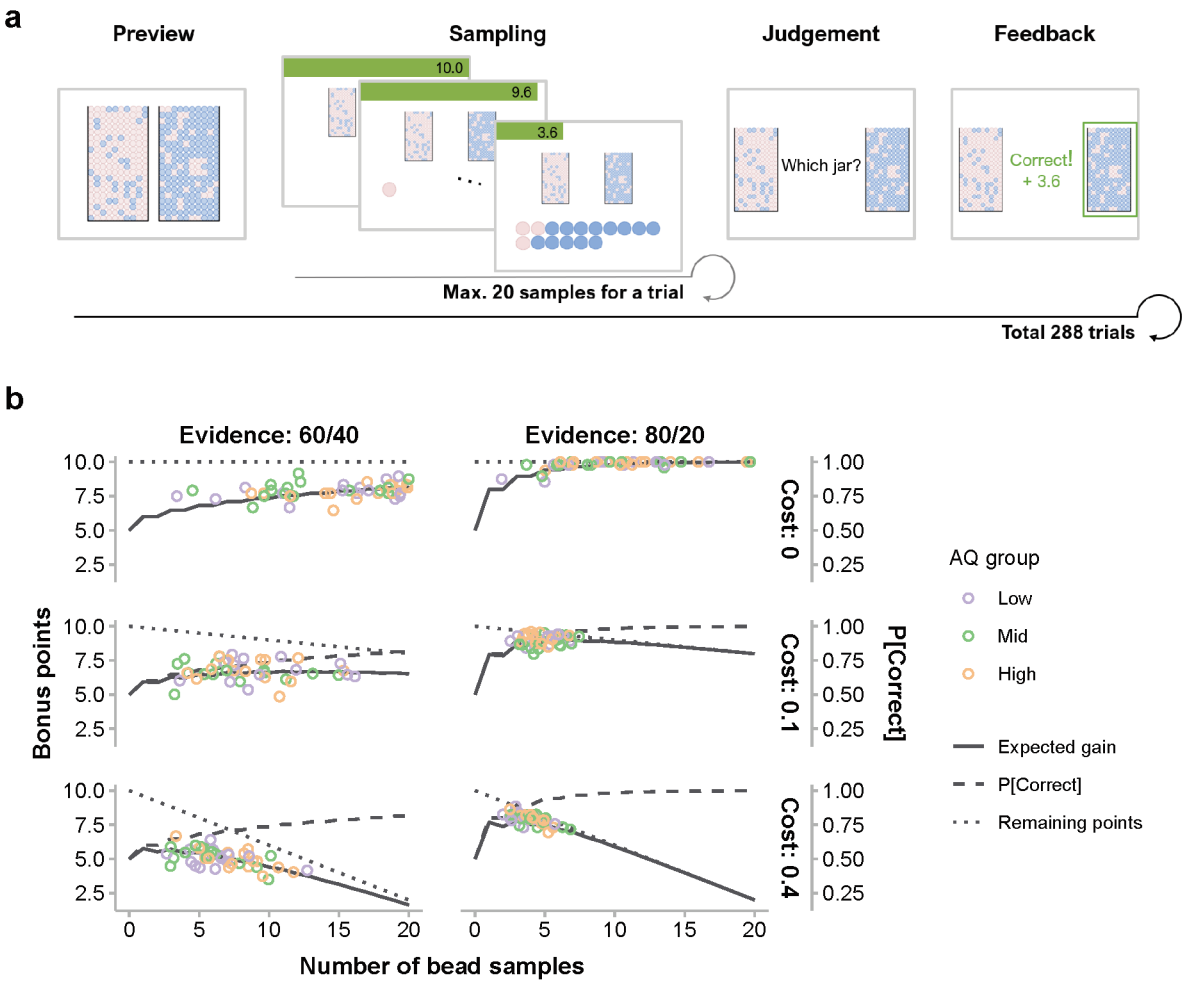

Fig 1. **The bead-sampling task**.

(a) Time course of one trial. "Preview" informed the participant of the pink-to-blue ratios of the two jars (80%:20% vs. 20%:80% in this example, corresponding to the 80/20 evidence condition). Then the participant could press Space to sample beads from the unknown pre-selected jar one at a time up to 20 beads ("sampling") or press Enter to quit sampling at any time. Afterwards the participant judged which jar had been selected ("judgment"). Feedback followed, showing the correctness of judgment and winning of the current trial. Feedback was presented for 1 s, while preview, sampling, and judgment were self-paced. During sampling, the remaining bonus points (green bar) as well as the array of bead samples were visualized and updated after each additional sample. (b) Optimal sampling strategy vs. participants' performance for each of the six cost by evidence conditions. On a specific trial, the expected probability of correctness (dashed lines) and the remaining bonus points (dotted lines) are respectively increasing and decreasing functions of the number of bead samples. As their multiplication product, the expected gain (solid lines) first increases and then decreases with the number of samples. Note that the sample size that maximizes expected gain varies across different cost and evidence conditions. In general, participants in all AQ groups (each circle for one participant) sampled in a way that their winnings were close to optimality.

People with high AQ are more likely to make costly oversampling

We computed efficiency—the expected gain for participants' sample sizes divided by the maximum expected gain—to quantify the optimality of participants' sampling choices and used linear mixed model analyses to identify the effects of AQ and its interactions with sampling cost and information gain (LMM1 for efficiency, see Methods). Participants' efficiency (Fig 2a) was on average 94% (i.e. close to optimality) but decreased with increasing cost ($F_{2,41} = 33.20$, $p < .001$) or decreasing evidence ($F_{1,42} = 66.24$, $p < .001$), and decreased more dramatically when high cost and low evidence occurred together (interaction $F_{2,84} = 47.88$, $p < .001$). Though the three AQ groups did not differ in overall efficiency, AQ influenced efficiency through its interaction with cost and evidence (three-way interaction $F_{4,48} = 6.85$, $p < .001$). In particular, a significant simple main effect of AQ group was found only under the high cost (0.4 points per sample) and low evidence (60/40) condition ($F_{2,48.98} = 5.28, p = .008$), with the high AQ group bearing a lower efficiency compared to those of the low $t_{48.98} = 2.64,\ p = .027$) and middle ($t_{48.98} = 2.97,\ p = .013$) AQ groups (single-step corrections were applied to the *p* values of these post-hoc comparisons and all the post-hoc analyses below, see Methods).

The overall high efficiency was accompanied by adaptive sampling behaviors that were modulated by both sampling cost and information gain: Participants drew fewer samples in costlier or more informative conditions (Fig 2b). A linear mixed model analysis on number of bead samples (LMM2, see Methods) showed main effects of cost ($F_{2,41} = 65.67$, $p < .001$) and evidence ($F_{1,42} = 155.44$, $p < .001$), as well as their interactions ($F_{2,84} = 20.20$, $p < .001$). Similar to its influence on efficiency, AQ did not

lead to a general tendency of oversampling or undersampling but had significant interactions with evidence ($F_{2,42} = 4.08$, $p = .024$) and cost ($F_{4,48.40} = 3.01$, $p = .027$). Although the simple main effects of AQ groups were not significant under any sampling cost ($ps > .13$) or evidence conditions ($ps > .18$), the post hoc interaction contrasts suggested participants of high AQ were less flexible to further reduce sampling when cost went from low to high ($t_{42} = -2.77,\ p = .044$) and sampled more beads than those of middle AQ when the evidence gain of each bead decreased ($t_{42} = 2.66, p = .029$).

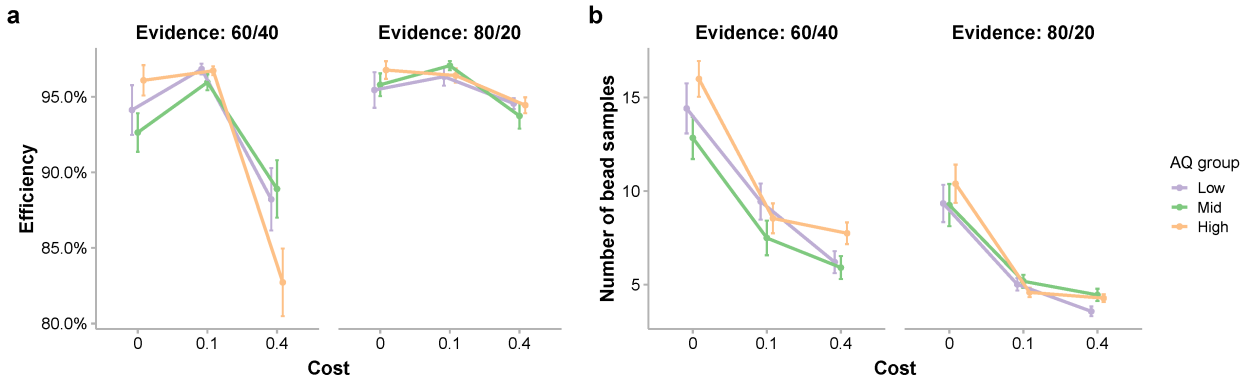

Fig 2. Mean efficiency (a) and number of bead samples (b) of the three AQ groups. Error bars denote standard error. (a) The high AQ group had a lower efficiency than the low and middle AQ groups under the high cost and low evidence condition. (b) This lower efficiency in high AQ participants was associated with a lower flexibility of reducing sampling when cost went from low (0.1) to high (0.4).

Bimodal decision times suggest two consecutive decision processes

Decision time (DT) for a specific sample—the interval between the onset of last bead (or, for the first sample, the start of the sampling phase) and the key press to draw the sample—provided further information about the cognitive process underlying sampling choices. Though decision or response times usually have a positively skewed unimodal distribution and are close to Gaussian when log-transformed [30,31], in our experiment,

log-transformed DTs for continuing sampling had a bimodal distribution (Hartigan's dip test for multimodality, $D = 0.003$, $p < .001$) and were well fit by the mixture of two Gaussian distributions (Fig 3a). Similar bimodal distributions were observed for individual participants (S1 Fig) and could not simply be artifacts of data aggregation.

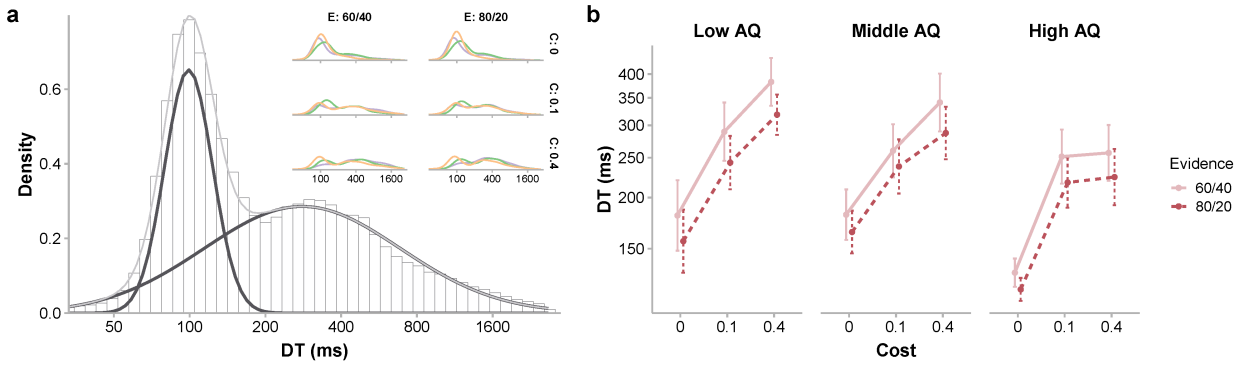

Fig 3. **Decision time (DT) of each sampling.**

(a) The distributions of DTs aggregated over all participants (main plot) and for each AQ group and condition (insets, where different AQ groups are color coded as in Fig 2). In the main plot, the distribution of DTs (histogram) was clearly bimodal, well fitted by a Gaussian mixture (gray curve) with two Gaussian components (black curves). The bimodality was also visible in most inset plots, though the relative weights of the two components varied with experimental conditions. (b) Mean DT for each AQ group compared across different cost and evidence conditions. Error bars represent standard error.

Linear mixed model analysis (LMM3) showed that the mean decision times (Fig 3b) increased with cost ($F_{2,41} = 50.30$, $p < .001$) and decreased with evidence ($F_{1,42} = 56.15$, $p < .001$).

A straightforward explanation for the bimodal DT distribution would be a probabilistic mixture of two cognitive processes. Next, we used computational modeling to explore the possibility of two decision stages and showed that it could quantitatively predict the effects of cost and evidence as well as the bimodal distribution of DTs.

Sampling is controlled by cost and evidence in two separate stages

We considered a variety of models for sampling choices, which fell into two categories: one-stage models and two-stage models (Fig 4a, see Methods). In one-stage models, the choice of whether to take a *(j+1)*-th sample after *j* samples is modeled as a Bernoulli random variable, with the probability of stopping controlled by cost- and evidence-related factors, including the expected cost and evidence for the prospective sample and the total cost and evidence of existing samples. To separate the influences of different factors on participants' sampling choices, we constructed a set of one-stage models that are controlled either by cost-related factors, or by evidence-related factors, or by both. To test the possibility that people of higher autistic traits may overweight recent evidence in evidence integration [4,17], we also considered models with an evidence decay parameter, in which the weight for an earlier sample decays as a function of the number of samples thereafter.

In two-stage models of sampling choices, we assumed that deciding whether to stop or continue sampling may involve two consecutive decision stages: if stopping sampling is the decision in the first stage, it is finalized and there is no second stage; otherwise, either continuing sampling becomes the final decision, or the decision is re-evaluated in a second stage. Whether to enter the second stage is probabilistic, whose probability ("second-thought probability") may vary with sampling cost. The decisions in the two stages are independent and controlled separately by the cost- and evidence-related factors and may be subject to evidence decay. In other words, the decision in each stage is similar to that of a one-stage model. Different two-stage models may differ in which factors control the first stage and which control the second stage. For clarity of

presentation, the results of two-stage models below only included those that have the same influencing factors as the best one-stage model (see S2 Fig for the results of additional two-stage models we considered).

We fit all the models to participants' sampling choices separately for each participant using maximum likelihood estimates and compared their goodness-of-fit using the Akaike Information Criterion corrected for small samples (AICc) [32,33] and group-level Bayesian model selection [34,35]. Lower AICc indicates better fit. Among the four one-stage models (Fig 4b), the best model was the one that is influenced by both cost and evidence as well as by the decay of evidence (Cost + Evidence w/ decay). However, its counterpart two-stage model with the first and second stages respectively controlled by cost- and evidence-related factors (Cost → Evidence w/ decay) performed even better, whose probability of outperforming all the other models (protected exceedance probability) was 0.69.

In general, two-stage models where cost was considered prior to evidence outperformed their counterpart one-stage models and two-stage models that assumed the reverse order (Fig 4b, but see one exception in S2 Fig). These results echoed our findings that DTs were bimodally distributed and were heavily influenced by cost.

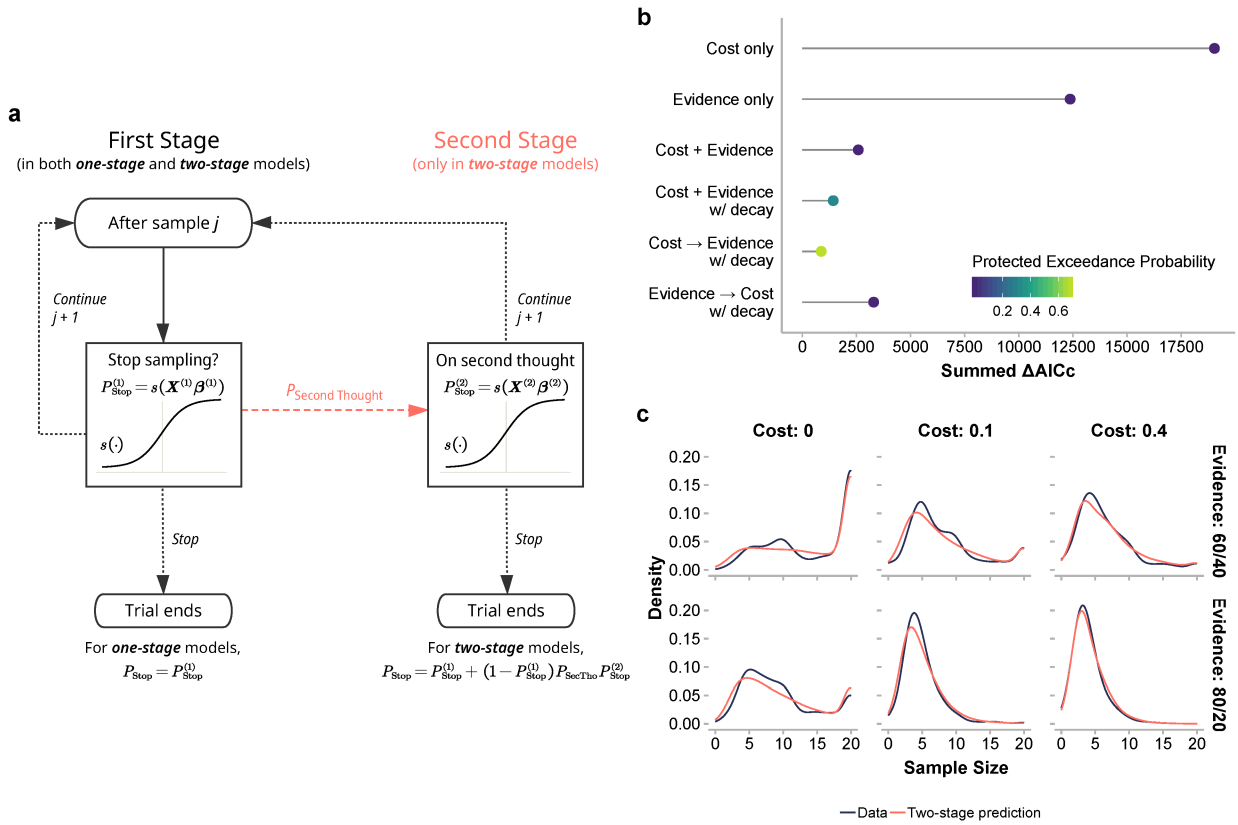

Fig 4. **Computational modeling of sampling choices**.

(a) Schematic of one-stage and two-stage models. One-stage models only consist of the steps on the left-hand side: Each time the participant decides whether to stop or continue sampling, the probability of stopping is a sigmoid function of a linear combination of multiple decision variables. Two-stage models assume that participants may probabilistically have a second thought to reconsider the choice (the coral dashed arrow) if they have not chosen to stop sampling during the first stage. The second stage (on the right-hand side) works in the same way as the first stage but the two stages are controlled by different sets of decision variables. (b) Results of model comparison. The ΔAICc for a specific model was calculated for each participant with respect to the participant's best-fitting (i.e. lowest-AICc) model and then summed across participants. Both fixed-effect (summed ΔAICc: lower is better) and random-effect (protected exceedance probability: higher is better) comparisons revealed that the best-fitting model was a two-stage model with cost-related variables considered in the first stage and evidence-related variables in the second stage (i.e. the Cost→Evidence w/ decay model). From top to down, the first four models are one-stage models and the last two are two-stage models. "Cost only" represents the one-stage model that is controlled only by cost-related information. Similarily, "Evidence only" represents the one-stage model controlled only by evidence-related information. "Cost+Evidence" and "Cost+Evidence w/ decay" represent one-stage models that are controlled by both information, with "w/ decay" referring to the decay of evidence across sampling. Models with "→" are two-stage models whose first-stage and second-stage are respectively controlled by decision variables on the left and right sides of the arrow. (c) Distribution of sample sizes (i.e. number of bead samples) for each condition: data vs. model

prediction. The best-fitted two-stage model well predicted the observed distribution of sample sizes for each cost and evidence condition. Both data and model prediction were aggregated across participants.

High AQ people less use the second stage to reduce costly sampling

How do autistic traits influence the decision process? The estimated parameters of the best model—a two-stage model with decaying evidence—allowed us to characterize individual participants' sampling choices from three aspects: cost- or evidence-related weights (11 parameters), second-thought probabilities (three parameters separately for the three cost conditions), and evidence decay rate (one parameter). We performed a MANOVA for each class of parameters (MANOVA was reduced to ANOVA for the one-parameter decay rate) and only second-thought probabilities significantly differed between AQ groups (Wilk's $\Lambda = 0.68$, $F_{6,82} = 2.79$, $p = .016$). Further one-way ANOVAs showed that the mean second-thought probabilities of the three AQ groups were comparable in the zero-cost condition ($F_{2,42} = 1.00$, $p = .84$) but had significant differences in the low-cost ($F_{2,42} = 3.84$, $p = .030$) and high-cost conditions ($F_{2,42} = 3.52$, $p = .038$). Compared with the other two AQ groups, low AQ participants had a lower probability to use the second stage under the 0.1 sampling cost, whereas high AQ participants had a lower probability to use the second stage under the 0.4 sampling cost (Fig 5a).

That is, participants with different levels of autistic traits seemed to differ in behavioral flexibility, that is, how they changed second-thought probabilities with cost. Indeed, when the cost of sampling went from zero to low (Fig 5b, left), most participants increased their second-thought probabilities and the increase was positively correlated

with AQ score (Spearman $r_S = .37, p = .013$). However, when the cost went from low to high (Fig 5b, right), a considerable proportion of participants barely increased their second-thought probabilities, and the increase was negatively correlated with AQ score (Spearman $r_S = -.36, p = .016$). That is, higher AQ did not impair participants' distinction between zero and low costs but impaired their ability to flexibly adjust second-thought probabilities for different levels of costs.

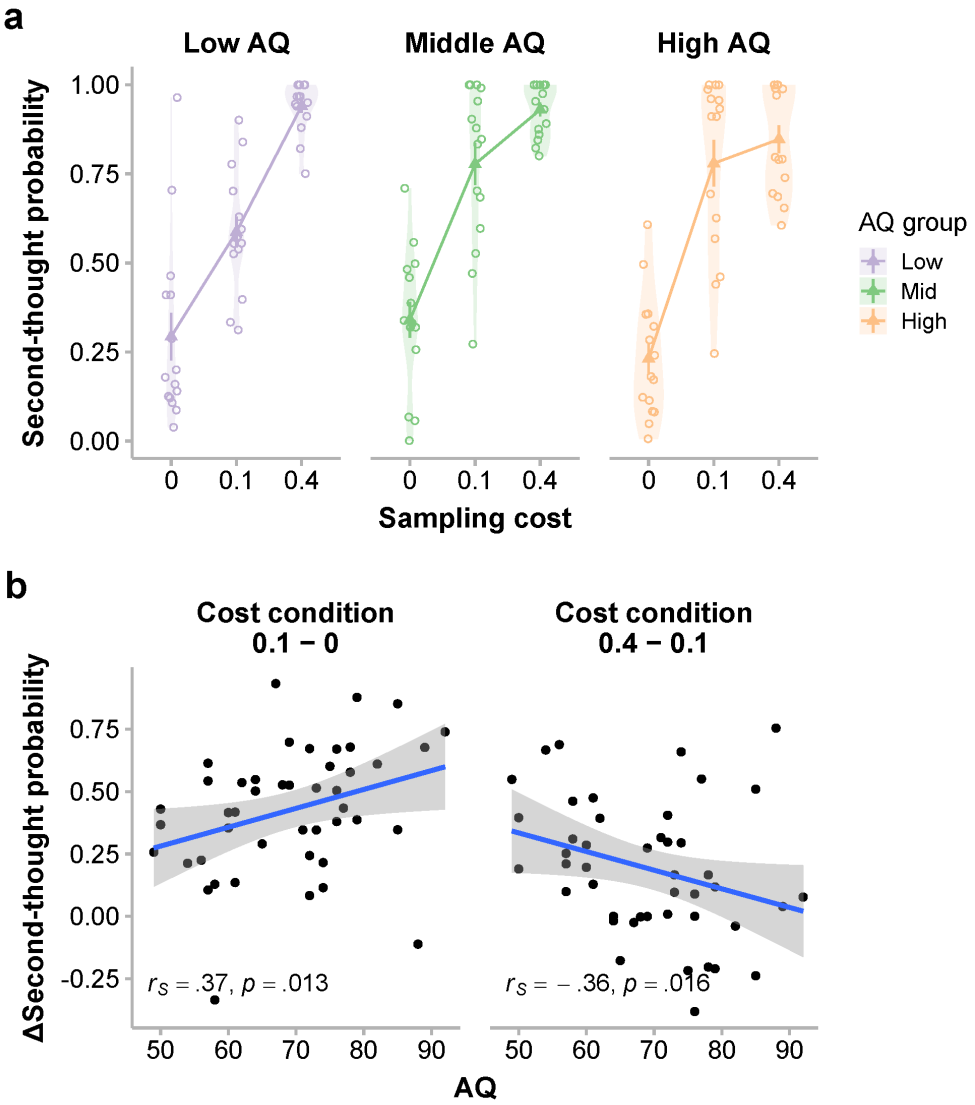

Fig 5. Autistic-trait-related differences in second-thought probability.

(a) Estimated second-thought probabilities for different AQ groups and cost conditions. Compared with the other two AQ groups, low AQ participants had a lower probability to use the second stage under low sampling cost, whereas high AQ participants had a lower probability to use the second stage under the high cost. Each circle denotes one participant's second-thought probability in a specific condition. Triangles denote the means across participants. Error bars denote standard error. Shaded areas indicate distributions of the data points via violin plots. (b) Correlation between AQ scores and cost-dependent changes of the second-thought probability. ΔSecond-thought probability refers to the difference in second-thought probability between two cost conditions. Each black dot is for one participant. Blue lines are regression lines and shade areas represent standard error.

That AQ influenced second-thought probabilities but not cost- or evidence-related weights also holds for other two-stage models we considered (S1 Table). The second decision stage provides an additional opportunity to choose stopping sampling when the first stage does not reach such a choice. High AQ participants' lower second-thought probability in the high-cost condition would thus lead to the costly oversampling and low efficiency we reported earlier (Fig 1).

Slow decisions correlate with the use of the second stage

Had the bimodal DT distribution arisen from the two-stage decision process assumed in two-stage models, as we conjectured earlier, the mean DT—as a proxy for the proportion of slow decisions—should increase with the probability of using the second stage. Indeed, the mean DT was positively correlated with second-thought probability (Fig 6; Spearman $r_S = .56, p < .001$). The positive correlation also held for each separate cost condition (zero cost: $r_S = 0.45, p = .001$; low cost: $r_S = .30, p = .044$), though the correlation under the high cost condition did not reach statistical significance due to the relatively narrower range of second-thought probability ($r_S = .27, p = .07$). Moreover, the effects of cost on mean DT (LMM3, as we reported earlier) could be

explained away by the effect of second-thought probability when the latter was added as a predictor (LMM4; second-thought probability, $F_{1,35.83} = 9.76, p = .004$; cost, $F_{2,33.53} = 1.39, p = .26$).

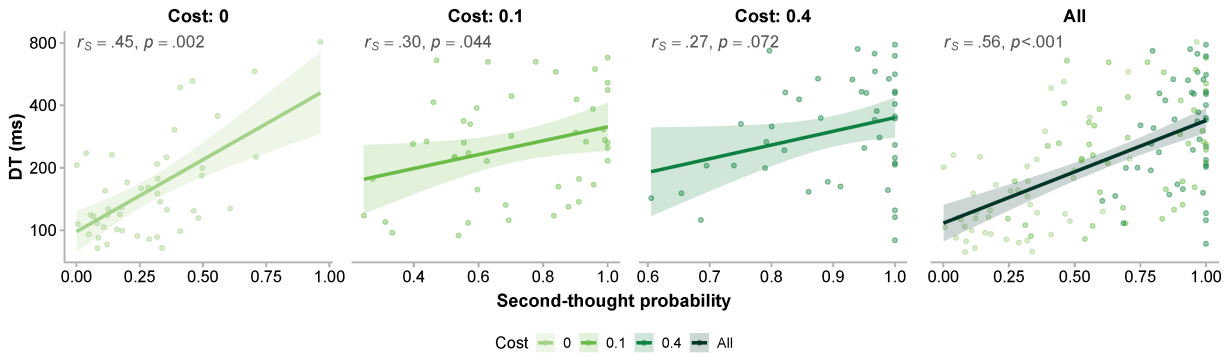

Fig 6. **Positive correlation between mean decision time and second-thought probability.** Each dot is for one participant in one specific cost condition. Lines and shaded areas respectively represent regression lines and standard error.

Discussion

Humans must sample the environment properly to balance the advantage of gaining additional information against the cost of time, energy, and money [36]. Previous research suggests that suboptimal information sampling may be a fundamental deficit in ASD [4,14–17,37]. In the current study, we tested healthy adults with different levels of autistic traits to investigate how autistic traits influence information sampling decisions. We found that participants adjusted their sample sizes according to both sampling cost and evidence gain and were overall close to optimality. However, in the high cost and low evidence condition where the tension between minimizing cost and gathering information was highest, individuals with high autistic traits were farther away from optimality compared to their peers due to oversampling. Further computational modeling revealed which components of the decision process were affected by and which were immune to autistic traits. Participants' sampling choices were best fit by a two-stage decision model, with the second stage serving as a second thought to stop sampling. Participants with high autistic traits differed from others only in that they were less likely to use the second thought to reduce sampling behaviors that had become costly.

Previous ASD studies that had used similar bead-sampling tasks yielded inconclusive results: one study found that adolescents with ASD sampled more than the control group [38], whereas a second study of adults with ASD found the reverse [27]. In contrast, we did not find overall oversampling but more subtle atypicality in healthy people with high autistic traits. These seemingly contradicting findings suggest that to ask whether people with ASD or high autistic traits oversample or undersample information is probably not a proper question. In fact, both oversampling and

undersampling may lower one's expected gain, depending on the rewarding structure of the environment. As we suggested in the Introduction, a more important question is whether autistic traits influence one's ability to sample optimally, that is, to balance sampling cost and information gain. In previous ASD studies [27,38], sampling incurred no explicit cost but implicit cost such as time or cognitive effort whose exact value to a specific individual is hard to measure, therefore we could hardly compare the optimality of different individuals' performances. By introducing explicit monetary cost for sampling (as Juni et al. did [36]) in our experiment, we were able to evaluate sampling cost as a potential moderator for autistic trait-related differences in information sampling. Indeed, we found that people with high autistic traits tend to be less optimal than those of low or middle autistic traits when sampling cost is high.

It is striking to find that the autistic traits of healthy individuals influence their abilities in performing a non-social and non-perceptual task like information sampling. However, our finding is consistent with the Bayesian theories that suboptimal inference and decision making are primary to various symptoms of ASD (see Palmer et al. for reviewer) [4]. It is important to further identify the autistic-trait-related cognitive processes that lead to suboptimal decisions. Sevgi, Diaconescu, Tittgemeyer, & Schilbach [39] demonstrated how computational modeling can be a powerful tool in deepening our understanding of such processes and separating the affected processes from the intact ones. They found that autistic traits do not, as usually believed, influence individuals' ability of learning social cues but only influence the weight assigned to social cues in decision making.

Here, we aimed to understand through which processes high levels of autistic traits may lead to suboptimal information sampling. Computational modeling allowed us to characterize the decision process of sampling choices by two stages. The existence of a probabilistic second stage is also supported by the bimodality of DT distributions as well as by the positive correlation between mean DT and the second-thought probability estimated from sampling choices.

The autistic-trait-related differences we found in decision processes are surprisingly selective. Participants with high autistic traits were indistinguishable from other participants in their ability to weigh sampling cost and evidence gain in the two stages. What they lacked was the flexibility to increase the second-thought probability appropriately under the increase of sampling cost. In other words, as sampling cost increased, participants need to increase their strategic complexity—that is, to have more second thoughts—to reduce unprofitable sampling behaviors but participants with high autistic traits failed to do it well.

Studies using autistic traits as a surrogate for studying ASD have revealed congruent and converging autistic-trait-related effects as those of ASD [9,10,39–43]. Given that participants with high-level autistic traits in our sample had AQ scores overlapping with people with ASD, we expect the latter to suffer from similar impairments in information sampling. Future research is needed to directly test this possibility in people with ASD.

To summarize, we find that people with high autistic traits are suboptimal when sampling cost is high and their suboptimality is associated with insufficient adjustment of strategic complexity for cost increase. Recent studies suggest that autistic traits may

influence an individual's ability of adaptively using her own information processing capability while not influencing the capability itself. For example, autistic traits may only influence the tendency to use social information but not the capability to perceive it [39], or may only influence the flexibility of updating learning rate but not probabilistic learning itself [10]. Our results add to this line of findings that autistic-trait-related suboptimality may come from impairments in higher-level cognitive functions rather than primary information processing.

In addition, our results point to a possible fundamental drive of all kinds of suboptimality: inappropriate adjustment to cost. It is not a total ignorance of cost, since people of high autistic traits do reduce their sampling in the presence of cost compared to zero cost. Their first-stage decision is also properly controlled by cost. People with high autistic traits only fail to appropriately adjust the use of second thoughts according to the quantitative change of cost. This kind of insensitivity to cost, though subtle, would in general hinder one's ability to choose actions that maximize expected gain, that is, to be Bayesian optimal. Therefore, failing to adjust one's strategic complexity with cost may have penetrative effects in human cognition and even be a major cause for the core symptoms of ASD.

Methods

Ethics Statement

The experiment had been approved by the Institutional Review Board of School of Psychological and Cognitive Sciences at Peking University. All participants provided written informed consent and were paid for their time plus performance-based bonus.

Experiment

Participants. Forty-five college student volunteers (18 male, aged 18-28) participated in our experiment. Three additional participants were excluded (one for being an IQ outlier, one for misunderstanding instructions, one for a strong bias towards one type of stimuli).

IQ test. Combined Raven Test (CRT) was used to measure participants' IQ for control purpose. Raw CRT scores of all 48 participants averaged 68.10 (s.d., 3.53) and ranged from 54 to 72. One of the participants (scoring 54) fell out of two standard deviations of the mean and was excluded from further analyses along with two other participants (as mentioned above). The remaining 45 participants had a mean CRT score of 68.58 (s.d., 2.81; ranging from 62 to 72), corresponding to a mean IQ score of 113.86.

AQ test. Autism Spectrum Quotient (AQ) questionnaire [18] was used to quantify participants' autistic traits. AQ questionnaire is a 4-point self-reported scale with 50 items measuring five type of autistic characteristics: social interaction, attentional switch, attention to detail, imagination, and communication. Though the 4-point scale was sometimes reduced to binary coding [18], we adopted the full 4-point scoring

system (“definitely disagree”, “slightly disagree”, “slightly agree”, “definitely agree” respectively scored 0–3) to maximize the coverage of latent autistic traits [25,28,44,45].

The AQ scores of the 45 participants were normally distributed (Shapiro-Wilk normality test, $W = 0.98$, $p = .78$; S3 Fig) with mean 69 and standard deviation 11. They were evenly divided into three groups of low-AQ (mean, 56.7; s.d., 4.2; range, 49-62), middle-AQ (mean, 69.8; s.d., 3.5; range, 64-74), and high-AQ (mean, 81; s.d., 5.5; range, 75-92), whose mean age (21.73 ± 1.91, 22.20 ± 3.00, 20.93 ± 1.71) and IQ (in CRT score, 68.53 ± 2.77, 69.80 ± 3.53, 69.07 ± 3.24) were statistically comparable (age: $F_{2,42} = 1.18, p = .32$; IQ: $F_{2,42} = 0.41, p = .67$).

Apparatus. All stimuli of the bead-sampling task were visually presented on a 21.5-inch computer screen controlled by MATLAB R2016b and PsychToolbox [46–48]. Participants were seated approximately 60 cm to the screen. Responses were recorded via the keyboard.

Procedure. On each trial of the experiment (Fig 1a), participants saw a pair of jars on the left and right of the screen, each containing 200 pink and blue beads. The pink-to-blue ratios of the two jars were either 60%:40% vs. 40%:60%, or 80%:20% vs. 20%:80%. Participants were told that one jar had been secretly selected, and their task was to infer which jar was selected. Each time they pressed the space bar, one bead was randomly sampled with replacement from the jar and presented on the screen, appended to the end of the sampled bead sequence. Participants were free to draw 0 to 20 bead samples, but each sample might incur a cost. The cost per sample on each trial could be 0, 0.1, or 0.4 points. A green bar on the top of the screen indicated how many

bonus points remained (10 points minus the total sampling cost by then). When participants were ready for inference, they pressed the Enter key to quit sampling and judged whether the pre-selected jar was the left or right jar by pressing the corresponding arrow key. Feedback followed immediately. If their judgment was correct, participants would receive the remaining bonus points; otherwise nothing. Bonus points accumulated across trials and would be converted into monetary bonus after the experiment. Participants were encouraged to sample wisely to maximize their winning.

The pink-dominant jar was pre-selected on half of the trials and the blue-dominant jar on the other half. Their left/right positions were also counterbalanced across trials. In the formal experiment, the two evidence (i.e. bead ratio) conditions (60/40 and 80/20) were randomly mixed within each block and the three cost conditions (0, 0.1, and 0.4) were blocked. Besides being visualized by the green bar on each trial, cost for each block was also informed at the beginning of the block. The order of cost blocks was counterbalanced across participants. Each of the six conditions was repeated for 48 times, resulting in 288 trials. The formal experiment were preceded by 24 practice trials. Participants first performed the experiment, then the Combined Raven Test and last the AQ questionnaire, which took approximately 1.5 hours in total.

Statistical Analyses

All statistical analyses (except for group-level Bayesian model comparison) were conducted in R 3.4.3 [49].

Linear mixed models (LMMs). Linear mixed models were estimated using “afex” package [50], whose *F* statistics, degrees of freedom of residuals (denominators), and *p*-values were approximated by Kenward-Roger method [51,52]. Specifications of random effects followed the suggestion of “parsimonious modeling”[53]. For significant fixed effects, “emmeans” package was used to test post hoc contrasts [54]. Interaction contrasts were performed for significant interactions and, when higher-order interactions were not significant, pairwise or consecutive contrasts were performed for significant main effects. Statistical multiplicity of the contrasts was controlled by a single-step adjustment, which used multivariate *t* distributions to estimate the critical value for conducted contrasts [55,56].

LMM1: decision efficiency is the dependent variable; fixed effects include an intercept, the main and interaction effects of groups, costs, and ratios (evidence); random effects include correlated random slopes of costs and ratios within participants and random participant intercept.

LMM2: number of bead samples is the dependent variable; the fixed and random effects are the same as LMM1.

LMM3: decision time is the dependent variable; the fixed and random effects are the same as LMM1.

LMM4: the dependent variable is the same as LMM3; in addition to the fixed and random effects of LMM1, the linear effect of second-thought probability is included in the fixed effects, and a random slope of the second-thought probability that is uncorrelated with the random intercept is included in the random effects.

Decision times (DTs). Because stopping sampling involved a different key press, only DTs for continuing sampling were analyzed. Before any analysis of DTs, the outliers of log-transformed DTs were excluded by nonparametric boxplot statistics based on interquartile ranges (IQR) of all the log-transformed DTs (i.e., valid data are less than 1.5 × IQR away from 1st/3rd quartile).

Modeling

Expected gain. Given a specific sequence of bead samples, an ideal observer would always judge the pre-selected jar to be the one whose dominant color is the same as that of the sample sequence. In the case of a tie, the observer would choose the two jars with equal probability. Suppose the sample size is *n*, the maximal reward is 10 points, the unit sampling cost is *c*, and the percentage of predominated beads in the pre-selected jar is *qq*. The expected probability of correct judgment is:

$$p(n|q) = \begin{cases} \sum_{i=\frac{n+1}{2}}^{n} \binom{n}{i} q^i (1-q)^{n-i} & n = 2k-1,\ k = 1, 2, ..., 10\ ; \\ \frac{1}{2}\binom{n}{\frac{n}{2}} q^{\frac{n}{2}} (1-q)^{\frac{n}{2}} + \sum_{i=\frac{n}{2}}^{n} \binom{n}{i} q^i (1-q)^{n-i} & n = 2k, k = 1, 2, ..., 10\ ; \\ \frac{1}{2} & n = 0\ . \end{cases} \quad (1)$$

The expected gain is $\mathrm{E}[Gain|n,q,c] = (10 - nc)p(n|q)$. For a specific cost and evidence condition, the optimal sample size is the value of n that maximizes $\mathrm{E}[Gain|n,q,c]$.

One-stage models. We modeled participants' each choice of whether to continue or stop sampling (i.e. whether to press the space bar or Enter key) as a Bernoulli random variable, with the probability of stopping sampling determined by cost- or evidence-related factors. Pressing the Enter key after 20 samples was not included as a choice of stopping sampling, because participants had no choice but to stop by then.

We considered two families of models: one-stage and two-stage models. In one-stage models, the probability of stopping sampling on the *i*-th trial after having drawn *j* beads is determined by a linear combination of *K* decision variables (DVs) via a logistic function:

$$p_{ij} = \frac{1}{1 + e^{-X_{ij}}}, \tag{2}$$

$$X_{ij} = \sum_{k=1}^{K} \beta_k DV_{ijk}. \tag{3}$$

Different one-stage models differed in whether cost-related variables, evidence-related variables, or both served as DVs.

Cost-only one-stage model: cost-related variables as DVs, including unit cost per bead (categorical: 0, 0.1, or 0.4), number of beads sampled (*j*), and total sampling cost (product of the former two DVs).

Evidence-only one-stage model: evidence-related variables as DVs, including unit log evidence per bead (i.e., ln(60/40) or ln(80/20)), absolute value of cumulative information (cumulative information refers to the difference between the numbers of pink and blue bead samples), total log evidence (product of the former two DVs), and the correctness and the number of bead samples in last trial.

Cost + evidence one-stage model: both cost-related and evidence-related variables as DVs.

Cost + evidence with decay one-stage model: both cost-related and decaying evidence-related variables as DVs.

In models with decaying evidence, cumulative information (CI) is modulated by a decay parameter α:

$$\mathrm{CI}_{ij} = \begin{cases} 0 & j = 0; \\ \alpha\,\mathrm{CI}_{i,j-1} + 1 & j > 0,\ \text{after a pink bead}; \\ \alpha\,\mathrm{CI}_{i,j-1} - 1 & j > 0,\ \text{after a blue bead}. \end{cases} \tag{4}$$

The DVs of absolute value of cumulative information and total log evidence in the decay models are modulated by the decay parameter accordingly.

Two-stage models. In two-stage models, sampling choices may involve two decision stages: when the first stage reaches the choice of stopping sampling, it is finalized; otherwise there is chance a second stage is engaged in the choice of continuing or stopping sampling. The overall probability of stopping sampling can thus be written as:

$$p_{ij} = p_{ij}^{Stage1} + (1 - p_{ij}^{Stage1})\, p(\mathrm{sec}|C)\, p_{ij}^{Stage2}\,, \tag{5}$$

$$p_{ij}^{Stage\,1} = \frac{1}{1+e^{-X_{ij}^{Stage1}}}\ , \tag{6}$$

$$p_{ij}^{Stage2} = \frac{1}{1+e^{-X_{ij}^{Stage2}}}\ . \tag{7}$$

Here $p_{ij}^{Stage\,1}$ and $p_{ij}^{Stage\,2}$ denote the stopping probabilities of the two stages and $p(\mathrm{sec}|C)$ denotes second-thought probability—the probability of using the second stage given that the first stage concludes with continuing sampling, which may vary with the unit sampling cost.

Each stage works in the same way as one-stage models do (Eq. 1-3) and is influenced by mutually exclusive sets of DVs. Different two-stage models differed in which DVs influenced the first stage and which for the second stage.

Cost → evidence two-stage model: cost-related variables as first-stage DVs and (decaying) evidence-related variables as second-stage DVs.

Evidence → cost two-stage model: evidence-related variables as first-stage DVs and (decaying) cost-related variables as second-stage DVs.

Model fitting. All coefficients β_k of decision variables, three second-thought probabilities $p(\mathrm{sec}|C)$, and the decay parameter α were estimated as free parameters using maximum likelihood estimates. All parameters were unbounded except for $p(\mathrm{sec}|C)$ and α (both bounded to [0, 1]). Optimization was implemented by the *fmincon* function with interior-point algorithm in MATLAB R2017a.

608 # References

609 1. Dall S, Giraldeau L, Olsson O, Mcnamara J, Stephens D. Information and its use by
610 animals in evolutionary ecology. Trends Ecol Evol. 2005;20: 187–193.
611 doi:10.1016/j.tree.2005.01.010

612 2. Stephens DW, Krebs JR. Foraging Theory. Princeton University Press; 1986.

613 3. American Psychiatric Association. Diagnostic and Statistical Manual of Mental
614 Disorders (DSM-5®). American Psychiatric Pub; 2013.

615 4. Palmer CJ, Lawson RP, Hohwy J. Bayesian approaches to autism: Towards volatility,
616 action, and behavior. Psychol Bull. 2017;143: 521–542. doi:10.1037/bul0000097

617 5. Au-Yeung SK, Kaakinen JK, Benson V. Cognitive Perspective-Taking During Scene
618 Perception in Autism Spectrum Disorder: Evidence From Eye Movements. Autism Res.
619 2014;7: 84–93. doi:10.1002/aur.1352

620 6. Song Y, Hakoda Y, Sang B. A selective impairment in extracting fearful information
621 from another's eyes in Autism. Autism Res. 2016;9: 1002–1011. doi:10.1002/aur.1583

622 7. Chambon V, Farrer C, Pacherie E, Jacquet PO, Leboyer M, Zalla T. Reduced sensitivity
623 to social priors during action prediction in adults with autism spectrum disorders.
624 Cognition. 2017;160: 17–26. doi:10.1016/j.cognition.2016.12.005

625 8. Goris J, Braem S, Nijhof AD, Rigoni D, Deschrijver E, Cruys SV de, et al. Sensory
626 Prediction Errors Are Less Modulated by Global Context in Autism Spectrum Disorder.
627 Biol Psychiatry Cogn Neurosci Neuroimaging. 2018;0. doi:10.1016/j.bpsc.2018.02.003

628 9. Lawson RP, Aylward J, Roiser JP, Rees G. Adaptation of social and non-social cues to
629 direction in adults with autism spectrum disorder and neurotypical adults with
630 autistic traits. Dev Cogn Neurosci. 2018;29: 108–116. doi:10.1016/j.dcn.2017.05.001

631 10. Lawson RP, Mathys C, Rees G. Adults with autism overestimate the volatility of the
632 sensory environment. Nat Neurosci. 2017;20: nn.4615. doi:10.1038/nn.4615

633 11. Manning C, Tibber MS, Charman T, Dakin SC, Pellicano E. Enhanced Integration of
634 Motion Information in Children With Autism. J Neurosci. 2015;35: 6979–6986.
635 doi:10.1523/JNEUROSCI.4645-14.2015

636 12. Palmer CJ, Paton B, Kirkovski M, Enticott PG, Hohwy J. Context sensitivity in action
637 decreases along the autism spectrum: a predictive processing perspective. Proc R Soc
638 Lond B Biol Sci. 2015;282: 20141557. doi:10.1098/rspb.2014.1557

639 13. Turi M, Burr DC, Igliozzi R, Aagten-Murphy D, Muratori F, Pellicano E. Children with
640 autism spectrum disorder show reduced adaptation to number. Proc Natl Acad Sci U S
641 A. 2015;112: 7868–7872. doi:10.1073/pnas.1504099112

14. Lawson RP, Rees G, Friston KJ. An aberrant precision account of autism. Front Hum Neurosci. 2014;8. doi:10.3389/fnhum.2014.00302

15. Pellicano E, Burr D. When the world becomes 'too real': a Bayesian explanation of autistic perception. Trends Cogn Sci. 2012;16: 504–510. doi:10.1016/j.tics.2012.08.009

16. van Boxtel JJA, Lu H. A predictive coding perspective on autism spectrum disorders. Front Psychol. 2013;4. doi:10.3389/fpsyg.2013.00019

17. van de Cruys S, Evers K, van der Hallen R, van Eylen L, Boets B, de-Wit L, et al. Precise minds in uncertain worlds: Predictive coding in autism. Psychol Rev. 2014;121: 649–675. doi:10.1037/a0037665

18. Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E. The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. J Autism Dev Disord. 2001;31: 5–17.

19. Hoekstra RA, Bartels M, Verweij CJH, Boomsma DI. Heritability of Autistic Traits in the General Population. Arch Pediatr Adolesc Med. 2007;161: 372–377. doi:10.1001/archpedi.161.4.372

20. Lundström S, Chang Z, Råstam M, Gillberg C, Larsson H, Anckarsäter H, et al. Autism Spectrum Disorders and Autisticlike Traits: Similar Etiology in the Extreme End and the Normal Variation. Arch Gen Psychiatry. 2012;69: 46–52. doi:10.1001/archgenpsychiatry.2011.144

21. Robinson EB, Koenen KC, McCormick MC, Munir K, Hallett V, Happé F, et al. 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. 2011;68: 1113–1121. doi:10.1001/archgenpsychiatry.2011.119

22. Ronald A, Hoekstra RA. Autism spectrum disorders and autistic traits: A decade of new twin studies. Am J Med Genet B Neuropsychiatr Genet. 2011;156: 255–274. doi:10.1002/ajmg.b.31159

23. Sucksmith E, Roth I, Hoekstra RA. Autistic Traits Below the Clinical Threshold: Re-examining the Broader Autism Phenotype in the 21st Century. Neuropsychol Rev N Y. 2011;21: 360–89. doi:http://dx.doi.org/10.1007/s11065-011-9183-9

24. Wheelwright S, Auyeung B, Allison C, Baron-Cohen S. Defining the broader, medium and narrow autism phenotype among parents using the Autism Spectrum Quotient (AQ). Mol Autism. 2010;1: 10. doi:10.1186/2040-2392-1-10

25. Bralten J, Hulzen KJ van, Martens MB, Galesloot TE, Vasquez AA, Kiemeney LA, et al. Autism spectrum disorders and autistic traits share genetics and biology. Mol Psychiatry. 2017; doi:10.1038/mp.2017.98

26. Ruzich E, Allison C, Smith P, Watson P, Auyeung B, Ring H, et al. Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. Mol Autism. 2015;6: 2. doi:10.1186/2040-2392-6-2

27. Jänsch C, Hare DJ. An Investigation of the "Jumping to Conclusions" Data-Gathering Bias and Paranoid Thoughts in Asperger Syndrome. J Autism Dev Disord. 2014;44: 111–119. doi:10.1007/s10803-013-1855-2

28. Hoekstra RA, Bartels M, Cath DC, Boomsma DI. Factor Structure, Reliability and Criterion Validity of the Autism-Spectrum Quotient (AQ): A Study in Dutch Population and Patient Groups. J Autism Dev Disord. 2008;38: 1555–1566. doi:10.1007/s10803-008-0538-x

29. Zhang L, Sun Y, Chen F, Wu D, Tang J, Han X, et al. Psychometric properties of the Autism-Spectrum Quotient in both clinical and non-clinical samples: Chinese version for mainland China. BMC Psychiatry. 2016;16. doi:10.1186/s12888-016-0915-5

30. Ratcliff R, Smith PL, Brown SD, McKoon G. Diffusion Decision Model: Current Issues and History. Trends Cogn Sci. 2016;20: 260–281. doi:10.1016/j.tics.2016.01.007

31. Wolfe JM, Palmer EM, Horowitz TS. Reaction time distributions constrain models of visual search. Vision Res. 2010;50: 1304–1311. doi:10.1016/j.visres.2009.11.002

32. Cavanaugh JE. Unifying the derivations for the Akaike and corrected Akaike information criteria. Stat Probab Lett. 1997;33: 201–208. doi:10.1016/S0167-7152(96)00128-9

33. Hurvich CM, Tsai C-L. Regression and time series model selection in small samples. Biometrika. 1989;76: 297–307. doi:10.1093/biomet/76.2.297

34. Rigoux L, Stephan KE, Friston KJ, Daunizeau J. Bayesian model selection for group studies — Revisited. NeuroImage. 2014;84: 971–985. doi:10.1016/j.neuroimage.2013.08.065

35. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. Bayesian model selection for group studies. NeuroImage. 2009;46: 1004–1017. doi:10.1016/j.neuroimage.2009.03.025

36. Juni MZ, Gureckis TM, Maloney LT. Information sampling behavior with explicit sampling costs. Decision. 2016;3: 147–168. doi:10.1037/dec0000045

37. Sinha P, Kjelgaard MM, Gandhi TK, Tsourides K, Cardinaux AL, Pantazis D, et al. Autism as a disorder of prediction. Proc Natl Acad Sci. 2014;111: 15220–15225. doi:10.1073/pnas.1416797111

38. Brosnan M, Chapman E, Ashwin C. Adolescents with Autism Spectrum Disorder Show a Circumspect Reasoning Bias Rather than 'Jumping-to-Conclusions.' J Autism Dev Disord. 2014;44: 513–520. doi:10.1007/s10803-013-1897-5

39. Sevgi M, Diaconescu AO, Tittgemeyer M, Schilbach L. Social Bayes: Using Bayesian Modeling to Study Autistic Trait–Related Differences in Social Cognition. Biol Psychiatry. 2016;80: 112–119. doi:10.1016/j.biopsych.2015.11.025

40. Robic S, Sonié S, Fonlupt P, Henaff M-A, Touil N, Coricelli G, et al. Decision-Making in a Changing World: A Study in Autism Spectrum Disorders. J Autism Dev Disord. 2015;45: 1603–1613. doi:10.1007/s10803-014-2311-7

41. Turi M, Burr DC, Binda P. Pupillometry reveals perceptual differences that are tightly linked to autistic traits in typical adults. eLife. 2018;7: e32399. doi:10.7554/eLife.32399

42. van Boxtel JJA, Dapretto M, Lu H. Intact recognition, but attenuated adaptation, for biological motion in youth with autism spectrum disorder. Autism Res. 2016;9: 1103–1113. doi:10.1002/aur.1595

43. van Boxtel JJA, Lu H. Impaired Global, and Compensatory Local, Biological Motion Processing in People with High Levels of Autistic Traits. Front Psychol. 2013;4. doi:10.3389/fpsyg.2013.00209

44. Austin EJ. Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). Personal Individ Differ. 2005;38: 451–460. doi:10.1016/j.paid.2004.04.022

45. Murray AL, Booth T, McKenzie K, Kuenssberg R. What range of trait levels can the Autism-Spectrum Quotient (AQ) measure reliably? An item response theory analysis. Psychol Assess. 2016;28: 673–683. doi:10.1037/pas0000215

46. Brainard DH. The Psychophysics Toolbox. Spat Vis. 1997;10: 433–436.

47. Kleiner M, Brainard D, Pelli D. What's new in Psychtoolbox-3? Perception 36 ECVP Abstract Supplement. 2007;

48. Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spat Vis. 1997;10: 437–442.

49. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2015. Available: http://www.R-project.org/

50. Singmann H, Bolker B, Westfall J, Aust F. afex: Analysis of Factorial Experiments [Internet]. 2018. Available: https://CRAN.R-project.org/package=afex

51. Halekoh U, Højsgaard S. A Kenward-Roger Approximation and Parametric Bootstrap Methods for Tests in Linear Mixed Models - The *R* Package **pbkrtest**. J Stat Softw. 2014;59. doi:10.18637/jss.v059.i09

52. Kenward MG, Roger JH. Small Sample Inference for Fixed Effects from Restricted Maximum Likelihood. Biometrics. 1997;53: 983–997. doi:10.2307/2533558

53. Bates D, Kliegl R, Vasishth S, Baayen H. Parsimonious Mixed Models. ArXiv150604967 Stat. 2015; Available: http://arxiv.org/abs/1506.04967

54. Lenth R. emmeans: Estimated Marginal Means, aka Least-Squares Means [Internet]. 2018. Available: https://CRAN.R-project.org/package=emmeans

55. Genz A, Bretz F, Hochberg Y. Approximations to multivariate t integrals with application to multiple comparison procedures. Institute of Mathematical Statistics Lecture Notes - Monograph Series. Beachwood, Ohio, USA: Institute of Mathematical Statistics; 2004. pp. 24–32. doi:10.1214/lnms/1196285623

56. Hothorn T, Bretz F, Westfall P. Simultaneous Inference in General Parametric Models. Biom J. 2008;50: 346–363. doi:10.1002/bimj.200810425