bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Beyond over- or under-sampling:

autistic children's inflexibility in sampling costly information

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

[1] School of Psychological and Cognitive Sciences, Peking University, Beijing, China; [2] PKU-IDG/ McGovern Institute for Brain Research, Peking University, Beijing, China; [3] Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China; [4] Chinese Institute for Brain Research, Beijing, China

*Correspondence: hang.zhang@pku.edu.cn (HZ); yilipku@pku.edu.cn (LY)

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Beyond over- or under-sampling:

autistic children's inflexibility in sampling costly information

Abstract

Background: Efficient information sampling is crucial for human inference and decision-making, even for young children. Information sampling is also closely associated with the core symptoms of autism spectrum disorder (ASD), since both the social interaction difficulties and repetitive behaviors suggest that autistic people may sample information from the environment distinctively. Previous research on information sampling in ASD focused mainly on adolescents and adults, and on whether they over- or under-sample. The specific ways in which autistic children sample information, especially when facing explicit costs and adapting to environmental changes, remain unclear.

Methods: We employed an adapted bead task to investigate the sampling behavior of 24 autistic and 41 neurotypical children, matched for age and IQ. In each trial of our experiment, children gathered information about an unknown target isle by drawing samples from it and then guessed the target between two isles based on their samples. In conditions where sampling was costly, children needed to weigh the benefits of information against the costs of acquiring additional samples. Through computational modeling and intricate behavioral measures, we revealed how the two groups of children differed in sampling decisions and underlying cognitive mechanisms.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Results: Under conditions involving costs, autistic children showed less efficient sampling than their neurotypical peers. This inefficiency was due to their increased variability in the number of samples taken across trials rather than a systematic bias. Computational models indicated that while both groups shared a similar decision process, autistic children's sampling decisions were less influenced by dynamic changes and more driven by recent evidence, thus leading to their increased sampling variation and reduced efficiency.

Limitations: To refine ASD subtyping and correlate symptom severity with behavioral characteristics and computational findings, future research may need larger participant groups and more comprehensive clinical assessments.

Conclusions: This study reveals an inefficiency of autistic children in information sampling and tracks down this inefficiency to their increased sampling variability, primarily due to their cognitive preference for more local and static information. These findings are consistent with several influential behavioral theories of ASD and highlight the needs of a multi-level understanding of cognitive flexibility in ASD.

(349 words)

Keywords: autism spectrum disorder; information sampling; computational modeling; cognitive flexibility; Bayesian account

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Background

Imagine you're at a newly opened gelato shop with a vast menu of flavors. While you might be tempted to inquire about each flavor to make the perfect choice, the key challenge lies not just in making the final choice, but in knowing when to stop gathering information and proceed to a decision. This scenario mirrors a critical aspect of everyday lives of everyone, including children. From basic physical laws and causality to language and social norms, children and even infants heavily rely on seeking novel information under uncertainty, especially given their limited prior experience to refer to. Previous studies have found that children and even infants are active information seekers, who further use the gathered information to guide their decisions under uncertainty [1–9]. Moreover, information often comes with a price that could be implicit (e.g., time and effort) or explicit (e.g., monetary cost). For example, children might ask questions to gain information, but they also risk experiencing embarrassment and rejection. Despite the importance of adaptive information sampling in childhood, how children modulate their sampling based on the interplay between information gain and costs has not been thoroughly examined. Researchers have only known that young children might be suboptimal than older children or adults in trading-off information gain and sampling cost [2, 5, 10]. The more intricate cognitive process underlying the adaptive information sampling in children, particularly in autistic children, is still elusive.

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by challenges with social interactions and restricted and repetitive behaviors [11, 12]. These core symptoms, though seemingly unrelated, are deeply intertwined with how autistic people interact with and sample information from their environment. For example, social interactions

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

necessitate constant information sampling, such as gathering information from others' facial expressions, tone of speech, body language, etc. to infer the mental state of others in order to keep the conversation going [13]. Repetitive behaviors could also be viewed as inefficient information sampling in reducing uncertainty due to the prolonged time spent on redundant details [14]. In a recent Bayesian account of ASD that explains a variety of perceptual, motor, and cognitive symptoms, the deviation from Bayesian optimality in information process is primary to ASD [14]. Under this Bayesian framework, information sampling is referred as "disambiguatory active inference", whose general impairments are believed to underlie autistic symptoms. Indeed, autistic people show atypical gaze patterns in ambiguous or social scenes, sampling the visual environment in a less efficient manner [15–17].

Previous studies that directly examine information sampling in autistic people predominantly test autistic adults or adolescents. Early studies presented mixed results, with some suggesting autistic adults sample less and decide faster than neurotypical adults [18, 19], while others reporting more extensive information gathering in similar tasks [20, 21]. These discrepancies suggest that ASD may affect one's ability to sample *optimally* dependent on particular situations, instead of oversampling versus undersampling. Using an inference task with explicit sampling costs, Lu et al. found that people with higher autistic traits have more varied and less efficient sampling behaviors, due to using a more inflexible strategy of balancing information gain and cost [22]. These studies on adults or adolescents have provided insights into autistic people's sampling behaviors, but because information sampling may change significantly from childhood through adolescence to adulthood [2, 5, 9, 10], the information sampling behavior of autistic children is still elusive.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In the present study, we investigated the possible differences in information sampling between autistic children and preschoolers with IQ and age matched. Based on the results of previous studies, we hypothesized that autistic children would exhibit inefficient sampling when facing the complicated trade-off between information gain and cost, which might result from the varied sampling across trials. We modified the bead inference task of Lu et al. to be more child-friendly and focused on how efficiently children would adapt their strategy to various settings of evidence strength and information cost [22]. We found that autistic children had lower sampling efficiency (i.e., lower gains) under costed conditions, due to more varied sampling choices across trials rather than systematic deviation from the optimal number of samples. Computational modeling further indicates that autistic children's lower efficiency and greater sampling variation are due to their reduced reliance on both the information from previous trials and the cost accumulated during sampling, along with a stronger influence of recently obtained information within a trial.

Methods

Participants

We collected data from 30 autistic children and 47 neurotypical children with ages ranging from five to eight years old. The sample size was determined by both resource limits and the typical sample size of related decision-making and learning studies in ASD [23] (see Limitations section for discussion). All the neurotypical children came from regular elementary schools. Twenty-six of the autistic children were from a special education school for autistic children in two waves and four autistic children were recruited via advertising posters in the

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

local community. Parents of neurotypical children reported no concerns of any developmental condition. All the autistic children had received a formal diagnosis of an autism spectrum condition from pediatrics based on Diagnostic and Statistical Manual - 5 criteria [24]. As the autistic children were recruited across multiple waves and sites, fifteen autistic children in our final sample received additional assessment by the Autism Diagnostic Observation Schedule (ADOS; [25]), five by both the Childhood Autism Rating Scale (CARS; [26]) and the Autism Spectrum Quotient—Children's Version (AQ-Child; [27]), and four only by AQ-Child (see Table 1). Nevertheless, the inclusion criteria were not solely based on meeting cut-offs of these instruments, as clinical judgment by professional clinicians or pediatrics has been shown to enhance the diagnosis stability [28]. The study abided by the Declaration of Helsinki and was approved by the Institutional Review Board at Peking University, and the study obtained oral consents from all children and written consents from their parents.

Table 1. Children demographic information

	ASD (N = 24)[1]	NT (N = 41)[1]	*p* value[2]
Gender			0.5
Male	20 (83%)	30 (73%)	
Female	4 (17%)	11 (27%)	
Age	6.3 (0.8)	6.5 (0.7)	0.5
Full-scale IQ	104.8 (16.9)	111.4 (14.8)	0.12
ADOS (raw score)	13.3 (4.2) [15]		
ADOS-SA	11.3 (3.2) [15]		
ADOS-RRB	2.0 (1.3)		

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	ASD (N = 24)[1]	NT (N = 41)[1]	*p* value[2]
	[15]		
CARS	34.4 (4.5)		
	[5]		
AQ-Child	80.9 (22.2)	60.5 (12.1)	0.026
	[9]	[34]	

[1]n (%); Mean (SD), [N] the number of valid participants when not all participants were measured
[2]Pearson's Chi-squared test; Welch Two Sample *t* test
Note. ASD: autism spectrum disorder; NT: neurotypical; ADOS: Autism Diagnostic Observation Schedule; SA: social affect; RRB: restricted & repetitive behaviors; CARS: Childhood Autism Rating Scale; AQ-Child: Autism Spectrum Quotient—Children's Version

Five autistic children did not understand the instructions to pass the practice trials, and thus they did not take test trials. Because of emotional distress, one autistic child quitted the experiment after completing half of the test trials, whose data were included in behavioral analyses but not in computational modeling. Six neurotypical children were not included for further analyses because of either not understanding the instruction, inattention to the task, or the experiment program crash. The final sample consists of 24 autistic children and 41 neurotypical children with two groups matched on ages and full-scale IQ (see Table 1 for detailed information). Full-scale IQ was measured by the Chinese version of the Wechsler scales [29, 30].

Stimuli and Procedure

The task was based on the "bead task", a classic instrumental information sampling task [31, 32], and adapted as an adventure game to be of interest to the children (Figure 1a). On each trial, children saw two isles on both sides of the monitor screen. Two isles were

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

habited by ten animals with the opposite ratio of dogs to cats, either 60% to 40% vs. 40% to 60% in a block of trials or 80% to 20% vs. 20% to 80%. Children were told that the goal on each trial was to find out on which isle they were on, based on the animals they had met. They were free to encounter 20 animals at most by pressing the button on an Xbox controller, and when they felt confident enough to make the decision, they could press the corresponding buttons to choose the left or right isles shown on the screen. They would receive 100 credits for correct responses and zero for incorrect responses. In some conditions of the experiment, it would cost children 0, 1, or 4 credits for each animal, in which case the actual gain for a correct response was 100 credits deducted by the cumulative cost incurred for finding animals; an incorrect response would still yield zero credit. At the end of the experiment, the experimenter would give children several of their favorite stickers in proportion to their accumulated credits across trials. Children were encouraged to sample (i.e., choose to encounter animals) wisely to maximize their game credits and final rewards, which would require proper trading-off between information gain and cost during the task (Figure 1b).

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

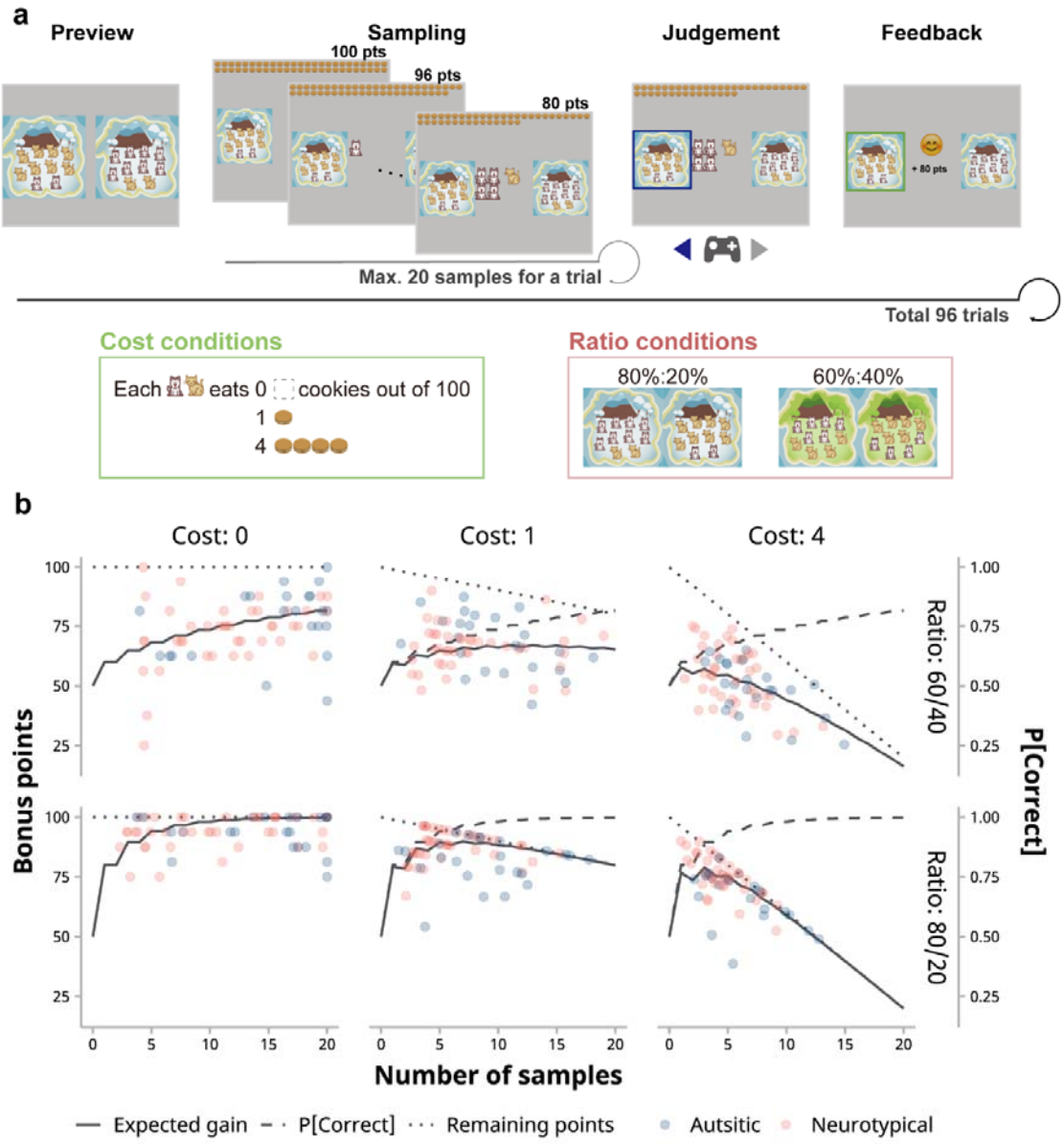


Figure 1. Information sampling “bead” task. (a) The child-friendly "bead task" substituted beads with cat and dog images to enhance engagement. Each block began with a preview of the cost and ratio conditions for subsequent trials ("Preview"). During a trial, children might sample up to 20 animals from a predetermined isle by pressing a button on the gamepad ("Sampling"). When children wanted to make a judgement, they could choose by pressing the left or right button on the gamepad ("Judgement"). The task was structured into six blocks, each a combination of three cost conditions and two ratio conditions. (b) Children typically faced a trade-off between information gain and cost: the expected probability of a correct judgement (dashed lines) increases, while the remaining bonus points (dotted lines) decrease as a function of the number of samples. Consequently, the expected gain, a product of these factors, initially increases but then decreases (as shown by solid lines, especially

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

noticeable in the "Cost: 1" and "Cost: 4" panels). Each point denotes the actual points won by each child averaged across trials.

The experiment consisted of three blocks corresponding to three cost conditions (Figure 1a; zero-cost: 0, low-cost: 1, or high-cost: 4 credits). The order of three blocks was counterbalanced across children. Two mini-blocks of different dog-to-cat ratios (low-evidence: 60%:40% and high-evidence: 80%:20%, respectively) were nested under each block in a random order. For most of the children, each mini block had 16 trials, so the experiment had 96 trials in total. However, because some autistic children were more prone to feel upset during the long task, six autistic children participated in a shorter version of the experiment with 8 trials in each mini block (i.e., 48 trials in total).

To ensure children had requisite understanding of the task, the experimenter asked a series of questions about the ratios, costs, and rewards after the introduction (see Additional File 1, S1 Text for the task instruction and comprehension questions). Only after correctly answering all the questions could children enter the practice task. The practice task was a mock test having only 12 trials but the same structure as the main task. To pass the practice, children had to be "theoretically correct" in more than 75% of the trials: being "theoretically correct" was to choose the isle where the majority of animals was matched with the majority of encountered animals, regardless of the actual outcome. For example, if children have met more dogs than cats, they should choose the dog isle because of higher likelihood of being correct. If children failed to pass the practice, the experimenter would repeat the instruction and the practice task for them. However, if children failed the practice again, they were not eligible to start the main task.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The stimuli were presented on a 24-inch monitor during the experiment via Psychtoolbox in MATLAB 2016b. Children sat about 60 cm in front of the monitor, holding an Xbox controller. We chose dog and cat icons that were similar in style and the same in size, average brightness, and contrast to mitigate the choice bias caused by familiarity or visual salience. To ease the difficulty of memorizing all related information for children, the ratio and cost information would be displayed before the beginning of the next mini-block, and the experimenter would repeat the information to them as well. Additionally, the remaining credits and encountered animals would always be shown during a trial so that children could know their accumulated costs and evidence at any time.

Statistical analyses

All statistical analyses were conducted in R 4.3.2 [33]. We used linear mixed models for model-free analyses to account for the imbalanced data resulting from the unequal numbers of children and trials. Linear mixed models were conducted on the trial level, which utilized more information of the data than condition average-based repeated measures ANOVA. Linear mixed models were estimated using "afex" package [34], and *F* statistics, degrees of freedom of residuals (denominators), and *p*-values of the models were approximated by Kenward-Roger method [35, 36]. Specifications of random effects followed parsimonious modeling [37]. For significant fixed effects, "emmeans" package was used to test follow-up simple effects and post hoc contrasts [38]. Statistical multiplicity of the post hoc contrasts was controlled by a single-step adjustment that used multivariate *t* distributions to estimate the critical value for the contrasts [39].

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Specifications of (generalized) linear mixed models for behavioral data are set as followed (using Wilkinson notation, [40]; $*$ represents main effects of predictors and interaction effects between all predictors plus all lower-order terms):

LMM1: $\text{Game credits}\sim\text{Group} * \text{Cost} * \text{Ratio} + (\text{Cost} * \text{Ratio}|\text{Participant})$

LMM2: $\text{Judgement}\,[\text{True/False}]\sim\text{Group} * \text{Cost} * \text{Ratio} + (\text{Cost} * \text{Ratio}|\text{Participant})$

LMM3: $\text{Efficieny}\sim\text{Group} * \text{Cost} * \text{Ratio} + (\text{Cost} * \text{Ratio}|\text{Participant})$

LMM4: $\text{Signed sampling deviation}\sim\text{Group} * \text{Cost} * \text{Ratio} + (\text{Cost} * \text{Ratio}|\text{Participant})$

LMM5: $\text{Sampling variation}\sim\text{Group} * \text{Cost} * \text{Ratio} + (\text{Cost} + \text{Ratio}|\text{Participant})$

Efficiency in LMM3 is the same as in Lu et al., which is the expected gain for children's animal sample sizes divided by the maximum expected gain given the sampling cost and ratio condition [22]. Suppose the animal sample size is n, the maximal reward is 100 credits, the unit sample cost is c, and the percentage of predominated animals in the preselected isle is q, then the expected gain is $\mathbb{E}[\text{Gain}|n,q,c] = (100 - nc)\,p(n|q)$. $p(n|q)$ is the expected probability of correct judgment, defined as follows:

$$p(n|q) = \begin{cases} \sum_{i=\frac{n+1}{2}}^{n} \binom{n}{i} q^i (1-q)^{n-i} & n = 1,3,5,\ldots,19 \\ \frac{1}{2}\binom{n}{\frac{n}{2}} q^{\frac{n}{2}} (1-q)^{\frac{n}{2}} + \sum_{i=\frac{n}{2}}^{n} q^i (1-q)^{n-i} & n = 2,4,6,\ldots,20 \\ \frac{1}{2} & n = 0 \end{cases} \quad (1)$$

Thus, the optimal sample size is the value of *n* that maximizes $\mathbb{E}[\text{Gain}|n,q,c]$. To investigate the sampling strategy that might result in inefficient sampling, we further calculate signed sampling deviation (LMM4), which is the signed sample size difference from the

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

optimal sample size, and sampling variation (LMM5), which captures the inter-trial standard deviation of sample sizes in a condition.

Behavioral modeling

To account for the difference in the cognitive mechanism behind the sampling decision of children, we adopted six computational models described in Lu et al. [22]. These six models could be divided into two one-stage models and four two-stage models (see Figure 3a). In short, one-stage models assumed the decision of stop sampling was determined by the linear combination of decision variables (such as cost-related and evidence-related variables) at once, whereas for two-stage models, the decision of stop sampling was first determined by a part of decision variables and probabilistically given a second thought to consider other variables. In one-stage models, the probability of stopping sampling on the *i*-th trial after having drawn *j* samples is as a function of the linear combination of *K* decision variables (DVs) via a logistic function:

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

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

Four one-stage models differed in the combination of decision variables:

1. Cost-only model (denoted $\mathrm{Cost\ only}$): cost-related variables, including unit cost per animal sample (categorical: 0, 1, 4; 0 was set as the reference level), number of animals sampled, and total sampling cost (product of the former two terms),
2. Cost and evidence with decay model (denoted $\mathrm{Cost + Evidence}$): both cost-related and decayed evidence-related variables. On top of $\mathrm{Cost\ only}$ model,

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$\mathrm{Cost}+\mathrm{Evidence}$ further includes unit log evidence per animal sample (i.e., $\ln\frac{80}{20}$ or $\ln\frac{60}{40}$), absolute value of cumulative information (cumulative information refers to the difference between the numbers of "cat" and "dog" samples), total log evidence (product of the former two terms), and the correctness and the number of animal samples in the last trial.

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

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

A value of the decay parameter near 0 indicates a tendency to consider only the most recently obtained information, whereas a value closer to 1 would mean considering all evidence with nearly equal weight.

In two-stage models, sampling choices may involve two decision stages, with the probability of reaching the decision of stopping sampling in each stage being

$$p_{ij}^{Stage1}=\frac{1}{1+e^{-X_{ij}^{Stage1}}}, \tag{5}$$

$$p_{ij}^{Stage2}=\frac{1}{1+e^{-X_{ij}^{Stage2}}}. \tag{6}$$

The overall probability of stopping sampling can be written as

$$p_{ij}=p_{ij}^{Stage1}+(1-p_{ij}^{Stage1})\,p_{ij}^{\mathrm{sec}}\,p_{ij}^{Stage2}. \tag{7}$$

Four second-stage models differed in both the combination of decision variables and the second-thought probability, including:

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3. $\text{Cost} \xrightarrow[\text{C-cond}]{} \text{Evidence}$: the first stage and the second-thought probability were both determined by cost-related variables, whereas the second stage was determined by evidence-related variables,
4. $\text{Cost} \xrightarrow[\text{E-cond}]{} \text{Evidence}$: the first stage was determined by cost-related variables, but the second-thought probability was controlled by evidence conditions,
5. $\text{Evidence} \xrightarrow[\text{C-cond}]{} \text{Cost}$: the first stage was determined by evidence-related variables, but the second-thought probability was controlled by cost conditions,
6. $\text{Evidence} \xrightarrow[\text{E-cond}]{} \text{Cost}$: the first stage and the second-thought probability were determined by evidence-related variables.

When the second-thought probabilities were conditional on cost conditions, $p_{ij}^{\text{sec}} = p^{\text{zero cost}}$, $p^{\text{low cost}}$, and $p^{\text{high cost}}$ respectively for the zero-, low-, and high-cost conditions; when conditional on evidence conditions, $p_{ij}^{\text{sec}} = p^{\text{low evidence}}$ and $p^{\text{high evidence}}$ respectively for the low- and high-evidence conditions. In general, model parameters to be estimated included β_k that represents the effects of decision variables, decay parameter α_{Decay}, and second-thought probabilities p_{ij}^{sec} if two-stage models (see Addition File 1, S1 Table).

Model fitting and comparison

We fitted models separately to the behavioral data of neurotypical and autistic children, using hierarchical Bayesian estimation with Hamiltonian Monte Carlo implemented in Stan and cmdstanr package in R [41, 42]. In contrast to Lu et al. using maximum likelihood estimation, the hierarchical Bayesian estimation allows incorporation of prior information and sharing information across participants, which is especially beneficial for relatively small sample sizes and provide more informative and robust results.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Each of four separate Markov chains with randomized initial values took an adaptive warm-up with 5000 samples from the posterior to prevent the dependence on the initial values, and then each chain took another 10000 samples. This resulted in 40000 samples from the posterior for each parameter. For complex hierarchical models, in accordance with the standard practice in Stan, parameters were sampled using non-centered parameterization; that is, parameters were independently sampled from a standard normal distribution before transformed to an appropriate range. Convergence between chains was confirmed based on Gelman-Rubin $\widehat{R}$ of all parameters being less than 1.01 and no systematic divergent transitions. All individual-level parameters were assumed to sampled from the corresponding group-level normal distributions, whose mean and standard deviation were estimated from the data (parameter specifications see Additional File 1, S1 Text).

Model comparison was performed based on the pointwise out-of-sample prediction accuracy of the model, which was evaluated by the Bayesian leave-one-out cross validation estimate of expected log predictive density (LOO ELPD) and implemented in "loo" package in R [43, 44]. Moreover, models could be compared in terms of the model weights from pseudo-Bayesian model averaging with Bayesian bootstrapping (Pseudo-BMA+; [45]). The model weights were calculated taking into account the penalty for model complexity. Models with higher ELPD or Pseudo-BMA+ model wights suggest better performance while having a parsimonious model specification. All models were compared within each group of children (i.e., in autistic and neurotypical children separately).

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Results

Autistic children differed from neurotypical children in sampling strategy rather than correct judgement

Overall, autistic children took fewer samples than neurotypical children overall (Figure 2a; M_{ASD} = 10.3, M_{NT} = 7.6, $F(1, 63.17) = 11.03$, $p < .001$), but their gain was not significantly fewer than neurotypical (NT) children (LMM1; M_{ASD} = 72.4, M_{NT} = 74.9, $F(1, 63.32) = 3.90$, $p = .053$). However, when we examined different conditions separately, we found that autistic children had won significantly fewer credits in the high-cost condition (M_{ASD} = 57.4, M_{NT} = 64.7, $t(66.6) = –3.14$, $p = .007$) and in the high-evidence condition (M_{ASD} = 80.3, M_{NT} = 85.7, $t(182.7) = –4.00$, $p < .001$). An analysis of the proportion of correct judgment allows us to exclude the possibility that autistic children made more incorrect judgement. There was no significant difference between the two groups in the proportion of correct judgement (LMM2; $\chi^2(1) = 0.33$, $p = .57$). Though a significant interaction effect was found between the group and the evidence conditions (LMM2; $\chi^2(1) = 5.08$, $p = .024$), post hoc comparisons did not support significant group differences under any of the two evidence conditions (low-evidence condition OR = 1.16, 95% CI: [0.92, 1.47], $p = .28$; high-evidence condition OR = 0.73, 95% CI: [0.46, 1.18], $p = .26$). When comparing two groups in terms of the gained credits in those correct trials, we found that autistic children won fewer credits particularly in those costly trials (interaction between groups and cost conditions: $F(2, 67.30) = 4.07$, $p = .021$; low-cost condition: $t(63.4) = -2.58$, $p = .033$; high-cost condition: $t(63.2) = -2.68$, $p = .026$). These results suggest that the main behavioral difference was pertinent to their sampling strategy rather than judgement.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Sampling behaviors: Autistic children had significantly lower efficiency than neurotypical children

An adaptive sampling strategy was used to reach the optimal decision by simultaneously considering and balancing the sampling costs and the information gained from sampling (see Figure 1b). We adopted the sampling decision efficiency to measure the optimality of children's information sampling behavior (LMM3). Efficiency was calculated as the expected gain for the number of samples taken by participants divided by the maximum expected gain. First, we looked at the effects of experiment conditions and found a significant interaction between sampling cost and evidence (Figure 2b; $F(2, 65.02) = 30.77$, $p < .001$). Both autistic and NT children had higher sampling efficiency in high-evidence conditions than low-evidence conditions when sampling was costless ($t(63.7) = 6.63$, $p < .001$), whereas children had better performance in the low-evidence conditions than the high-evidence conditions in those low-cost trials ($t(64.6) = -3.07$, $p = .010$).

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

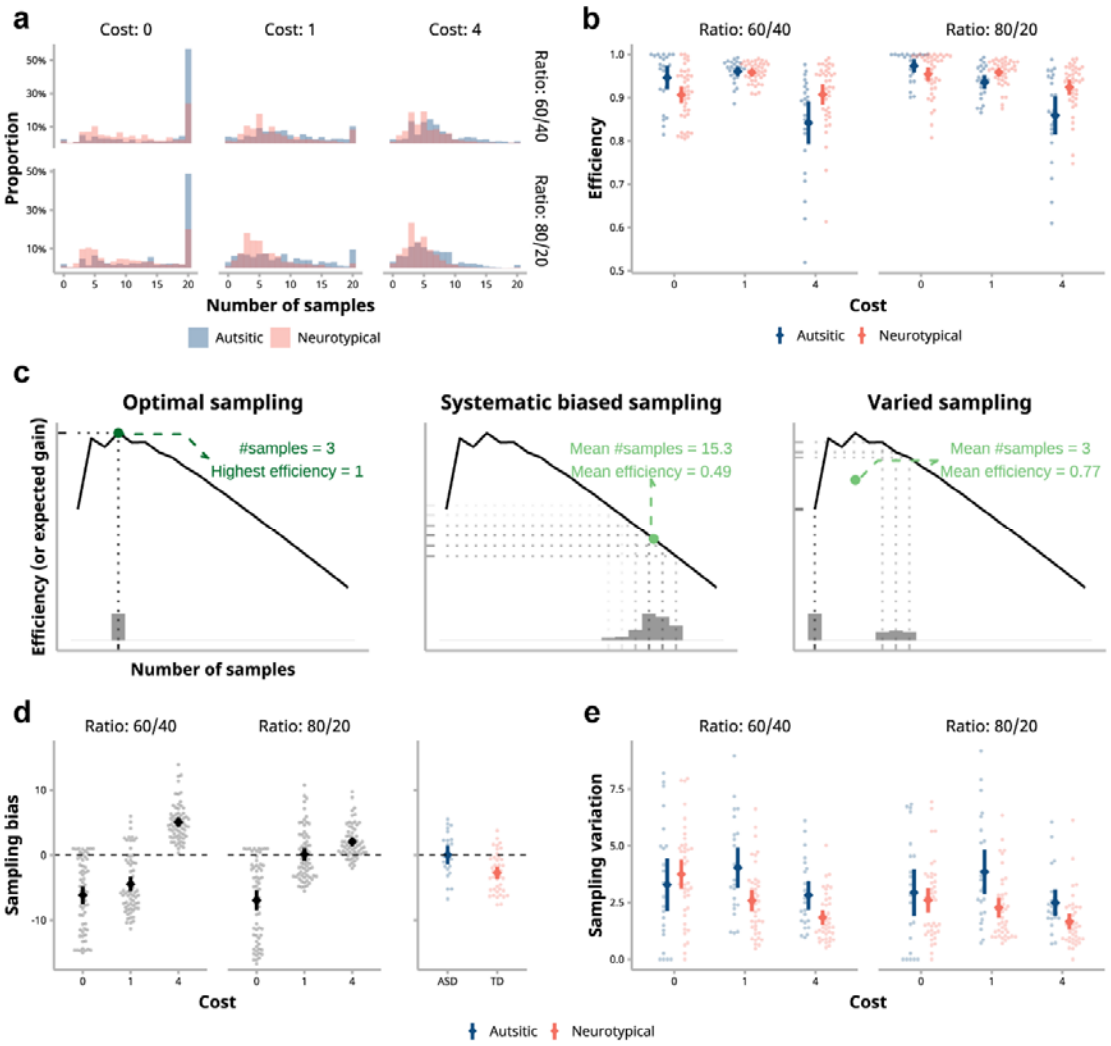


Figure 2. Sampling performance of autistic and neurotypical children**.** (a) Distributions of the average number of samples of each child. Children took fewer samples as the sampling cost or sample ratio increased. Autistic children in general sampled more than neurotypical (NT) children. Histograms of two groups are overlaid with one another. (b) Sampling efficiency is the ratio of expected gain from a given number of samples to the maximum expected gain from the optimal sample number. The sampling efficiency of autistic children was significantly lower than that of NT children, and the difference was even larger in the high-cost condition. (c) When children could stick to the (near-)optimal sample (three samples, in the case of the figure), they would achieve the highest efficiency (left panel). However, both systematic bias and large variation in sampling would lead to poorer efficiency (middle and right panels). Notice that even when the mean number of samples equates to the optimal, that individuals sometimes take too few samples and at other times too many would nevertheless reduce the efficiency (i.e., varied sampling, right panel). Solid lines are the expected gain as the function of the number of samples, illustrating the case under the high-cost and high-evidence condition. Histograms are example distributions of samples. (d) The sampling bias was

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

382 modulated by sampling cost and evidence. Autistic children did not show systematic bias,
383 whereas NT children showed undersampling behavior in general. The difference between the
384 two groups was significant. (e) The sampling variation of autistic children was significantly
385 greater than that of NT children and was also moderated by the amount of costs. Under costly
386 conditions, the sampling variation of the autistic children was greater. In (b), (d), and (e), large
387 points and error bars are means and 95% confidence intervals, and each small point
388 represent the mean of each child.

389

390 Overall, the sampling efficiency of autistic children was marginally significantly lower
391 than that of NT children (M_{ASD} = 92.0%, M_{NT} = 93.5%; F(1, 63.26) = 3.96, p = .051). The
392 significant interaction between group and cost suggests that the group difference increased
393 with increasing cost (Figure 2b; F(2, 63.35) = 6.65, p = .002): there was no significant
394 difference between the two groups in zero- and low-cost trials (t(63.3) = 2.26, p = .75, and
395 t(62.9) = -1.55, p = 0.31 respectively), whereas in high-cost trials, the sampling efficiency of
396 autistic children was lower than that of NT children (t(63.0) = -3.17, p = .007). We also found a
397 significant interaction between group and evidence level (F(1, 63.50) = 4.28, p = .043) that
398 autistic children had a significantly lower efficiency than NT children in high-evidence
399 conditions (high-evidence: t(63.2) = -2.63, p = .019; low-evidence: t(63.5) = -0.97, p = .50).

400 **Sampling behaviors: Autistic children's lower sampling efficiency came from their**
401 **higher sampling variation**

402 In the above analysis of sampling efficiency, we found that the differences between
403 the two groups were modulated by sampling cost: the higher the cost, the worse autistic
404 children performed relative to NT children (Figure 2b). What erred in autistic children's
405 sampling strategy? Greater systematic biases (i.e., oversampling or undersampling) was one
406 possibility, but not the only one. According to our simulations, higher sampling variation may

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

also lead to lower efficiency (see Figure 2c). Or, both bias and variation may play their roles. To elucidate these different possibilities, we decomposed children's sampling behaviors into biases from optimality (Figure 2d) and variation (Figure 2e).

In general, children exhibited both oversampling and undersampling behavior under different conditions (LMM4; Cost x Ratio interaction: $F(2, 62.89) = 250.12$, $p < .001$): for low-evidence conditions, both groups undersampled in zero- and low-cost trials while oversampled in high-cost trials. For high-evidence conditions, children undersampled in zero-cost trials, while they oversampled in high-cost trials; in the low-cost trials, the number of samples taken by both groups was close to optimal (Figure 2d). Autistic children did not exhibit a systematic sampling bias and their signed sampling bias from the optimal was not significantly different from zero ($M = -0.004$, $t(63.6) = -0.006$, $p = 0.99$). In contrast, the NT group exhibited significant undersampling, with a sampling bias significantly less than zero ($M = -2.72$, $t(62.4) = -5.49$, $p < .001$). The difference of sampling bias between the two groups was significant ($F(1, 63.12) = 11.03$, $p = .001$). No significant interaction effects involved groups.

The analysis of sampling variation revealed a significant main effect of group, with autistic children showing greater variation in the number of samples taken across trials (Figure 2e; LMM5: $F(1, 63.70) = 6.87$, $p = .011$). There was also a significant interaction between group and sampling cost ($F(2, 63.38) = 7.73$, $p = .007$), with the variation of the autistic group being particularly high in low- and high-cost conditions compared to the NT group ($t(64.2) = 3.72$, $p = .001$ and $t(75.4) = 2.96$, $p = .012$, respectively). No other interaction effects including group predictor were significant. Considering both sampling bias and

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

sampling variation, it appears that the high variation in sampling across trials in costly conditions may be a major cause of the lower sampling performance observed in autistic children in those conditions.

Computational modeling: Autistic children were less influenced by dynamic and global information when making sampling decisions

To understand the cognitive mechanisms behind sampling decisions in autistic and neurotypical children, we applied models from Lu et al. [22], which were four first-stage models and four second-stage models (Figure 3a). The first-stage models described what information children integrated to achieve the sampling decision. The second-stage models further distinguished how they made the decision in two steps, with different information considered in each step. The information that the children were assumed to integrate could be categorized into cost-related (e.g., cumulative costs during sampling) and evidence-related decision variables (e.g., cumulative log evidence; Figure 3b). Using hierarchical Bayesian models, posterior probability distributions for each parameter for each child at the individual level and the entire group at the population level were obtained simultaneously. Model comparisons were performed based on leave-one-out expected log predictive density (LOO ELPD) and pseudo-Bayesian model averaging weights using resampling method (Pseudo-BMA+; see Methods for details).

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

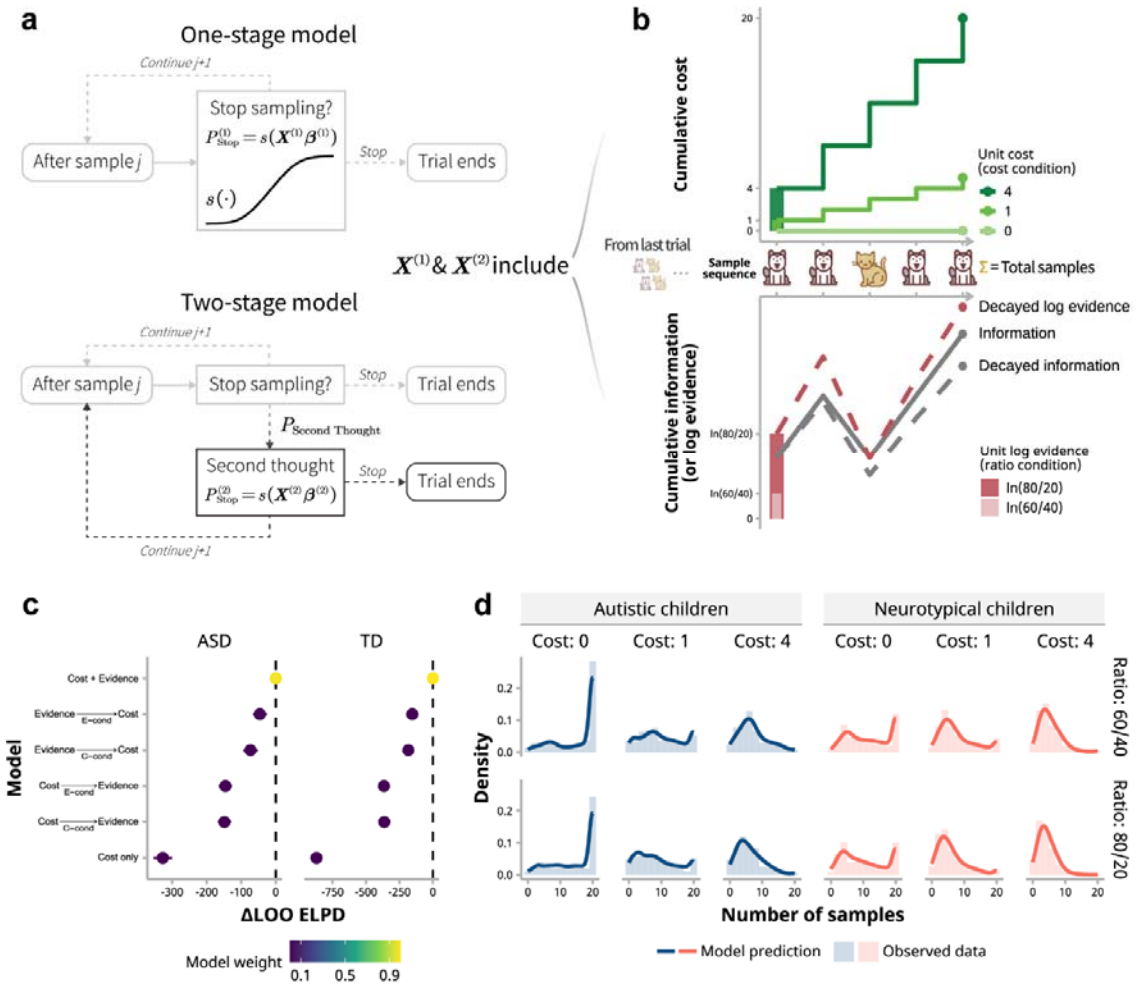


Figure 3. Computational modeling of sampling choices in autistic and neurotypical children. (a) Diagram of one-stage and two-stage models. The one-stage model consists of the steps shown on the top: Each time a child opts to stop sampling, the probability of doing so is determined by a sigmoid function of a linear combination of various decision variables. Additionally, the two-stage model includes a potential second stage (on the bottom), where children might re-evaluate their decision to stop sampling via the second-thought probability, and the second decision is influenced by a different set of decision variables. (b) Decision variables are categorized into cost-related and evidence-related groups. Cost-related variables encompass unit cost, total samples, and cumulative cost (primarily in the upper sub-panel), whereas evidence-related variables consist of unit log evidence, total decayed information, cumulative decayed log evidence, along with the sample number and correctness of the last trial (primarily in the lower sub-panel). (c) LOO ELPD (leave-one-out expected log predictive density) signifies the model evidence, with higher values indicating a better fit to the data. ΔLOO ELPD represents the difference in LOO ELPD relative to the best model (i.e., Cost + Evidence), where a value nearer to zero suggests a closer alignment with

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the best model. Error bars denote the standard errors of the LOO ELPD differences. Point colors reflect model weights assigned using the Pseudo-BMA+ method; heavier weights suggest a better model fit. (d) Comparison between the best model's predictions and the actual data shown as the distribution of sample numbers under various conditions. The model accurately predicts the sampling behavior of both groups, with solid lines depicting observed data and histograms indicating model predictions.

In both groups of children, LOO ELPD and Pseudo-BMA+ model weights indicated that Model $\mathrm{Cost} + \mathrm{Evidence}$ had the highest model evidence and was the best fit for the sampling behavior data based on the number of samples (Figure 3c). In the autistic group, LOO ELPD for Model $\mathrm{Cost} + \mathrm{Evidence}$ was - 4134.68 and Pseudo-BMA+ weight for Model 4 was 0.987; in the NT group, the LOO ELPD was -7777.16 and the Pseudo-BMA+ weight was greater than 0.999. In addition, based on the estimated model and the complete stimulus sequence during the experiment (including unpresented "cat" and "dog" stimuli that would have been shown if children had not stopped sampling), we can predict the sampling choices of children. The posterior predictive checking confirmed that Model $\mathrm{Cost} + \mathrm{Evidence}$ was robust to predict decision choices well for both groups of children (Figure 3d). As the best model $\mathrm{Cost} + \mathrm{Evidence}$ suggested, both the autistic and NT children tended to combine both the cost-related and the evidence-related decision variables all together to determine the sampling choice.

To explore the computational mechanisms that lead to differences in behavioral outcomes between autistic and NT children, the population-level parameters of the winning Model $\mathrm{Cost} + \mathrm{Evidence}$ were compared between the two groups (this method of between-group comparison at the population level is similar to that used in [46]; Figure 4). Results showed that the two groups of children differed in the coefficients of the following

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

decision variables: cost condition differences ($\beta_{H\text{-}0}$ for high-cost relative to zero-cost condition, and $\beta_{L\text{-}0}$ for low- to zero-cost difference), cumulative sampling cost ($\beta_{CumuCost}$), the number of samples drawn in the last trial ($\beta_{LastDraw}$), as well as in the *decay* term of accumulated evidence (α_{Decay}). The rest of the model parameters did not show noticeable group difference (for detailed summary statistics of two groups and the difference see Additional File 1, S2 Table).

In addition to the cost- and evidence-related taxonomy, these model parameters can also be classified as “static” and “dynamic”, by the nature of their coupled decision variables. This classification offers another perspective to understand the cognitive process differences between groups. Static parameters include $\beta_{L\text{-}0}$ and $\beta_{H\text{-}0}$ since they parallel the fixed differences between experimental conditions, which do not vary along with the sampling behavior. In contrast, dynamic parameters, such as $\beta_{CumuCost}$, $\beta_{LastDraw}$, and α_{Decay}, influence sampling decisions dynamically with each sample and trial. For the static parameters, the median group difference $Mdn_{ASD} - Mdn_{NT}$ for $\beta_{H\text{-}0}$ was 0.83, 95% highest density interval (HDI) [0.11, 1.59], and for $\beta_{L\text{-}0}$ it was 0.77, 95% HDI [0.16, 1.41]. For the dynamic parameters, the median differences were $\beta_{LastDraw}$ at 0.54, 95% HDI [0.28, 0.78], $\beta_{CumuCost}$ at -0.48, 95% HDI [-0.95, -0.01], and α_{Decay} at -0.46, 95% HDI [-0.81, -0.04]. Then, we further examine how these static and dynamic parameters affect the sampling behavior of autistic and NT children.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

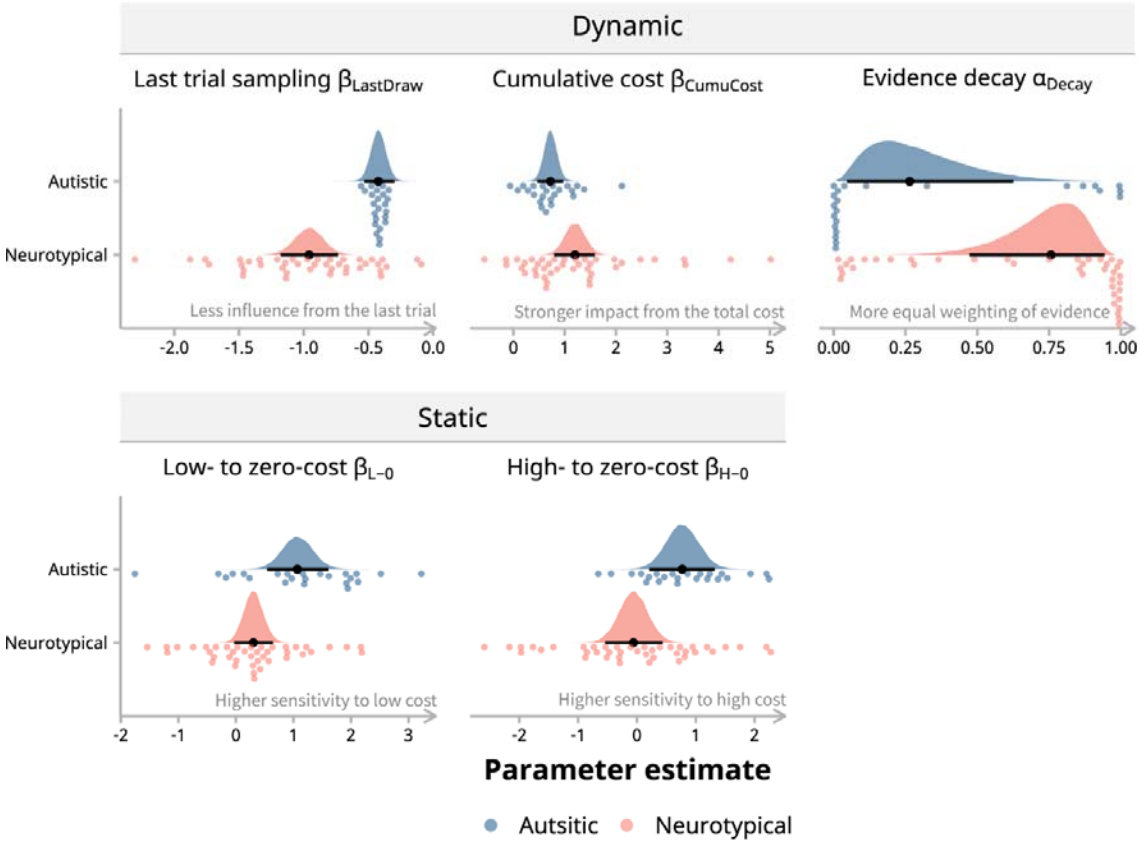


Figure 4. Parameter estimates from the best model in autistic and neurotypical children. Two groups of children showed significant difference in five parameters from the best model. These parameters fall into two categories that reflect dynamic and static influences of decision variables on when stop sampling. The top row ("Dynamic") includes parameters that gauge behavioral adaptability: $\beta_{LastDraw}$ represents the influence from the number of samples of the last trial, $\beta_{CumuCost}$ measures the impact of the total cost accrued during the current trial, and α_{Decay} determines how information gathered during a trial is weighted. These parameters depict a flexible, adaptive process as they change with the actions taken across and within trials. The bottom row ("Static") captures cost sensitivity: β_{L-0} and β_{H-0} indicate the general inclination to stop sampling in costly conditions compared to a zero-cost condition. These two parameters are considered "static" because they remain consistent across trials within each cost context. Overall, compared with neurotypical children, the autistic group demonstrated a reduced adaptation to dynamic shifts within the task but heightened sensitivity to the static factor of "cost as a condition". Data points denote individual parameter estimates, shaded areas represent group-level parameter distributions, and black markers with lines illustrate medians and 95% highest density intervals (HDIs) for these distributions.

Reduced adaption to dynamic information led to increased sampling variation in autistic children

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

First, we focused on the effects of the static parameters, $\beta_{H\text{-}0}$ and $\beta_{L\text{-}0}$, on sampling decisions. The parameters determined the average tendency to stop sampling under different cost conditions, as evidenced by the differences in average sample numbers between these conditions. For the autistic group, both $\beta_{H\text{-}0}$ and $\beta_{L\text{-}0}$ were higher compared to the NT group, indicating a heightened sensitivity to cost of autistic children. Specifically, in costly conditions, autistic children reduced their sampling more significantly (the number of samples: $M_H - M_0$ = -8.33, $t(63.4) = -8.44$, $p < .001$; $M_L - M_0 = -5.64$, $t(62.8) = -6.00$, $p < .001$; see Figure 2a), in contrast to NT children who showed a less pronounced decrease in sampling under the same conditions ($M_H - M_0 = -6.04$, $t(62.1) = -8.15$, $p < .001$; $M_L - M_0 = -4.02$, $t(61.8) = -5.72$, $p < .001$). However, these static parameters alone could not explain why the sampling variation differed between the groups since they predominantly govern the general sampling tendency within a given condition rather than influencing individual trial variations.

Therefore, we turned to the dynamic parameters, $\beta_{LastDraw}$, $\beta_{CumuCost}$, and α_{Decay}. These parameters reflected a more dynamic and flexible cognitive process, because they would affect sampling decisions within each trial based on the previous trial or the sample history within a trial. To better illustrate their influence, we conducted simulations based on the best-fitting model and by varying interested parameters one by one within the range estimated from the children's date while keeping other parameters fixed (Figure 5a-c). For $\beta_{LastDraw}$, autistic children showed values that were larger and closer to zero compared to neurotypical children, suggesting they were less influenced by the sample numbers from the last trial in their current sampling decisions. From the equation of calculating the stopping probability (Equations 1 & 2), a more negative $\beta_{LastDraw}$ value indicates a decreased likelihood

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

549 of stopping sampling as the previous trial's sample number increases, potentially leading to
550 more samples in the current trial. Conversely, a $\beta_{LastDraw}$ value closer to zero means the
551 previous trial's sample count has little to no impact on the current decision, allowing for
552 greater variation in sample numbers from trial to trial, which is consistent with the simulation
553 (Figure 5a). For $\beta_{CumuCost}$, a higher value indicates that children are more likely to stop
554 sampling as the cumulative cost increases. This means that as children encounter higher
555 costs, they become more cautious, reducing the number of samples they take (Figure 5b). In
556 the case of α_{Decay}, a value approaching one suggests that children consider all pieces of
557 evidence more uniformly when deciding to stop sampling. A balanced α_{Decay} (i.e., closer to
558 one) means the decision to sample is less swayed by the recency or initial pieces of evidence,
559 potentially leading to a more consistent number of samples across trials. This is more evident
560 in zero-cost conditions and high-evidence conditions where sampling decisions should not be
561 limited by the cost and the evidence is easier to accumulate in one direction (Figure 5c).

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

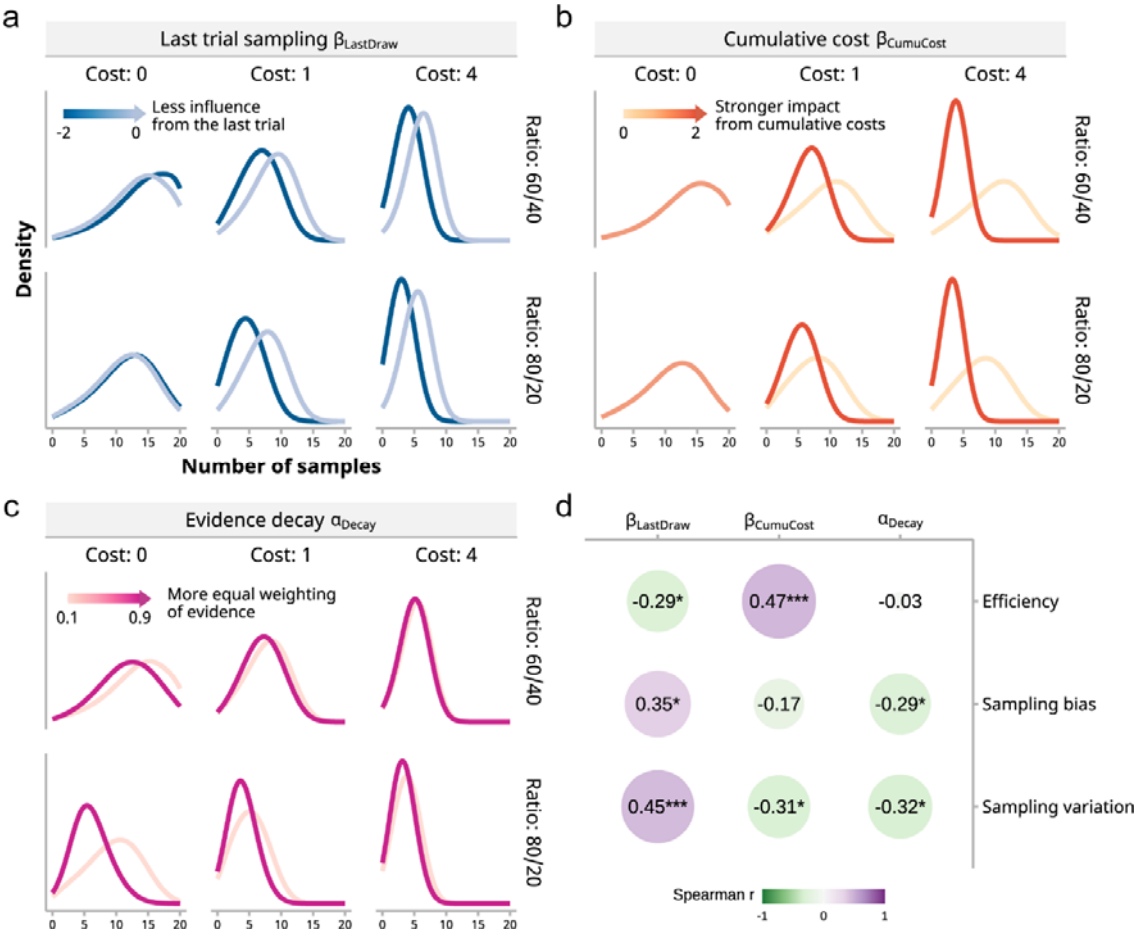


Figure 5. Influence of dynamic decision parameters on sampling behavior. (a) – (b) To illustrate the effect of dynamic parameters, we conducted simulations of children's sampling processes by holding other parameters constant and varying the interested variables individually within the observed range of the children's estimates. The influence of $\beta_{LastDraw}$ and $\beta_{CumuCost}$ was evident: a stronger impact from the last trial's sampling and cumulative costs resulted in fewer samples and reduced variation, particularly at higher costs. (c) Simulations for α_{Decay} indicated that a more equal weighting of evidence led to a reduction in the number of samples, particularly under conditions with more evidence and no cost. (d) Correlation between parameter estimates and behavioral metrics: Spearman correlation coefficients reveal the relationships between dynamic parameters (individual-level estimates of $\beta_{LastDraw}$, $\beta_{CumuCost}$, α_{Decay}) and behavioral metrics (efficiency, sampling bias, and sampling variation). Significance levels are marked by asterisks (*p < 0.05, ***p < 0.001), with p-values adjusted for false discovery rate (FDR).

Based on the qualitative findings from the simulations, we further conducted correlation analyses between these dynamic parameters and children's sampling behaviors,

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and the results exhibited consistent patterns (Figure 5d). $\beta_{LastDraw}$ showed a significant negative correlation with the average efficiency throughout the experiment (r_S = -.29, adjusted 95% CI [-.51, -.04], p = .028). Such a negative correlation would be even stronger in high-cost conditions (r_S = -.39, adjusted 95% CI [-.58, -.15], p = .002) or in high-evidence conditions (r_S = -.38, adjusted 95% CI [-.58, -.14], p = .006). The direction of these correlations was consistent with the group differences in sampling efficiency and model parameters: autistic children had lower sampling efficiency overall and in high-cost conditions or high-evidence conditions, and their $\beta_{LastDraw}$ was larger than that of NT children. $\beta_{CumuCost}$ was positively correlated with the efficiency (r_S = .47, adjusted 95% CI [.24, .64], p < .001), which mirrored the group difference in both the parameter and overall efficiency. Although α_{Decay} was not correlated with the overall efficiency, but as the simulation predicts, it was negatively correlated with the efficiency under the zero-cost conditions (r_S = -.35, adjusted 95% CI [-.55, -.11], p = .021; see Figure 5c left panels).

The direction of the correlations between these parameters and the sampling bias and the sampling variation was also consistent with the group differences. $\beta_{LastDraw}$ showed a significant positive correlation with the average sampling bias (r_S = .35, adjusted 95% CI [.011, .56], p = .012); and a positive correlation with the average sampling variation (r_S = .45, adjusted 95% CI [.22, .63], p < .001). $\beta_{CumuCost}$ only had a significant correlation with the sampling variation (r_S = -.31, adjusted 95% CI [-.52, -.06], p = .025). α_{Decay} also had significant negative correlations with sampling bias (r_S = -.29, adjusted 95% CI [-.50, -.04], p = .032) and sampling variation (r_S = -.32, adjusted 95% CI [-.53, -.08], p = .013), which was also compatible with the group differences in α_{Decay} (autistic children had lower α_{Decay} than

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

601 neurotypical children) and the measures of sampling behavior. These correlations between
602 the estimated parameters and the behavioral measures, combined with the simulations,
603 suggested that autistic children were less flexible and adaptive to the dynamic changes in the
604 environment but more sensitive to the static and local nature of conditions. Such cognitive
605 characteristics of autistic children further lead to their higher variation in sampling decisions,
606 and ultimately relatively poorer sampling efficiency.
607

608 **Discussion**

609 The current study adapted the classic bead task into a child-friendly version to
610 investigate differences between autistic and neurotypical children in information sampling
611 with explicit costs. Model-free analyses showed that both groups showed suboptimal
612 sampling but for distinct reasons. Neurotypical children tended to undersample overall. In
613 contrast, the autistic group did not exhibit consistent oversampling or undersampling but
614 instead higher trial-to-trial variability in costly conditions, which resulted in lower sampling
615 efficiency compared to their neurotypical peers. Computational modeling further revealed that
616 both autistic and neurotypical children considered cost- and evidence-related factors in their
617 decision to stop sampling. However, the autistic group, compared to the neurotypical children,
618 were less influenced by last trial sample numbers and cumulative costs but favored recent
619 evidence in a trial. Model simulations and correlational analyses confirmed that these factors,
620 particularly the influence of the last trial samples, contributed to the differences in sampling
621 decision efficiency, sampling bias, and variation observed between the two groups.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Experimental conditions in the current study were divided into distinct blocks, where trials within the same block shared the same information gain and cost, though the samples in different trials were independent of each other. This setting should have facilitated children's learning of the optimal policy, but autistic children did not exhibit a stable sampling strategy within these blocks. Our modeling results suggest this could result from autistic children's struggles to effectively use historical and dynamic information to learn and adjust their strategies, coupled with their more weight assigned to recent evidence over earlier evidence in a trial. These findings are reminiscent of the Bayesian account of ASD, including the "hypo-prior" hypothesis [47] and the "high sensory precision" hypothesis [48–50]. According to this framework, autistic people give relatively less weight to prior knowledge or top-down predictions in the inference process, or more weight to sensory inputs, leading to an imbalanced reliance on incoming sensory evidence over prior beliefs during perceptual inference [13, 49, 51] and "continuously high learning rates weakening the influence of past perception on current perception" [14, 50]. From this perspective, our findings of a diminished influence of sample numbers from the previous trial can be regarded as weak prior information for decisions in the current trial, and a stronger weight on recent evidence indicates a higher learning rate in terms of integrating information stream. Our findings also extend this unifying Bayesian theoretical framework for ASD, in the sense that the belief updating in previous research was only about specific percept or value, while ours are also about more abstract sampling strategy.

Further extending these insights, our results also resonate with the Weak Central Coherence (WCC) account of ASD. Traditionally focusing on perceptual processing, WCC

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

theory suggests that autistic individuals tend to focus on details and local information, sometimes at the expense of broader and global context [52–54]. For example, compared to neurotypical individuals, autistic people have shown superior performance in visual search and embedded figures tasks [55], as well as an attenuated McGurk effect, indicative of less integration of visual and auditory cues [56]. Prior research on these theories predominantly addressed perceptual inference [57–66], but both the detail-focused cognitive style and the improper weighting between prior beliefs and new information could manifest in various facets. Our findings align with this detail-oriented cognitive style, where autistic children prioritize recently gathered evidence (i.e., local information) over historical data from previous trials and samples (i.e., global information) in their sampling decisions. Our study expands beyond perception in ASD to delves into information sampling as a form of disambiguatory active inference, which involves individuals actively collecting information from the environment for better inferences in the future [14, 67]. Such actions (i.e., active inference) could link sensory processing differences to a broader range of autistic symptoms, suggesting ASD can be understood as a difference in the balance between perception and action [14]. Thus, our study not only provides new supportive evidence for this perspective but also underscores the importance of extending the Bayesian and WCC accounts of ASD beyond perception to include action. This extension is crucial, given that the core symptoms of ASD are tightly related to the interaction with the world.

Our findings may also shed light on the mixed results found in previous studies on information sampling in ASD. Jänsch & Hare found adults with Asperger syndrome sample less evidence compared to neurotypical adults in a classic “bead task” [19], where

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

participants choose from which bead jar a sequence of beads is more likely to be drawn. Similarly, Farmer et al. found that autistic adults sample less information and make faster decisions, rather than adopting a deliberative sampling style [18]. However, in a similar bead task, Brosnan et al. showed that autistic adolescents gather more beads (i.e., more information) compared to their neurotypical peers [20]. Vella et al. found in a similar box task that autistic participants collect more information to make more correct choices [21]. While the inconsistency might stem from the differences in sample characteristics like age and diagnosis, it is noteworthy that these studies provided limited insights beyond just how much information participants sampled. In particular, the absence of explicit sampling costs or internal cost measurements in these studies makes it hard to assess what role the costs played and how participants responded, because whether oversampling or undersampling is observed depends on the sampling cost, as shown in our results of sampling bias (Figure 2d). Moreover, given our findings that autistic participants base their sampling decisions on a more local context, the seemingly inconsistent observations of oversampling or undersampling in previous studies may arise from subtle differences in their experimental designs.

Age differences also play an important role in information sampling. For example, studies have established distinctive profiles for adolescents, including risk-taking [10, 68–70] and information sampling [2, 5]. These aspects might contribute to the discrepancies observed in Jänsch & Hare and Brosnan et al. [19, 20], where even neurotypical adolescents and adults showed different information sampling behaviors, even under a condition with the same bead ratio and no explicit costs. Comparing our findings with Lu et al. [22], who also

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

688 used a bead task with explicit costs and multiple ratios, we noted similarities and differences
689 in information sampling between children and adults on the autistic spectrum. Both autistic
690 children and adults with more autistic traits had greater sampling variability under costly
691 conditions, leading to the lower efficiency. However, while our modeling results and those of
692 Lu et al. indicate reduced flexibility in adapting to experimental changes in autistic individuals,
693 the underlying cognitive mechanisms differ between children and adults. We found that
694 autistic children had a cognitive process similar to neurotypical children but were more
695 influenced by static and local factors in the experiment. In contrast, Lu et al. showed adults
696 with more autistic traits adhered to a rigid and fixed strategy. Future research should
697 systematically investigate how information sampling behavior in autistic individuals at
698 different developmental stages is regulated by contextual factors and how it develops. For
699 instance, Crawley et al. [71] demonstrated through computational models that individuals at
700 different age stages, whether they had an ASD diagnosis, used different learning strategies in
701 a probabilistic reversal learning task. This suggests that the learning mechanisms may
702 become more complex as cognitive functions develop. Although direct comparisons between
703 our study and Lu et al. could be challenging due to differences in diagnosis and experimental
704 details, both studies collectively provide strong evidence that variations in information
705 sampling under different information gain and cost conditions may emerge early in autistic
706 individuals and even persist across ages.

707 In our research, we focused on one type of information sampling, which is
708 instrumental information sampling that by obtaining more information, the probability of
709 making correct decisions can be increased and higher expected returns can be obtained [72,

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

73]. However, it is also true that the obtained information sometimes cannot change the result of decisions but can only reduce the related uncertainty in advance. This type of information sampling is called non-instrumental information sampling [72, 74–76]. Studies have suggested that the intrinsic value of non-instrumental information may partly be related to ameliorating negative affects caused by uncertainty [74, 77, 78], which makes non-instrumental information highly relevant to autism. Research has found that handling uncertainty about events is challenging for individuals with autism and often leads to anxiety [79–81]. Some theories also suggest that such anxiety may induce common restricted and repetitive behaviors in autism, such as shifting attention to limited interests, sensory-seeking, or repetitive behaviors to regulate and attenuate the anxiety [82]. Therefore, it is crucial for future research to investigate how autism affects non-instrumental information sampling behavior and the relationship between non-instrumental information sampling and repetitive behaviors in autistic people.

Limitations

Though all the autistic children in our study had a formal ASD diagnosis by professional clinicians, we enrolled the participants from different waves and sites, leading to a situation where children, particularly autistic children, underwent additional assessments that served similar functions but were different in the form (e.g., among the 24 autistic children, 15 children had ADOS, 5 children had CARS and AQ-child measures, and the remaining 4 children had AQ-child measures). In addition to the formal ASD diagnosis, those assessed with ADOS or CARS met the respective cutoffs, and those tested with AQ-child exhibited significantly higher scores than the neurotypical children (Table 1). Nevertheless,

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

this inconsistency in assessment tools prevents us from adopting a dimensional approach to correlate our modeling results with autistic symptom severity, as recently advocated in psychiatry [83, 84]. A second limitation is the sample size (i.e., the number of participants). The sample size of our study lines up with those of recent related studies on decision-making, learning, and inference in autistic children or adults [61, 65, 85–93], leading to robust findings that accorded with those found in adults with varying autistic traits [22]. However, if we had had resources to test more participants (the high functional autistic children is relatively rare), we would be able to refine subtyping based on our computational model parameters, which could offer more insights into each individual and ASD per se [94–96]. Future studies could benefit from a larger sample size and more comprehensive assessments to achieve not only greater statistical power but also deeper insights into individual and subgroup differences, hopefully contributing more to precise interventions.

Conclusion

In summary, we show that autistic children perform differently in an information sampling task compared to neurotypical children of the same age and intelligence level. Using computational modeling, we find that autistic children are less affected by dynamic changes between and within trials and more inclined to weigh recent pieces of evidence, resulting in greater behavioral instability and lower sampling efficiency. The information sampling behaviors of autistic children resemble those of adults with more autistic traits, despite the strategic differences probably due to ages. These findings reveal developmental

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

753 differences in a fundamental cognitive function at an early stage, and may provide empirical
754 support for and even expand the Bayesian framework of ASD.

755

756 **Declarations**

757 **Ethics approval and consent to participate**

758 The Institutional Review Board of the School of Psychological and Cognitive Sciences
759 at Peking University approved the study, and all the child participants and their parents
760 provided informed consent according to the Declaration of Helsinki.

761 **Consent for publication**

762 Not applicable

763 **Availability of data and materials**

764 The codes and datasets generated and analyzed during the current study are
765 available from the corresponding author on reasonable request and in accordance with
766 ethical procedures for the reuse of sensitive data.

767 **Competing interests**

768 The authors declare that they have no competing interests.

769 **Funding**

770 HZ was partly supported by the National Natural Science Foundation of China
771 (32171095), National Science and Technology Innovation 2030 Major Program
772 (2022ZD0204803), and funding from Peking-Tsinghua Center for Life Sciences. LY was
773 supported by National Natural Science Foundation of China (32271116), the Fundamental

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

774 Research Funds for the Central Universities, and Key-Area Research and Development
775 Program of Guangdong Province (2019B030335001).

776 **Authors' contributions**

777 HL collected, analyzed, and interpreted the data for statistical analyses and
778 computational modeling, and was a major contributor in drafting the manuscript. HZ and LY
779 provided supervision and funding acquisition. All authors participated in conceptualization
780 and reviewing and editing the manuscript.

781 **Acknowledgements**

782 We would like to thank all the children and their parents for participating in our study.
783 We are also thankful to Tianbi Li, Yixiao Hu, Zheng Wang, and staff in Qingdao Elim School
784 for their generous assistance for the study.

785

786

787

788

789

790

791

792

793

794

795

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Bazhydai M, Westermann G, Parise E. “I don’t know but I know who to ask”: 12-month-olds actively seek information from knowledgeable adults. Dev Sci. 2020;23:e12938.

2. Bowler A, Habicht J, Moses-Payne ME, Steinbeis N, Moutoussis M, Hauser TU. Children perform extensive information gathering when it is not costly. Cognition. 2021;208:104535.

3. Goupil L, Romand-Monnier M, Kouider S. Infants ask for help when they know they don’t know. Proc Natl Acad Sci. 2016;113:3492–6.

4. Iwasaki S, Kuroshima H, Arahori M, Fujita K. Prospective information-seeking in human children (Homo sapiens): When to seek and what to seek. J Comp Psychol. 2020;134:341–8.

5. Jones PR, Landin L, McLean A, Juni MZ, Maloney LT, Nardini M, et al. Efficient visual information sampling develops late in childhood. J Exp Psychol Gen. 2019;148:1138–52.

6. Meder B, Wu CM, Schulz E, Ruggeri A. Development of directed and random exploration in children. Dev Sci. 2021;24:e13095.

7. Mills CM, Landrum AR. Learning Who Knows What: Children Adjust Their Inquiry to Gather Information from Others. Front Psychol. 2016;7.

8. Pelz MC, Allen KR, Tenenbaum JB, Schulz LE. Foundations of intuitive power analyses in children and adults. Nat Hum Behav. 2022;:1–12.

9. Ruggeri A, Lombrozo T, Griffiths TL, Xu F. Sources of developmental change in the efficiency of information search. Dev Psychol. 2016;52:2159–73.

10. van den Bos W, Hertwig R. Adolescents display distinctive tolerance to ambiguity and to uncertainty during risky decision making. Sci Rep. 2017;7:40962.

11. Lai M-C, Lombardo MV, Baron-Cohen S. Autism. The Lancet. 2014;383:896–910.

12. Lord C, Elsabbagh M, Baird G, Veenstra-Vanderweele J. Autism spectrum disorder. The Lancet. 2018;392:508–20.

13. Palmer CJ, Seth AK, Hohwy J. The felt presence of other minds: Predictive processing, counterfactual predictions, and mentalising in autism. Conscious Cogn. 2015;36:376–89.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

14. Palmer CJ, Lawson RP, Hohwy J. Bayesian approaches to autism: Towards volatility, action, and behavior. Psychol Bull. 2017;143:521–42.

15. Au-Yeung SK, Kaakinen JK, Benson V. Cognitive Perspective-Taking During Scene Perception in Autism Spectrum Disorder: Evidence From Eye Movements. Autism Res. 2014;7:84–93.

16. Król ME, Król M. A novel machine learning analysis of eye-tracking data reveals suboptimal visual information extraction from facial stimuli in individuals with autism. Neuropsychologia. 2019;129:397–406.

17. Song Y, Hakoda Y, Sang B. A selective impairment in extracting fearful information from another's eyes in Autism. Autism Res. 2016;9:1002–11.

18. Farmer GD, Smith P, Baron-Cohen S, Skylark WJ. The effect of autism on information sampling during decision-making: An eye-tracking study. Judgm Decis Mak. 2021;16:24.

19. 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;44:111–9.

20. 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;44:513–20.

21. Vella L, Ring HA, Aitken MR, Watson PC, Presland A, Clare IC. Understanding self-reported difficulties in decision-making by people with autism spectrum disorders. Autism. 2018;22:549–59.

22. Lu H, Yi L, Zhang H. Autistic traits influence the strategic diversity of information sampling: Insights from two-stage decision models. PLOS Comput Biol. 2019;15:e1006964.

23. Lakens D. Sample Size Justification. 2021.

24. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). American Psychiatric Association Publishing; 2013.

25. Lord C, Risi S, Lambrecht L, Cook, Jr. EH, Leventhal BL, DiLavore PC, et al. The Autism Diagnostic Observation Schedule–Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. J Autism Dev Disord. 2000;30:205–23.

26. Schopler E, Reichler RJ, DeVellis RF, Daly K. Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). J Autism Dev Disord. 1980;10:91–103.

27. Auyeung B, Baron-Cohen S, Wheelwright S, Allison C. The Autism Spectrum Quotient: Children's Version (AQ-Child). J Autism Dev Disord. 2007;38:1230–40.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

858 28. Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism From 2 to 9 Years of
859 Age. Arch Gen Psychiatry. 2006;63:694–701.

860 29. Wechsler D. Wechsler intelligence scale for children–Fourth edition (WISC-IV). The
861 Psychological Corporation; 2003.

862 30. Wechsler D. Wechsler preschool and primary scale of intelligence–Fourth edition
863 (WPPSI-IV). The Psychological Corporation; 2012.

864 31. Garety PA, Freeman D, Jolley S, Dunn G, Bebbington PE, Fowler DG, et al. Reasoning,
865 Emotions, and Delusional Conviction in Psychosis. J Abnorm Psychol. 2005;114:373–84.

866 32. Huq SF, Garety PA, Hemsley DR. Probabilistic Judgements in Deluded and Non-Deluded
867 Subjects. Q J Exp Psychol Sect A. 1988;40:801–12.

868 33. R Core Team. R: A language and environment for statistical computing. Manual. Vienna,
869 Austria: R Foundation for Statistical Computing; 2023.

870 34. Singmann H, Bolker B, Westfall J, Aust F, Ben-Shachar MS. afex: Analysis of factorial
871 experiments. Manual. 2023.

872 35. Halekoh U, Højsgaard S. A Kenward-Roger Approximation and Parametric Bootstrap
873 Methods for Tests in Linear Mixed Models - The *R* Package **pbkrtest**. J Stat Softw. 2014;59.

874 36. Kenward MG, Roger JH. Small Sample Inference for Fixed Effects from Restricted
875 Maximum Likelihood. Biometrics. 1997;53:983–97.

876 37. Bates D, Kliegl R, Vasishth S, Baayen H. Parsimonious Mixed Models. ArXiv Prepr
877 ArXiv150604967. 2015.

878 38. Lenth RV. emmeans: Estimated marginal means, aka least-squares means. Manual.
879 2023.

880 39. Genz A, Bretz F, Hochberg Y. Approximations to multivariate t integrals with application to
881 multiple comparison procedures. In: Institute of Mathematical Statistics Lecture Notes -
882 Monograph Series. Beachwood, Ohio, USA: Institute of Mathematical Statistics; 2004. p.
883 24–32.

884 40. Wilkinson GN, Rogers CE. Symbolic Description of Factorial Models for Analysis of
885 Variance. J R Stat Soc Ser C Appl Stat. 1973;22:392–9.

886 41. Gabry J, Češnovar R, Johnson A. cmdstanr: R interface to “CmdStan.” Manual. 2022.

887 42. Stan Development Team. Stan Modeling Language Users Guide and Reference Manual.
888 2.34. 2023.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

889 43. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out
890 cross-validation and WAIC. Stat Comput. 2017;27:1413–32.

891 44. Vehtari A, Gabry J, Magnusson M, Yao Y, Bürkner P-C, Paananen T, et al. loo: Efficient
892 leave-one-out cross-validation and WAIC for Bayesian models. Manual. 2023.

893 45. Yao Y, Vehtari A, Simpson D, Gelman A. Using stacking to average Bayesian predictive
894 distributions. 2018.

895 46. Ahn W-Y, Vasilev G, Lee S-H, Busemeyer JR, Kruschke JK, Bechara A, et al.
896 Decision-making in stimulant and opiate addicts in protracted abstinence: evidence from
897 computational modeling with pure users. Front Psychol. 2014;5.

898 47. Pellicano E, Burr D. When the world becomes 'too real': a Bayesian explanation of autistic
899 perception. Trends Cogn Sci. 2012;16:504–10.

900 48. Brock J. Alternative Bayesian accounts of autistic perception: comment on Pellicano and
901 Burr. Trends Cogn Sci. 2012;16:573–4.

902 49. Lawson RP, Rees G, Friston KJ. An aberrant precision account of autism. Front Hum
903 Neurosci. 2014;8.

904 50. Van de Cruys S, Evers K, van der Hallen R, van Eylen L, Boets B, de-Wit L, et al. Precise
905 minds in uncertain worlds: Predictive coding in autism. Psychol Rev. 2014;121:649–75.

906 51. van Boxtel JJA, Lu H. A predictive coding perspective on autism spectrum disorders.
907 Front Psychol. 2013;4.

908 52. Happé F. Autism: cognitive deficit or cognitive style? Trends Cogn Sci. 1999;3:216–22.

909 53. Happé F, Frith U. The Weak Coherence Account: Detail-focused Cognitive Style in Autism
910 Spectrum Disorders. J Autism Dev Disord. 2006;36:5–25.

911 54. Happé F. Studying Weak Central Coherence at Low Levels: Children with Autism do not
912 Succumb to Visual Illusions. A Research Note. J Child Psychol Psychiatry. 1996;37:873–7.

913 55. Mottron L, Burack JA, Iarocci G, Belleville S, Enns JT. Locally oriented perception with
914 intact global processing among adolescents with high-functioning autism: evidence from
915 multiple paradigms. J Child Psychol Psychiatry. 2003;44:904–13.

916 56. Mottron L, Peretz I, Ménard E. Local and Global Processing of Music in High-functioning
917 Persons with Autism: Beyond Central Coherence? J Child Psychol Psychiatry.
918 2000;41:1057–65.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

57. Coll M-P, Whelan E, Catmur C, Bird G. Autistic traits are associated with atypical precision-weighted integration of top-down and bottom-up neural signals. Cognition. 2020;199:104236.

58. Croydon A, Karaminis T, Neil L, Burr D, Pellicano E. The light-from-above prior is intact in autistic children. J Exp Child Psychol. 2017;161:113–25.

59. Ewbank MP, Rhodes G, Hagen VD, H EA, Powell TE, Bright N, et al. Repetition Suppression in Ventral Visual Cortex Is Diminished as a Function of Increasing Autistic Traits. Cereb Cortex. 2015;25:3381–93.

60. Karvelis P, Seitz AR, Lawrie SM, Seriès P. Autistic traits, but not schizotypy, predict increased weighting of sensory information in Bayesian visual integration. eLife. 2018;7:e34115.

61. Noel J-P, Shivkumar S, Dokka K, Haefner RM, Angelaki DE. Aberrant causal inference and presence of a compensatory mechanism in autism spectrum disorder. eLife. 2022;11:e71866.

62. Powell G, Meredith Z, McMillin R, Freeman TCA. Bayesian Models of Individual Differences: Combining Autistic Traits and Sensory Thresholds to Predict Motion Perception. Psychol Sci. 2016;27:1562–72.

63. Skewes JC, Jegindø E-M, Gebauer L. Perceptual inference and autistic traits. Autism. 2015;19:301–7.

64. Turi M, Karaminis T, Pellicano E, Burr D. No rapid audiovisual recalibration in adults on the autism spectrum. Sci Rep. 2016;6.

65. Van de Cruys S, Lemmens L, Sapey-Triomphe L-A, Chetverikov A, Noens I, Wagemans J. Structural and contextual priors affect visual search in children with and without autism. Autism Res. 2021;14:1484–95.

66. Zaidel A, Goin-Kochel RP, Angelaki DE. Self-motion perception in autism is compromised by visual noise but integrated optimally across multiple senses. Proc Natl Acad Sci. 2015;112:6461–6.

67. Seth AK. The cybernetic bayesian brain. In: Metzinger TK, Windt JM, editors. Open MIND. Frankfurt am Main: MIND Group; 2015.

68. Defoe IN, Dubas JS, Figner B, Van Aken MAG. A meta-analysis on age differences in risky decision making: Adolescents versus children and adults. Psychol Bull. 2015;141:48–84.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

951 69. South M, Dana J, White SE, Crowley MJ. Failure is not an option: Risk-taking is
952 moderated by anxiety and also by cognitive ability in children and adolescents diagnosed with
953 an autism spectrum disorder. J Autism Dev Disord. 2011;41:55–65.

954 70. Steinberg L. A social neuroscience perspective on adolescent risk-taking. Dev Rev.
955 2008;28:78–106.

956 71. Crawley D, Zhang L, Jones EJH, Ahmad J, Oakley B, Cáceres ASJ, et al. Modeling
957 flexible behavior in childhood to adulthood shows age-dependent learning mechanisms and
958 less optimal learning in autism in each age group. PLOS Biol. 2020;18:e3000908.

959 72. Gottlieb J, Oudeyer P-Y, Lopes M, Baranes A. Information-seeking, curiosity, and
960 attention: computational and neural mechanisms. Trends Cogn Sci. 2013;17:585–93.

961 73. Gottlieb J, Oudeyer P-Y. Towards a neuroscience of active sampling and curiosity. Nat
962 Rev Neurosci. 2018;19:758.

963 74. Bennett D, Bode S, Brydevall M, Warren H, Murawski C. Intrinsic Valuation of Information
964 in Decision Making under Uncertainty. PLOS Comput Biol. 2016;12:e1005020.

965 75. Bromberg-Martin ES, Hikosaka O. Midbrain Dopamine Neurons Signal Preference for
966 Advance Information about Upcoming Rewards. Neuron. 2009;63:119–26.

967 76. Charpentier CJ, Bromberg-Martin ES, Sharot T. Valuation of knowledge and ignorance in
968 mesolimbic reward circuitry. Proc Natl Acad Sci. 2018;115:E7255–64.

969 77. Bennett D, Sutcliffe K, Tan NP-J, Smillie LD, Bode S. Anxious and obsessive-compulsive
970 traits are independently associated with valuation of noninstrumental information. J Exp
971 Psychol Gen. 2020;150:739–55.

972 78. Daddaoua N, Lopes M, Gottlieb J. Intrinsically motivated oculomotor exploration guided
973 by uncertainty reduction and conditioned reinforcement in non-human primates. Sci Rep.
974 2016;6:20202.

975 79. Bervoets J, Milton D, Van de Cruys S. Autism and intolerance of uncertainty: an ill-fitting
976 pair. Trends Cogn Sci. 2021. https://doi.org/10.1016/j.tics.2021.08.006.

977 80. Boulter C, Freeston M, South M, Rodgers J. Intolerance of Uncertainty as a Framework
978 for Understanding Anxiety in Children and Adolescents with Autism Spectrum Disorders. J
979 Autism Dev Disord. 2014;44:1391–402.

980 81. Hwang YI (Jane), Arnold S, Srasuebkul P, Trollor J. Understanding anxiety in adults on
981 the autism spectrum: An investigation of its relationship with intolerance of uncertainty,
982 sensory sensitivities and repetitive behaviours. Autism. 2019;:136236131986890.

bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

983 82. Stark E, Stacey J, Mandy W, Kringelbach ML, Happé F. Autistic Cognition: Charting
984 Routes to Anxiety. Trends Cogn Sci. 2021;25:571–81.

985 83. Casey BJ, Oliveri ME, Insel T. A Neurodevelopmental Perspective on the Research
986 Domain Criteria (RDoC) Framework. Biol Psychiatry. 2014;76:350–3.

987 84. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research Domain
988 Criteria (RDoC): Toward a New Classification Framework for Research on Mental Disorders.
989 Am J Psychiatry. 2010;167:748–51.

990 85. Chrysaitis NA, Jardri R, Denève S, Seriès P. No increased circular inference in adults with
991 high levels of autistic traits or autism. PLOS Comput Biol. 2021;17:e1009006.

992 86. Greene RK, Zheng S, Kinard JL, Mosner MG, Wiesen CA, Kennedy DP, et al. Social and
993 nonsocial visual prediction errors in autism spectrum disorder. Autism Res. 2019;12:878–83.

994 87. Król M, Król M. The world as we know it and the world as it is: Eye-movement patterns
995 reveal decreased use of prior knowledge in individuals with autism. Autism Res. 2019;0.

996 88. Maurer C, Chambon V, Bourgeois-Gironde S, Leboyer M, Zalla T. The influence of prior
997 reputation and reciprocity on dynamic trust-building in adults with and without autism
998 spectrum disorder. Cognition. 2018;172:1–10.

999 89. Randeniya R, Vilares I, Mattingley JB, Garrido MI. Reduced Context Updating but Intact
1000 Visual Priors in Autism. Comput Psychiatry. 2021;5:140.

1001 90. Sapey-Triomphe L-A, Sonié S, Hénaff M-A, Mattout J, Schmitz C. Adults with Autism Tend
1002 to Undermine the Hidden Environmental Structure: Evidence from a Visual Associative
1003 Learning Task. J Autism Dev Disord. 2018. https://doi.org/10.1007/s10803-018-3574-1.

1004 91. Sapey-Triomphe L-A, Weilnhammer VA, Wagemans J. Associative learning under
1005 uncertainty in adults with autism: Intact learning of the cue-outcome contingency, but slower
1006 updating of priors. Autism. 2021;:136236132110450.

1007 92. Wang Q, DiNicola L, Heymann P, Hampson M, Chawarska K. Impaired Value Learning for
1008 Faces in Preschoolers With Autism Spectrum Disorder. J Am Acad Child Adolesc Psychiatry.
1009 2018;57:33–40.

1010 93. Zhang F, Jaffe-Dax S, Wilson RC, Emberson LL. Prediction in infants and adults: A
1011 pupillometry study. Dev Sci. 2018;:e12780.

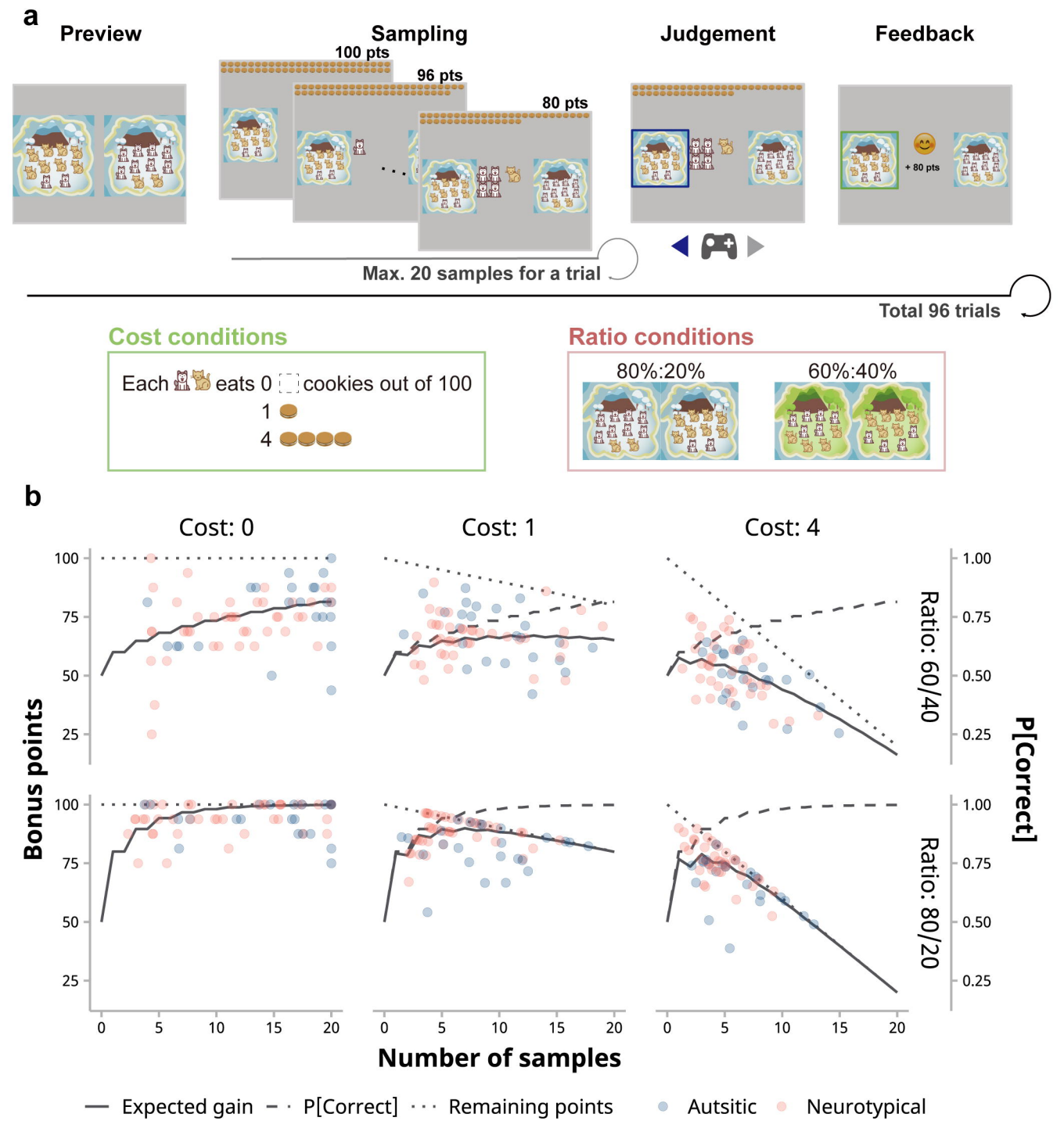
1012 94. Adams RA, Huys QJM, Roiser JP. Computational Psychiatry: towards a mathematically
1013 informed understanding of mental illness. J Neurol Neurosurg Psychiatry. 2016;87:53–63.

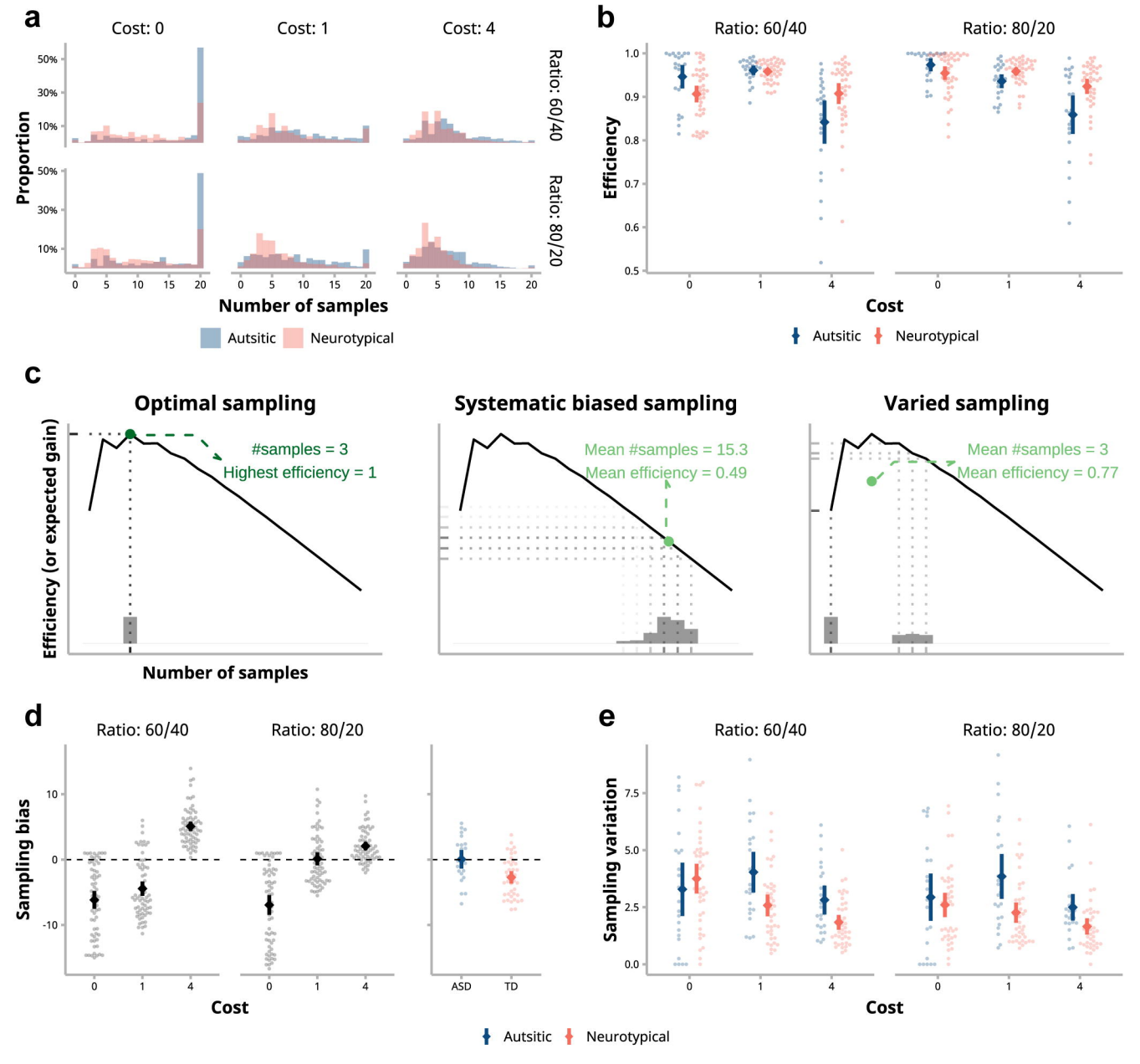
1014 95. Huys QJM, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience
1015 to clinical applications. Nat Neurosci. 2016;19:404–13.

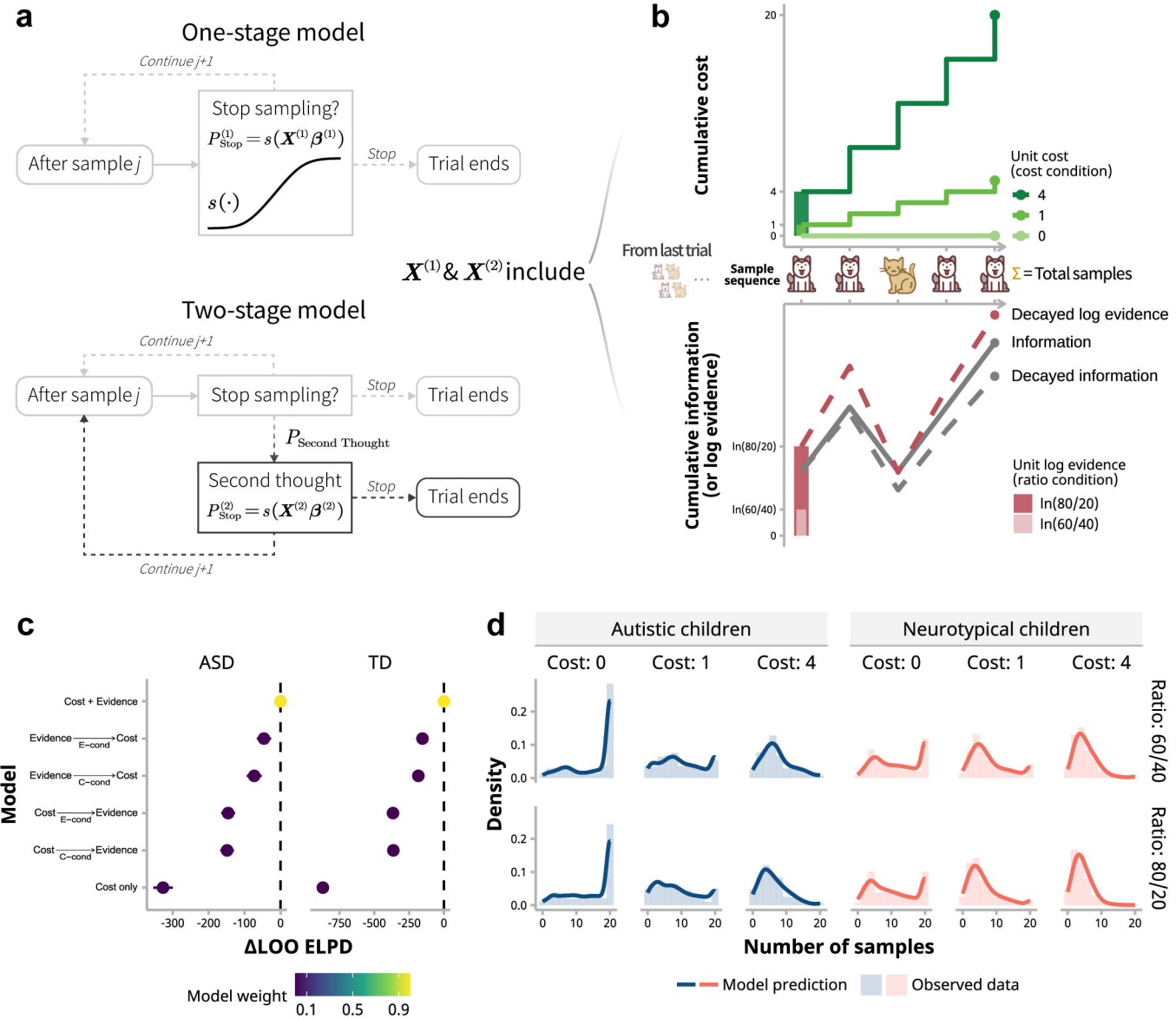
bioRxiv preprint doi: https://doi.org/10.1101/2024.02.04.578786; this version posted February 6, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

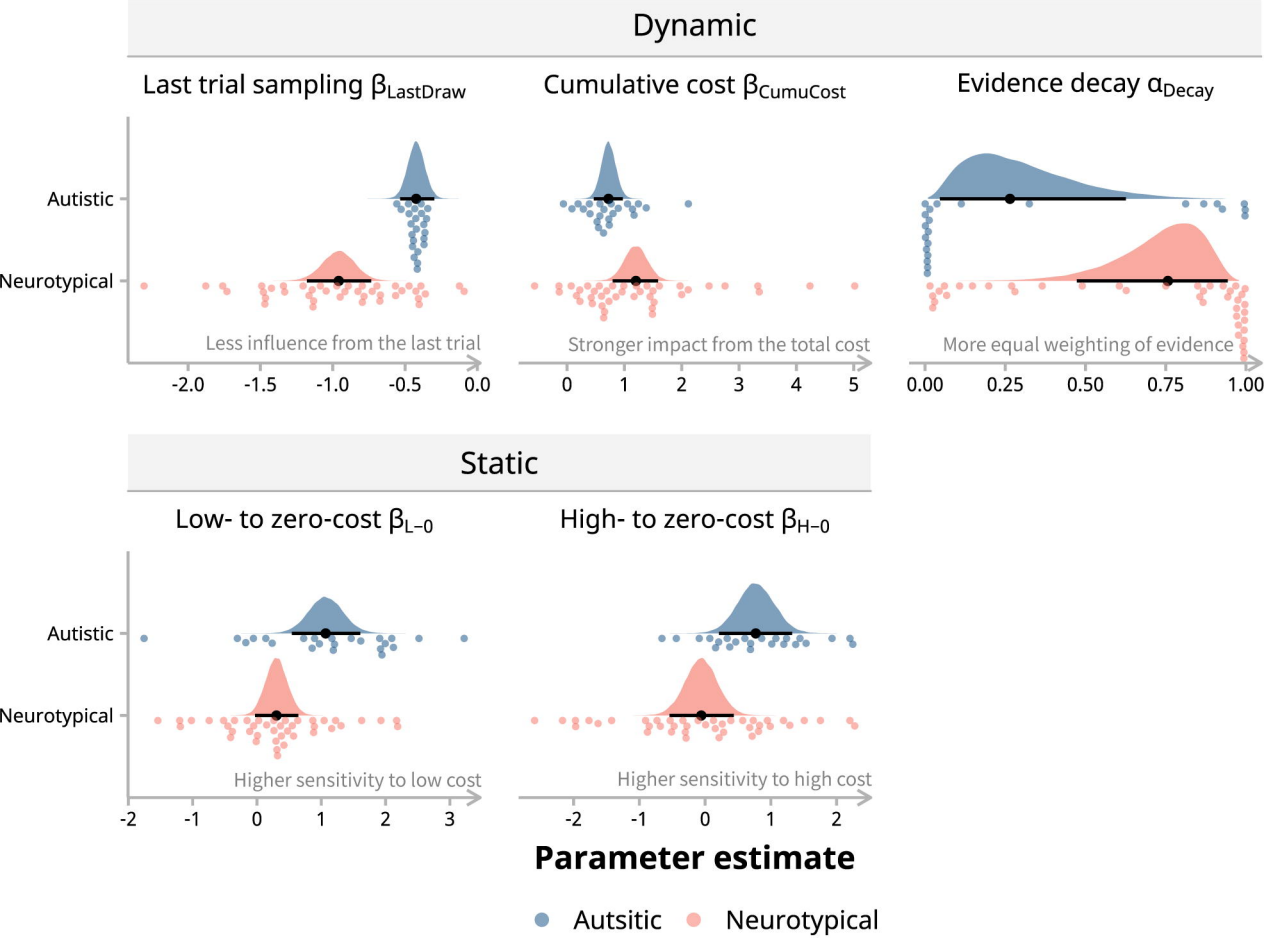
1016 96. Uljarević M, Frazier TW, Jo B, Billingham WD, Cooper MN, Youngstrom EA, et al. Big
1017 Data Approach to Characterize Restricted and Repetitive Behaviors in Autism. J Am Acad
1018 Child Adolesc Psychiatry. 2022;61:446–57.

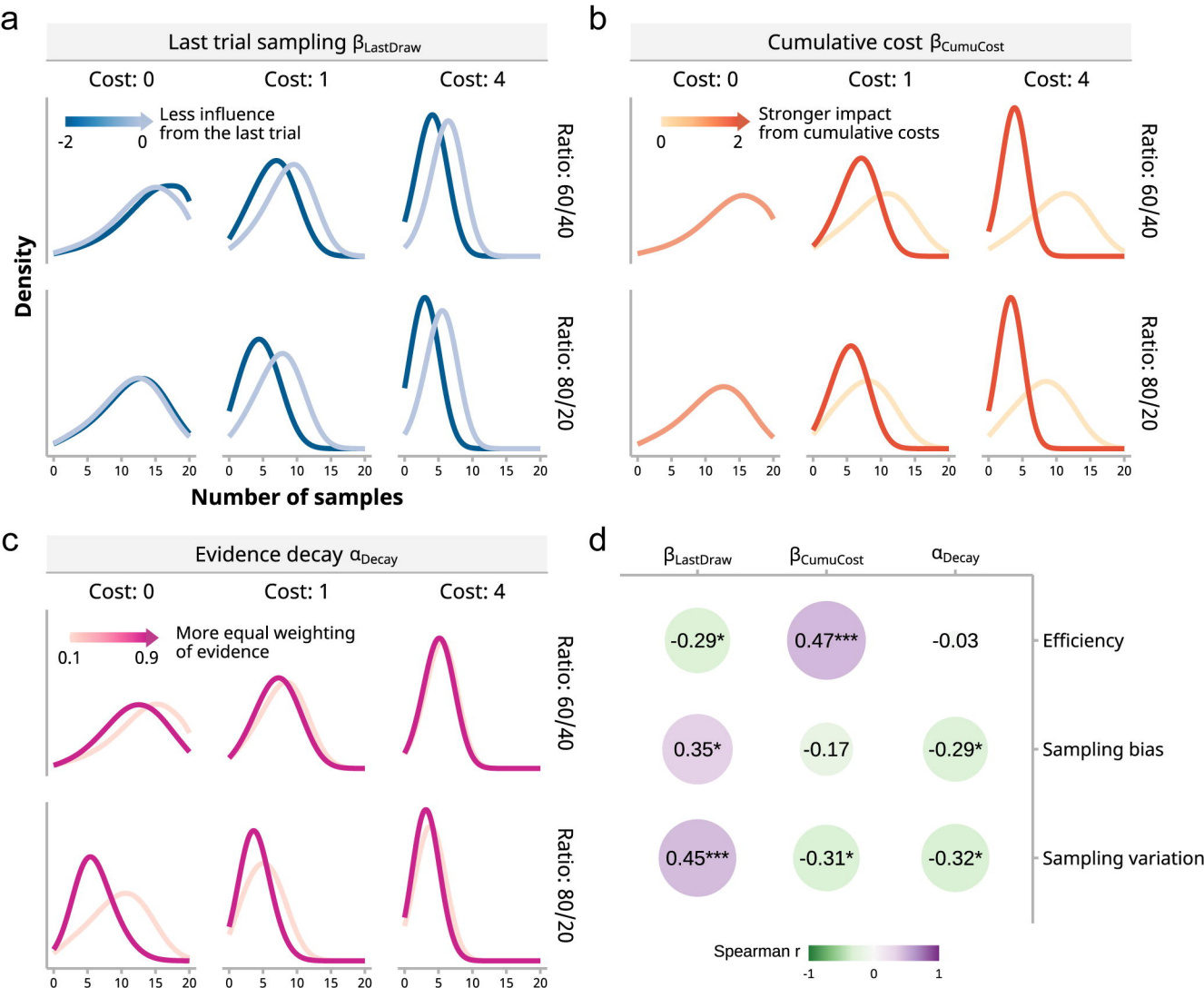
1019


a
Preview
Sampling
Judgement
Feedback
100 pts
96 pts
80 pts
+ 80 pts
Max. 20 samples for a trial
Total 96 trials
Cost conditions
Each eats 0 cookies out of 100
1
4
Ratio conditions
80%:20%
60%:40%
b
Cost: 0
Cost: 1
Cost: 4
Bonus points
Ratio: 60/40
Ratio: 80/20
P[Correct]
Number of samples
Expected gain
P[Correct]
Remaining points
Autsitic
Neurotypical


a
Cost: 0
Cost: 1
Cost: 4
Ratio: 60/40
Ratio: 80/20
Proportion
50%
30%
10%
0 5 10 15 20
Number of samples
Autsitic
Neurotypical
b
Ratio: 60/40
Ratio: 80/20
Efficiency
1.0
0.9
0.8
0.7
0.6
0.5
0 1 4
Cost
Autsitic
Neurotypical
c
Optimal sampling
Systematic biased sampling
Varied sampling
Efficiency (or expected gain)
#samples = 3
Highest efficiency = 1
Mean #samples = 15.3
Mean efficiency = 0.49
Mean #samples = 3
Mean efficiency = 0.77
Number of samples
d
Ratio: 60/40
Ratio: 80/20
Sampling bias
10
0
-10
0 1 4
ASD TD
Cost
e
Ratio: 60/40
Ratio: 80/20
Sampling variation
7.5
5.0
2.5
0.0
0 1 4
Cost
Autsitic
Neurotypical


a
One-stage model
Continue j+1
After sample j
Stop sampling?
$P^{(1)}_{\text{Stop}} = s(\boldsymbol{X}^{(1)}\boldsymbol{\beta}^{(1)})$
$s(\cdot)$
Stop
Trial ends
$\boldsymbol{X}^{(1)}$ & $\boldsymbol{X}^{(2)}$ include
Two-stage model
Continue j+1
After sample j
Stop sampling?
Stop
Trial ends
$P_{\text{Second Thought}}$
Second thought
$P^{(2)}_{\text{Stop}} = s(\boldsymbol{X}^{(2)}\boldsymbol{\beta}^{(2)})$
Stop
Trial ends
Continue j+1
b
Cumulative cost
20
4
1
0
Unit cost
(cost condition)
4
1
0
From last trial
Sample sequence
Σ = Total samples
Cumulative information
(or log evidence)
Decayed log evidence
Information
Decayed information
ln(80/20)
ln(60/40)
0
Unit log evidence
(ratio condition)
ln(80/20)
ln(60/40)
c
ASD
TD
Cost + Evidence
Evidence → Cost (E−cond)
Evidence → Cost (C−cond)
Cost → Evidence (E−cond)
Cost → Evidence (C−cond)
Cost only
Model
-300 -200 -100 0
-750 -500 -250 0
ΔLOO ELPD
Model weight
0.1 0.5 0.9
d
Autistic children
Neurotypical children
Cost: 0
Cost: 1
Cost: 4
Ratio: 60/40
Ratio: 80/20
Density
0.2
0.1
0.0
0 10 20
Number of samples
Model prediction
Observed data


Dynamic
Last trial sampling $\beta_{LastDraw}$
Cumulative cost $\beta_{CumuCost}$
Evidence decay α_{Decay}
Autistic
Neurotypical
Less influence from the last trial
Stronger impact from the total cost
More equal weighting of evidence
-2.0
-1.5
-1.0
-0.5
0.0
0
1
2
3
4
5
0.00
0.25
0.50
0.75
1.00
Static
Low- to zero-cost β_{L-0}
High- to zero-cost β_{H-0}
Autistic
Neurotypical
Higher sensitivity to low cost
Higher sensitivity to high cost
-2
-1
0
1
2
3
Parameter estimate
Autsitic
Neurotypical


a
Last trial sampling $\beta_{LastDraw}$
Cost: 0
Cost: 1
Cost: 4
Less influence from the last trial
-2
0
Ratio: 60/40
Ratio: 80/20
Density
0 5 10 15 20
Number of samples
b
Cumulative cost $\beta_{CumuCost}$
Stronger impact from cumulative costs
0
2
c
Evidence decay α_{Decay}
More equal weighting of evidence
0.1
0.9
d
$\beta_{LastDraw}$
$\beta_{CumuCost}$
α_{Decay}
-0.29*
0.47***
-0.03
Efficiency
0.35*
-0.17
-0.29*
Sampling bias
0.45***
-0.31*
-0.32*
Sampling variation
Spearman r
-1
0
1