# RESEARCH

**Molecular Autism** 

**Open Access** 



# The preference for surprise in reinforcement learning underlies the differences in developmental changes in risk preference between autistic and neurotypical youth

Motofumi Sumiya<sup>1,2\*</sup>, Kentaro Katahira<sup>3,4</sup>, Hironori Akechi<sup>5</sup> and Atsushi Senju<sup>1,2</sup>

# Abstract

**Background** Risk preference changes nonlinearly across development. Although extensive developmental research on the neurotypical (NTP) population has shown that risk preference is highest during adolescence, developmental changes in risk preference in autistic (AUT) people, who tend to prefer predictable behaviors, have not been investigated. Here, we aimed to investigate these changes and underlying computational mechanisms.

**Method** We ran a game-like risk-sensitive reinforcement learning task on 75 participants aged 6–30 years (AUT group, n = 31; NTP group, n = 44). Focusing on choices between alternatives with the same objective value but different risks, we calculated the risk preference and stay probability of a risky choice after a rewarding or non-rewarding outcome. Analyses using *t*-tests and multiple regression analyses were conducted. Using the choice-related data of each participant, we fit four reinforcement learning models and compared the fit of each model to the data. Furthermore, we validated the results of model fitting with multiple methods, model recovery, parameter recovery, and posterior predictive check.

**Results** We found a significant difference in nonlinear developmental changes in risk preference between the AUT and NTP groups. The computational modeling approach with reinforcement learning models revealed that individual preferences for surprise modulated such preferences.

**Conclusions** These findings indicate that for NTP people, adolescence is a developmental period involving risk preference, possibly due to lower surprise aversion. Conversely, for AUT people, who show opposite developmental change of risk preference, adolescence could be a developmental period involving risk avoidance because of low surprise preference.

Keywords Risk preference, Developmental changes, Autism, Reinforcement learning, Surprise

\*Correspondence: Motofumi Sumiya motofumisumiya@gmail.com Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

## Background

Risk preference, the propensity to take or avoid risk, is a fundamental driver of decision-making, which reportedly undergoes complex developmental changes across one's lifespan. Psychometric questionnaire surveys and laboratory task studies provide empirical evidence supporting a developmental peak in risk-taking in mid-adolescence, that is, nonlinear developmental changes in risk-taking that increase and decrease after adolescence [1–5]. Conversely, adolescents are also known to make more prudent decisions than adults, depending on the task and context [6–8]. Although extensive research on neurotypical (NTP) individuals has revealed the developmental trajectory of changes in risk preference, our understanding of how these processes unfold in autistic (AUT) individuals remains limited.

Clinically, AUT is defined as a neurodevelopmental condition that is primarily characterized by difficulties in social communication and interaction as well as restricted, repetitive patterns of behavior, interests, or activities [9]. In recent years, multiple studies with large numbers of research participants have shown that everyday risky behaviors, such as binge drinking or using illicit drugs, which are considered problematic in NTP individuals, are less common in AUT individuals [10, 11]. Such risk preferences in AUT individuals have also been investigated in experimental studies. Recently, van der Plas et al. [12] conducted a narrative review of 104 decision-making studies on AUT people and found that their performance in reward-learning paradigms (e.g., learning which deck of cards provides the best reward) was similar to that of NTP individuals, but their performance in value-based paradigms (e.g., making a decision based on a choice between two outcomes that differ in subjective value) was different from that of NTP individuals. For example, in several studies that adopted valuebased paradigms, financial risk-taking tasks focused on choices between alternatives of equivalent objective value but different risk for rewards (i.e., a choice between a sure option to win \$1 versus a risky option to have a 50% chance to win \$2) have shown that AUT adults are more risk-averse than NTP adults [13-15]. Although these studies suggest that differential subjective value processing, not reward learning, in risk-taking decisions underlies the risk aversion in AUT people, the core underlying mechanism that affects subjective decision-making is not known.

The computational modeling approach has helped us investigate the possible computational mechanisms underlying risk-taking decisions primarily in the NTP population. Multiple reinforcement learning models, such as the utility model, which incorporates nonlinear subjective utilities for different amounts of reward, or risk-sensitive model in which positive and negative prediction errors (differences between a decision outcome and its predicted outcome) have asymmetric effects on learning, have been proposed to explain risk-taking decision-making [16, 17]. Recently, we proposed that the preference for surprise in a decision outcome is a critical factor that can modulate risk preference; surprise occurs because of prediction errors, regardless of whether the error is positive or negative [18]. Based on the cognitive evolutionary model of surprise [19] and experimental studies [20-22], we proposed a reinforcement learning model in our previous study, the surprise model, by introducing a parameter that modulates the value of the outcome based on surprise in each decision [18]. Using two datasets of risk-taking tasks with monetary outcomes [16, 23] and behavioral simulations, we showed that the surprise model had a better fit, indicating that surprise in each decision leads individuals to risk-averse behavior. Therefore, we hypothesized that the tendency toward risk aversion in AUT people is explained by their aversion to surprise. This is because AUT individuals tend to have a strong preference for routine, and repetitive behaviors are hypothesized to stem from the different ways in which prediction errors are processed, which could lead to a preference for predictable behaviors or situations [23, 24]. Accordingly, it is possible that the preference for surprise, as a core mechanism underlying the developmental change in risk preference, leads to a difference in risk preference between the two neurodiverse populations.

In this study, we aimed to investigate the nonlinear developmental changes in risk preference in AUT and NTP individuals as well as to elucidate the underlying computational mechanism from the perspective of surprise preference by adopting a multi-method approach that integrates online experimental paradigms, crosssectional data, and computational modeling. We hypothesized that the risk preferences (i.e., choices between alternatives with the same objective value but different risks) of AUT people would differ from those of NTP people, which change nonlinearly across development. Furthermore, we hypothesized that this difference in risk preference between the two neurodiverse populations would be underpinned by a preference for surprise.

#### Methods

## Participants

This study was conducted as part of a multiple-study protocol project targeting students or alumni of a private school, Musashino Higashi Gakuen, and included NTP participants and AUT participants who had been clinically diagnosed with autism spectrum disorder by at least one pediatrician, child psychiatrist, or

clinical psychologist. In this study, 125 participants aged 6–30 years completed an online task in their homes. After multiple data exclusion steps, we analyzed the data of 75 participants (17 females, 57 males, and 1 person who did not prefer to report; mean age =  $16.03 \pm 5.86$  years; AUT group, n = 31; NTP group, n = 44) (Supplementary material 1: Table S1). The parents of all participants completed the Japanese version of the Social Communication Questionnaire (SCQ [25]). Among these participants, 27 and 25 in the AUT and NTP groups, respectively, reported their intelligence quotient (IQ) scores measured for previous projects (e.g., [26-28]) using either the Wechsler Intelligence Scale for Children III or Wechsler Adult Intelligence Scale, Revised. Since age and IQ (those for reported) were significantly different between the groups, we performed the correlation analysis between age and IQ in each group to exclude the possibility if the nonlinear developmental changes in risk preference can be explained by the changes in IQ, and confirmed linearly the correlation in each group (ASD: r = -0.386, t(25) =-2.095, p=0.046, NTP: r=-0.456, t(23)=-2.454, p = 0.022). These data suggest that IQ is not a significant indicator of risk preference.

## **Exclusion criteria**

We excluded data from eight participants who did not press the button in > 10% of the trials. We then screened the data for a reaction time that was too fast (mean of all trials < 300 ms) and the same button being pressed in > 90% of the trials; these applied to none of the participants. Additionally, based on the study by Rosenbaum et al. [17], we excluded the data of 42 participants whose mean accuracy in the test trials in the second and third blocks was < 60%. After applying these criteria, we used regression analysis to confirm that participants learned the probability of each option block by block (t=6.065, p < 0.001) (mean accuracy of test blocks: 1=0.63, 2=0.77, 3=0.79). Subjects who performed the task intensively should increase their correct response rate as the task progresses, and should exceed 50% of the random correct response rate from the second block onward, when they have probably made some progress in their learning. Rigorous exclusion methods can improve data quality by filtering out inattentive or fraudulent participants without introducing significant bias into online experiment data [29]. Therefore, the criteria were critical for selecting participants who took the task seriously and made the results of this study valid, as a project with an inclusive approach that does not pre-select participants by IQ or language development. To exclude the possibility that our criteria did not systematically exclude the certain pool of participants, such as participants with low IQ or high severity of symptoms, we examined whether the excluded participants differed systematically from the remaining participants. We confirmed that we did not find any significant difference in the IQ or SCQ score in each group (Supplementary material 1: Table S2, S3).

## **Experimental procedure**

Most participants completed the task on their personal computers (PCs), and the remainder completed the task on their tablets (PC, 62 participants; tablets, 13 participants) (Supplementary material 1: Table S1). The participants accessed Pavlovia (pavlovia.org) to complete the experimental task, which was developed using PsychoPy v2021.2.3. The stimuli used in the task (aliens, treasures, and backgrounds) were adapted from open-access materials created by Kool et al. [30], whereas the other stimuli (rockets) were adapted from an online site (iStockphoto. com).

#### **Experimental task**

In this study, we created a treasure task based on the risk-sensitive reinforcement learning task of Rosenbaum et al. [17] (Fig. 1). As a cover story, the participants rode a rocket to a planet and received treasures from an alien



#### Fig. 1 Task sequence

who gave them 0-8 treasures. The number of treasures the participants received depended on the rocket they chose, and their challenge was to collect as many treasures as possible and obtain the large rocket. There were five rockets, including three deterministic rockets for 0, 2, and 4 treasures, respectively, and two probabilistic rockets for 0 or 4 and 0 or 8 treasures, respectively (50% each). In the task, there were 3 blocks and 183 trials, including 42 sure vs. risky choices (24 choices for a 100% 2-treasure rocket vs. a 50% 0- or 4-treasure rocket and 18 choices for a 100% 4-treasure rocket vs. a 50% 0- or 8-treasure rocket), 24 other choices (a 100% 2-treasure rocket vs. a 50% 0- or 8-treasure rocket), 42 test choices that ensured that participants learned the features of each rocket (e.g., a 100% 2-treasure rocket vs. a 100% 4-treasure rocket), and 75 forced-choice trials in which participants were forced to learn the features of each rocket. The trial order was pseudo-randomized based on seven templates created from the order of the seven participants in the study by Rosenbaum et al. [17].

## Basic analysis Risk preference

First, we compared group differences in risk preference using a *t*-test. Subsequently, using multiple regression analysis, we examined whether there was a group difference in risk preference that changed nonlinearly with age by modeling the interaction term of the quadratic term of age and groups. As confounding covariates, in addition to the reported gender (male, female, or preferred not to report) and execution device (PC or tablet), we modeled the accuracy of the test trials in the second and third blocks, which was significantly correlated with IQ scores among participants who reported their IQ score (r=0.34, p=0.012), to remove the effect of the intellectual component associated with this task. Accordingly, the interaction term of the quadratic term of age and groups, interaction term of age and groups, execution device, reported gender, and accuracy of the test trials were modeled. Models were assessed using the lm\_robust command from the "estimatr" package [31] with heteroskedasticity-consistent 0 robust standard errors. This method estimates the standard errors under heterogeneous variances and allows for valid inference even when the assumption of constant variance is violated. In these analyses, categorical variables were coded as 1 = AUT and -1 = NTP for the group and as 1 = PCand -1 = tablet for the device; continuous variables (age and accuracy) were scaled to mitigate multicollinearity issues. We checked for the multicollinearity of regressors within models using the "performance" package [32], and confirmed that the variance inflation factor (VIF) of each regressor did not exceed 10, while VIF values greater

than 10 are a sign of high, unacceptable correlation of model predictors.

Furthermore, as a consecutive analysis to confirm the nonlinear developmental change in risk preference in each group, we conducted a multiple regression analysis for risk preference within each group using the model with the quadratic term of age, in addition to linear term of age, reported sex, execution device, accuracy of the test trials, and SCQ score.

## Stay probability

We calculated the stay probability (i.e., the probability of choosing the same option consecutively) of the sure and risky choices both after the rewarding and non-rewarding outcomes. In these calculations, we did not consider the outcome of the forced-choice trials that appeared between the re-risk choice trials. We investigated the group differences in each stay probability that changed nonlinearly with age by modeling the interaction term of the quadratic term of age and groups. The other regressors were the same as those used in the analysis for risk preference. Furthermore, as a consecutive analysis to confirm the nonlinear developmental change in each stay probability in each group, we conducted a multiple regression analysis within each group using the same regressors as in the analysis for risk preference.

## **Computational modeling**

## Model description

In this study, we fit three widely used models, the Q-learning (QL), utility, and risk-sensitive QL (RSQL) models [16–18], and our proposed model, the surprise model [18]. For each model, the learning rate was constrained to the range,  $0 \le \alpha$ ,  $\alpha^+$ ,  $\alpha^- \le 1$ , with a beta (2,2) prior distribution, and the inverse temperature was constrained to the range,  $0 \le \beta \le 20$ , with a gamma (2,3) prior distribution. The utility parameter was constrained to the range,  $0 \le \rho \le 2.5$ , with a gamma (1.5,1.5) prior distribution. Additionally, the modulation rate of the surprise model was constrained to  $-1 \le d \le 1$  with a uniform prior distribution.

## QL model

The QL model was used as the base model for the other three models. The QL model incorporates the Rescorla-Wagner rule, where only the *Q*-value of the chosen option is updated based on a prediction error that explains the observed behavior by computing the action value Q(t) for each trial *t*, which represents the expected outcome of the action. The *Q*-value of the chosen action is iteratively updated based on a prediction error, which is the difference between the expected outcome Q(t) and the received outcome *r*, by a learning rate  $\alpha$ .

#### Utility model

The utility model is a QL model that incorporates nonlinear subjective utilities for different amounts of rewards. In this model, the reward outcome is exponentially transformed by  $\rho$ , which represents the curvature of the subjective utility function for each individual.

$$Q(t+1) = Q(t) + \alpha(r(t)^{\rho} - Q(t))$$

#### RSQL model

In the RSQL model, which is a QL model, positive and negative prediction errors have asymmetric effects on learning. Specifically, there are separate learning rates:  $\alpha^+$  and  $\alpha^-$  for positive and negative prediction errors, respectively.

 $Q(t + 1) = Q(t) + \alpha^+(r(t) - Q(t))$  for positive prediction error

 $Q(t + 1) = Q(t) + \alpha^{-}(r(t) - Q(t))$  for negative prediction error

#### Surprise model

In the surprise model, the received outcome r is affected by surprise (absolute value of the prediction error). In this model, S(t) is the subjective utility modulated by surprise. The degree of modulation is controlled by the parameter d as follows:

$$S(t) = r(t) - d|r(t) - Q(t)|,$$
$$Q(t+1) = Q(t) + \alpha(S(t) - Q(t)).$$

For all models, the probability of choosing option *i* during trial *t* is provided by the softmax function:

$$P(a(t) = i) = \frac{\exp(\beta \cdot Q_i(t))}{\sum_{i=1}^{K} \exp(\beta \cdot Q_i(t))},$$

where  $\beta$  is the inverse temperature parameter that determines the sensitivity of the choice probabilities to differences in the values, and *K* represents the number of possible actions (in the present study, *K*=2); moreover, *a*(*t*) denotes the option chosen in trial *t*.

## **Parameter estimation**

We fit the parameters of each model using the maximum a posteriori (MAP) estimation, which improves parameter estimates by incorporating prior information on parameter values [33, 34]. We also approximated the log marginal likelihood (model evidence) for each model using the Laplace approximation [35]. We used the R function "solnp" in the "Rsolnp" package [36] to estimate the fitting parameters.

#### Model comparison

For model selection, the model evidence (log marginal likelihood) for each model and participant was subjected to random-effects Bayesian model selection (BMS) using the "spm\_BMS" function in SPM12 [37]. BMS provides a less biased and statistically more accurate way to identify the best model at the group level by estimating the protected excess probability, which is defined as the probability of a particular model being more frequent in the population among a set of candidate models [34]. We conducted a model comparison for all participants in each group. To visualize how well the winning model fit the data, we also determined the number of participants that best fit each model [34].

## **Estimated parameters**

After the model comparison, we investigated group differences in the estimated parameters of the winning model, which changed nonlinearly with age, by modeling the interaction term of the quadratic term of age and groups. The other regressors were the same as those used in the risk preference analysis. Furthermore, as a consecutive analysis to confirm the nonlinear developmental change in the estimated parameters in each group, we conducted a multiple regression analysis within each group with the same regressors as in the risk preference analysis.

#### Parameter and model recovery

We conducted parameter recovery to assess the reliability of parameter estimation procedures; specifically, we determined how accurately parameters were estimated when the true generative model and its parameter values were known [34, 38, 39]. Model recovery was performed to test the discriminability of each model [34, 38, 39]. Details of each analysis are provided in Supplementary material 1 (Supplementary material 1: Text S1, S2).

## Posterior predictive check

We performed a posterior predictive check that analyzed the simulated data in the same way as the analyses of the empirical data to validate that each model adequately captured behavioral data [34, 40].Detailed information of the analysis is provided in Supplementary material 1 (Supplementary material 1 Text S3).

#### Supplemental analysis with open data

As the task used in this study was based on the study by Rosenbaum et al. [17], to confirm that the surprise model was better fitted to the risk preference data with similar age diversity, we conducted model fitting and model comparisons. For model fitting, the parameters were set to be the same as those used in this study. Additionally, to confirm the nonlinear relationship between age and the surprise parameter, we conducted a regression analysis of the surprise parameter using the quadratic term of the scaled age.

## Results

#### **Risk preference**

We did not find a significant group difference in risk preference (t(62.282) = 0.219, p = 0.827, d = 0.052). Meanwhile, the multiple regression analysis revealed a significant interaction of group and the quadratic term of age ( $\beta = 0.098$ , t(65) = 2.927, p = 0.005, 95% CI [0.031, 0.165]) (Fig. 2a). The results of the other regressors are summarized in Fig. 3 and Supplementary material 1: Table S4. We calculated the multicollinearity of the regressors and confirmed weak correlations between them (Figure S1a).

With consecutive multiple regression analyses within each group, we confirmed the tendency of an inverted U-curve developmental change in the NTP group ( $\beta$ = -0.039, t(36)= -1.953, p=0.059, 95% CI [-0.080, 0.002]) and a U-curve developmental change in the AUT group ( $\beta$ =0.055, t(24)=1.970, p=0.060, 95% CI [-0.003, 0.113]) (Supplementary material 1: Figure S2 and S3, Table S5).

## Stay probability

We performed a multiple regression analysis for each stay probability for the sure and risky choices. We found a significant interaction of group and the quadratic term of age for the stay probability of sure choices ( $\beta = -0.092$ , t(65) = -2.290, p = 0.025, 95% CI [-0.172, -0.012]) (Fig. 4a). For the stay probability of risky choices after a non-rewarding outcome, we removed the covariates, sex and device, from the model to address the excessive correlation among the explanatory variables. We found a significant interaction of group and the quadratic term of age ( $\beta = 0.094$ , t(67) = 2.151, p = 0.035, 95% CI [0.007, 0.177]) (Fig. 4b). Furthermore, we found no significant relationship, only a tendency toward one, between group and the quadratic term of age for the stay probability of risky choices after a rewarding outcome ( $\beta = 0.056$ , t(63) = 1.795, p = 0.077, 95% CI [-0.006, 0.118]) (Fig. 4c). The results for the other regressors are summarized in Fig. 5 and Supplementary material 1: Table S6. Information on the multicollinearity of the regressors is summarized in Supplementary material 1: Figure S9.

We conducted consecutive multiple regression analyses for each stay probability within each group using the same regressors as those in the risk preference analysis in each group; however, in the analysis of the stay probability for sure choices in the NTP group, a regressor (device) was removed to deal with multicollinearity. We confirmed the significance of a U-curve developmental change in sure choices in the NTP group ( $\beta$ =0.071, t(37)=3.088, p=0.004, 95% CI [0.024, 0.118]) but not in the AUT group ( $\beta$ = -0.021, t(24)= -0.595, p=0.558, 95% CI [-0.092, 0.051]). Further, we confirmed the significance of a U-curve developmental change in risky choices after a non-rewarding outcome in the AUT



Fig. 2 Relationship of age with (a) risk preference and (b) the estimated surprise parameter. The regression lines are from a linear regression model including linear and quadratic age terms. Data points represent individual participants. The lines and 95% confidence intervals were estimated under heterogeneous variances to reduce the influence of outlier data



Fig. 3 Model coefficients for risk preference (blue) and the surprise parameter (orange). Parameters within models were coded as group (AUT = 1, NTP = -1), device (PC = 1, tablet = -1), and sex (M = male, F = female, N = not reported)



Fig. 4 Relationship between age and the stay probability of a (A) sure choice, (B) risky choice after a non-rewarding outcome, and (C) risky choice after a rewarding outcome. The regression line is from a linear regression model including linear and quadratic age terms. Data points represent individual participants. The lines and 95% confidence intervals were estimated under heterogeneous variances to reduce the influence of outlier data



Fig. 5 Model coefficient for stay probability for a risky choice after a non-rewarding outcome (blue) and rewarding outcome (orange). Parameters within models were coded as group (AUT = 1, NTP = -1), device (PC = 1, tablet = -1), and sex (M = male, F = female, N = not reported)

group ( $\beta$ =0.080, *t*(23)=3.532, *p*=0.002, 95% CI [0.033, 0.126]) but not in the NTP group ( $\beta$ = -0.016, *t*(36)= -0.424, *p*=0.674, 95% CI [-0.094, 0.062]). Furthermore, we confirmed the significance of a U-curve developmental change in risky choices after a rewarding outcome in the AUT group ( $\beta$ =0.073, *t*(24)=2.782, *p*=0.010, 95% CI [0.019, 0.127]) but not in the NTP group ( $\beta$ =0.020, *t*(34)=1.016, *p*=0.317, 95% CI [-0.020, 0.061]). Instead, the linear effect of age on risky choices after a rewarding outcome was significant in the NTP group ( $\beta$ =0.082, *t*(34)=2.846, *p*=0.007, 95% CI [0.023, 0.141]) (Supplementary material 1: Figure S10 and S11, Table S7).

## Model comparison

We compared the model evidence for each model (log marginal likelihood) and found that the surprise model had the highest value (Fig. 6a). We then performed a Bayesian model comparison to determine the best model to explain choice behavior and found that the surprise model had a significantly higher protected exceedance probability than the other models, indicating that it was more frequent in this population (Fig. 6b). We further

checked the fitness of each model in each group and confirmed that the surprise model had the best fit among the four models (Supplementary material 1: Figure S12). We also confirmed that the fit of the surprise model was the best in a relatively large proportion of participants (42% in the AUT group, 41% in the NTP group) compared with that of the other models (utility model: 35% in the AUT group, 20% in the NTP group; QL model: 13% in the AUT group, 27% in the NTP group; RSQL model: 9.7% in the AUT group, 11% in the NTP group) (Supplementary material 1: Table S8). The distribution of each estimated parameter in the surprise model is shown in Supplementary material 1: Figure S13 to illustrate the potential problem fitting [34, 41]. These results are discussed in the Supplementary material 1 (Supplementary material 1: Text S4).

#### **Estimated parameters**

We used the surprise parameter of the surprise model, the winning model in the model comparisons described below, as the dependent variable in the multiple regression analysis. We found a significant interaction between



Fig. 6 Model comparison. a Evidence of each model (log marginal likelihood). b Bayesian model comparison

group and the quadratic term of age ( $\beta = -0.342$ , t(65) = -2.977, p = 0.004, 95% CI [-0.571, -0.113]) (Fig. 2b). The results of the other regressors are summarized in Fig. 3 and Supplementary material 1: Table S4. Information on the multicollinearity of the regressors is summarized in Supplementary material 1: Figure S1b.

With consecutive multiple regression analyses within each group, we confirmed the significance of a U-curve developmental change in the NTP group ( $\beta$ =0.142, t(36)=2.063, p=0.046, 95% CI [0.002, 0.282]) and an inverted U-curve in the AUT group ( $\beta$ = -0.203, t(24)= -2.249, p=0.034, 95% CI [-0.389, -0.017]) (Supplementary material 1: Figure S4 and S5, Table S9).

In addition to the surprise parameter, we conducted multiple regression analysis with the other parameters in the surprise model: learning rate and inverse temperature. To address multicollinearity, we only modeled the interaction term of the group and age and of the group and the quadratic term of age. The effect of the interaction between group and the quadratic term of age was significant on the inverse temperature ( $\beta = -1.871$ , t(69) = -2.298, p = 0.025, 95% CI [-3.496, -0.247]) but not the learning rate ( $\beta = -0.029$ , t(69) = -0.876, p = 0.384, 95% CI [-0.095, 0.037]) (Supplementary material 1: Figure S6-S8, Table S10). These data are consistent with those in previous literature on the NTP population [42-44] and have added new findings about the AUT population. These results were discussed in the supplemental text (Supplementary material 1: Text S5).

#### Parameter and model recovery

For most of the parameters in all models, the recovered and true parameters were highly correlated (r > 0.91) (QL: alpha, utility: alpha & utility, RSQL: alphaP & alphaN, surprise: alpha & surprise), confirming that these parameters were identifiable (Supplementary material 1: Text S6). Additionally, all models showed high recovery rates (Supplementary material 1: Text S7).

## Posterior predictive check

We conducted a multiple regression analysis of real preference and stay probability data with simulated data from all models, and confirmed that these models, especially utility and surprise models, can capture data sufficiently well to predict real behaviors (Supplementary material 1: Text S8).

#### Supplemental analysis with open data

As a result of model comparisons using the data of Rosenbaum et al. [17], who found a nonlinear developmental change in risk preference in NTP participants of similar age as that in this study and used a similar task but with a different monetary context, we confirmed that the surprise model had a better fit based on the model evidence (Supplementary material 1: Figure S14a) and protected exceedance probability (Supplementary material 1: Figure S14b) than the other models. Accordingly, adding evidence to this study, the surprise model can explain the risk preference data well, including from the developmental perspective. Furthermore, with the regression analysis, we confirmed a significant nonlinear relationship between the quadratic term of age and the surprise parameter ( $\beta = -0.228$ , t(60) = -2.858, p = 0.006, 95% CI [-0.387, -0.068]) (Supplementary material 1: Figure S15).

## Discussion

This is the first study to investigate the age-related nonlinear changes in risk preference in AUT and NTP participants and propose the underlying computational mechanism that best explains the risk preference. Contrary to our hypothesis, we did not find a group difference in the mean risk preference. Instead, we found a significant difference in the relationship between risk preference and the quadratic term of age between the AUT and NTP groups. This finding indicated that risk preference in the NTP group increased toward adolescence and decreased later, in line with the results of previous studies investigating different contexts [1–5]. By contrast, the result suggested that the AUT group showed a developmental curve in the opposite direction: risk preference decreased toward adolescence and increased afterward. Critically, the estimated parameter from the surprise model, the better-fitting model based on model comparison, revealed that the preference for surprise underlies the opposite patterns of developmental change in risk preference between the AUT and NTP groups.

Adolescence is a time of dramatic emotional and social change [45, 46], which also poses vulnerabilities to physiological and neural development [47, 48] in both NTP and AUT individuals. Although no studies have explicitly investigated the age-related nonlinear change in risk preference in AUT, South et al. [49] reported the risk preference of AUT and NTP children and adolescents, which partly aligns with the current study. Their findings implied that risk preferences were similar between NTP and AUT adolescents, but there was a higher risk preference in AUT individuals than in NTP individuals during childhood. This result indicates that age is an important moderator of risk preference differences between AUT and NTP individuals.

For further interpretation of risk preference, the stay probabilities for each option helped us better evaluate the factors influencing risk preference in AUT. In the AUT group, regardless of whether the reward was acquired, the stay probability of risky choices decreased toward adolescence and increased afterwards, but it was not the case for sure choices. These results indicate that among AUT youths, individuals who subjectively prefer risk choose the risky option regardless of the preceding objective outcome value, which may originate from the conformity of both positive and negative prediction errors, as a surprise [50]. Additionally, this finding is consistent with AUT features and preferences for repetitive behavioral patterns [23, 24]. In the NTP group, individuals who subjectively avoided risk tended to stay with the sure choice, but with increasing age, the objective outcome value affected their stay probability for risky choices. Accordingly, in both groups, it seemed that risk-related choices were weighted not only by the objective outcome value but also by subjective value processing.

To uncover the computational mechanisms contributing to these behavioral indicators, we conducted computational modeling using reinforcement learning models that incorporated possible additional factors that could account for risk preference. The model comparison showed that the best-fitting model was the surprise model, which incorporated the surprise parameter that alters the reward sensitivity such that a larger prediction error further attenuates the reward value. Moreover, the relationship between the surprise parameter and quadratic term of age showed a significant group difference in the opposite pattern to that of the relationship between risk preference and the quadratic term of age. Furthermore, these results were supported by those of an additional analysis of the data of a previous risk-taking study on NTP youth [17] that found that the surprise model had a better fit than previous winning models, such as the RSQL and utility models. We also found a significant nonlinear relationship between the quadratic term of age and the surprise parameter. These findings indicate that the preference for surprise is one of the key computational mechanisms underlying developmental changes in risk preference. Adolescence is often considered a developmental period with heightened sensitivity to rewards, resulting in risky behaviors [1-5]. Such reward sensitivity may be represented by the surprise parameter, which reflects individual differences in confidence regarding surprise in decision outcomes.

Moreover, we can consider the association with the surprise parameter with the trait recently focused in the literature of AUT, intolerance to uncertainty (IU). IU is defined as a dispositional trait involving maladaptive responses under conditions of uncertainty, and low IU is often reported in AUT individuals [51]. This trait has been captured in a questionnaire designed to capture general difficulty in coping with the unexpected or unknown, including vulnerability to surprise [52]. However, because it captures the general characteristics of individuals, the background factors that cause surprise and the perspective of reward prediction error (positive or negative prediction error) have not been considered in terms of IU. Therefore, although we cannot directly address the relationship between the surprise parameter and IU at this point, it is possible that the preference for surprise may be one of the hierarchical mechanisms of IU, such that we dislike the occurrence of surprise and thus dislike uncertainty. In future research, it is important to define IU more precisely from a computational point of view [53, 54].

Importantly, previous studies have addressed the issue of developmental change in risk preference from the perspective of the interaction between cognitive control and emotional-incentive processing [5]. NTP individuals demonstrate an aversion to risk as a consequence of vague anxiety about risks during childhood. As they mature, they tend to take a risk as a means of enhancing their sensitivity to rewards and success during

adolescence. However, as their executive functioning develops in adulthood, they subsequently demonstrate an aversion to risk and a preference for certainty. Our finding about NTP individuals is consistent with this proposal. Accordingly, a surprise parameter may partially represent emotional-incentive processing, suggesting that those who dislike surprises may be vaguely anxious and thus risk-averse, while those who prefer surprises may not evoke the anxiety associated with risk. Furthermore, if we consider the AUT, which showed an inverse developmental change trend from the NTP, it is possible that AUT individuals may prefer risks associated with their preference for curiosity and exploratory behavior [55], as anxiety is not aroused in situations such as game play. And it may be that risk aversion in adolescence is due to the difficulties experienced during the developmental process that make them more susceptible to anxiety, and that in adulthood, as executive functions develop, they again develop a preference for risk.

In addition, self-report studies of NTP individuals have reported that risk preference generally increases during adolescence and decreases during adulthood ([56]; for further discussion, see [5]). The trend for NTP individuals in the present study was similar to the findings of these studies, suggesting that game-like tasks such as those used in the present study measure such a general risk preference. It has been previously reported that gamelike tasks (with monetary rewards) in AUT children and adolescents produced results similar to those in the present study [49]. Thus, developmental changes in risk preference in AUT individuals may be generally different from those in NTP individuals, and future studieIns of general risk preference in AUT individuals are warranted. On the other hand, in a study using a financial task that, like the present study, equalized expected values for risky options (probabilistic outcomes) and safe options (deterministic outcomes), risk preference declined interestingly during adolescence [17]. Although there are no studies of developmental changes in risk preference using monetary rewards in AUT individuals, several studies using a similar paradigm in which the expected value of options is set equal have consistently found that AUT adults are more risk averse than NTP adults [13–15]. It is possible that, similar to the findings of this study, developmental changes in financial tasks may differ between NTP and AUT individuals. Future studies should also consider the context of such tasks.

Finally, these two findings were inconsistent with our initial hypothesis. First, AUT participants were not risk-averse, as in previous studies with financial risk preference tasks targeting AUT adults [13–15], indicating that the target age and context of the task (i.e., presence or absence of financial reward) may be important factors

for risk preference. Another finding that did not support our hypothesis was the preference for surprise. In our previous study [18], we assumed that surprises negatively affected the outcome value. However, in the present study, which used a game-like risk preference task, approximately half of the participants in both groups preferred the risky option, and the estimated surprise parameter was below zero, suggesting that surprises increased the value of the outcome for these participants. It is known that some individuals who self-report that they like surprises prefer mysterious consumption-the opportunity to be surprised-over non-mysterious consumption of equal expected value [50]. Accordingly, risk preference may differ in different contexts and individuals from the perspective of surprise preference. Future studies should investigate developmental changes in risk preference in different contexts, such as financial or game-like tasks.

## Limitations

Our study has limitations. First, we performed online tasks as this study was conducted during the coronavirus pandemic. This resulted in a lack of control over the experimental environment, and participant performance varied to such an extent that much data had to be excluded based on the performance criteria. Our target sample size was more than 60 participants in each group, similar to that of Rosenbaum et al. [17]. After multiple data exclusion steps, data from only 60% of participants could be included in the final sample. Using these criteria, we confirmed that the accuracy increased as the task proceeded, similar to that of Rosenbaum et al. [17], and our results from the remaining data are rigid. However, the small number of participants for the analysis limited our ability to detect subtle effects of age, such as the nonlinearity developmental change of risk preference in each group, and the generalizability of our findings. Although it was not possible to provide a prior explanation of the task as part of a multi-study protocol project, such as that provided by Rosenbaum et al. [17], it may be necessary.

Another limitation was the variability of participant characteristics between groups, such as age and IQ, which were significantly different between the groups. Moreover, although not possible for the same reason, it would have been desirable to obtain IQ directly from all participants and to specify the device used to perform the task, such as limiting the use of personal computers. In this study, the parameters age and accuracy, instead of IQ, were statistically controlled for as covariates in the multiple linear regressions. Future studies should examine these conditions in detail after controlling for them. In addition, we believe that conducting longitudinal studies of everyday risk-taking behaviors would allow for more ecologically valid findings.

## Conclusion

The current study is the first to demonstrate a significant difference in nonlinear developmental changes in risk preference between AUT and NTP participants. This finding indicated that, during adolescence, risk preference was similar between the AUT and NTP groups, but the opposite was true in childhood and adulthood. Using a computational modeling approach, we revealed the underlying mechanism of risk preference from the perspective of surprise preference. These findings indicate that in NTP individuals, adolescence is a developmental period in which risk is preferred because of the lowest aversion to surprise, whereas in AUT individuals, adolescence is a developmental period in which risk is avoided because of the highest aversion to surprise.

#### Abbreviations

- AUT Autistic BMS Bayesian model selection
- MAP Maximum a posteriori
- NTP Neurotypical
- RSQL Risk-sensitive Q-learning
- QL Q-learning
- VIF Variance inflation factor

## **Supplementary Information**

The online version contains supplementary material available at https://doi.org/10.1186/s13229-025-00637-5.

Supplementary material 1.

#### Acknowledgements

We would like to acknowledge all the participants, their families, and the teachers at Musashino Higashi Gakuen.

#### Author contributions

MS conceived the project, participated in data collection, data analysis, data interpretation, writing the first version of the manuscript, and provided founding. KK participated in data analysis, data interpretation, and manuscript writing. HA participated to data collection and manuscript writing. AS conceived and supervised the project, participated in manuscript writing and data interpretation, and provided founding. All authors read and approved the final manuscript.

#### Funding

This work was supported by JSPS KAKENHI Grants (21K13748 and 24K16870) to M.S. and KAKENHI Grant (23K22373) to A.S.

#### Availability of data and materials

The experimental task and all codes used in the analysis are available from the Open Science Framework (https://osfio/b5rt2/). Data are available upon request owing to privacy/ethical restrictions.

## Declarations

#### Ethics approval and consent to participate

This study adhered to the Declaration of Helsinki and was approved by the Committee on Ethics of Experimental Research on Human Subjects, Graduate School of Arts and Sciences, University of Tokyo (156–17).

#### **Consent for publication** Not applicable.

## Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Research Center for Child Mental Development, Hamamatsu University School of Medicine, Hamamatsu, Japan. <sup>2</sup>United Graduate School of Child Development, Osaka University, Kanazawa University, Hamamatsu University School of Medicine, Chiba University and University of Fukui, Osaka, Kanazawa, Hamamatsu, Chiba, Fukui, Japan. <sup>3</sup>Human Informatics and Interaction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan. <sup>4</sup>Graduate School of Comprehensive Human Sciences, University of Tsukuba, Tsukuba, Japan. <sup>5</sup>Graduate School of Education, Kyoto University, Kyoto, Japan.

#### Received: 4 September 2024 Accepted: 4 January 2025 Published online: 16 January 2025

#### References

- Steinberg L. Risk taking in adolescence: what changes, and why? Ann N Y Acad Sci. 2004;1021:51–8.
- Steinberg L. Cognitive and affective development in adolescence. Trends Cogn Sci. 2005;9:69–74.
- Ernst M, Paulus MP. Neurobiology of decision making: a selective review from a neurocognitive and clinical perspective. Biol Psychiatry. 2005;58:597–604.
- Luna B, Paulsen DJ, Padmanabhan A, Geier C. Cognitive control and motivation. Curr Dir Psychol Sci. 2013;22:94–100.
- Shulman EP, Smith AR, Silva K, Icenogle G, Duell N, Chein J, et al. The dual systems model: review, reappraisal, and reaffirmation. Dev Cogn Neurosci. 2016;17:103–17.
- Wilbrecht L, Davidow JY. Goal-directed learning in adolescence: neurocognitive development and contextual influences. Nat Rev Neurosci. 2024;25:176–94.
- van Duijvenvoorde ACK, van Hoorn J, Blankenstein NE. Risks and rewards in adolescent decision-making. Curr Opin Psychol. 2022;48: 101457.
- Defoe IN, Dubas JS, Figner B, van Aken MA. A meta-analysis on age differences in risky decision making: adolescents versus children and adults. Psychol Bull. 2015;141:48–84.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Washington, DC: American Psychiatric Press; 2013.
- Weir E, Allison C, Baron-Cohen S. Understanding the substance use of autistic adolescents and adults: a mixed-methods approach. Lancet Psychiatry. 2021;8:673–85.
- Mangerud WL, Bjerkeset O, Holmen TL, Lydersen S, Indredavik MS. Smoking, alcohol consumption, and drug use among adolescents with psychiatric disorders compared with a population based sample. J Adolesc. 2014;37:1189–99.
- van der Plas E, Mason D, Happé F. Decision-making in autism: a narrative review. Autism. 2023;27(6).
- De Martino B, Harrison NA, Knafo S, Bird G, Dolan RJ. Explaining enhanced logical consistency during decision making in autism. J Neurosci. 2008;28:10746–50.
- 14. Gosling CJ, Moutier S. Brief report: risk-aversion and rationality in autism spectrum disorders. J Autism Dev Disord. 2018;48:3623–8.
- Wu HC, White S, Rees G, Burgess PW. Executive function in high-functioning autism: decision-making consistency as a characteristic gambling behaviour. Cortex. 2018;107:21–36.
- Niv Y, Edlund JA, Dayan P, O'Doherty JP. Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. J Neurosci. 2012;32:551–62.
- Rosenbaum GM, Grassie HL, Hartley CA. Valence biases in reinforcement learning shift across adolescence and modulate subsequent memory. Elife. 2022;11: e64620.
- 18. Sumiya M, Katahira K. Surprise acts as a reducer of outcome value in human reinforcement learning. Front Neurosci. 2020;14:852.

- Reisenzein R, Horstmann G, Schutzwohl A. The cognitive-evolutionary model of surprise: a review of the evidence. Top Cogn Sci. 2019;11:50–74.
- Knight EJ, Klepac KM, Kralik JD. Too good to be true: rhesus monkeys react negatively to better-than-expected offers. PLoS ONE. 2013;8: e75768.
- Topolinski S, Strack F. Corrugator activity confirms immediate negative affect in surprise. Front Psychol. 2015;6:134.
- Koch C, Zika O, Schuck NW. Influence of surprise on reinforcement learning in younger and older adults. PLoS Comput Biol. 2024;20: e1012331.
- Goris J, Brass M, Cambier C, Delplanque J, Wiersema JR, Braem S. The relation between preference for predictability and autistic traits. Autism Res. 2020;13:1144–54.
- Palmer CJ, Lawson RP, Hohwy J. Bayesian approaches to autism: towards volatility, action, and behavior. Psychol Bull. 2017;143:521–42.
- 25. Rutter M, Bailey A, Lord C. Social communication questionnaire. Western Psychological Services. 2003.
- 26. Akechi H, Kikuchi Y, Tojo Y, Hakarino K, Hasegawa T. Mind perception and moral judgment in autism. Autism Res. 2018;11:1239–44.
- Asada K, Akechi H, Kikuchi Y, Tojo Y, Hakarino K, Saito A, Hasegawa T, & Kumagaya S. Longitudinal study of personal space in autism. Child Neuropsychol. 2024;1–9.
- Kikuchi Y, Akechi H, Senju A, Tojo Y, Osanai H, Saito A, Hasegawa T. Attention to live eye contact in adolescents with autism spectrum disorder. Autism Res. 2022;15:702–11.
- Thomas KA, Clifford S. Validity and Mechanical Turk: an assessment of exclusion methods and interactive experiments. Comput Human Behav. 2017;77:184–97.
- Kool W, Gershman SJ, Cushman FA. Cost-benefit arbitration between multiple reinforcement-learning systems. Psychol Sci. 2017;28:1321–33.
- Blair G, Cooper J, Coppock A, Humphreys M, Sonnet L, Fultz N, et al. Package 'estimatr'. Stat. 2018;7:295–318.
- Lüdecke D, Ben-Shachar M, Patil I, Waggoner P, Makowski D. Performance: an R package for assessment, comparison and testing of statistical models. JOSS. 2021;6:3139.
- Katahira K. How hierarchical models improve point estimates of model parameters at the individual level. J Math Psychol. 2016;73:37–58.
- Wilson RC, Collins AG. Ten simple rules for the computational modeling of behavioral data. Elife. 2019;8: e49547.
- 35. Kass RE, Raftery AE. Bayes factors. J Am Stat Assoc. 1995;90:773–95.
- Ghalanos A, Theussl S. Package 'Rsolnp'. Vienna, Austria: R Foundation for Statistical Computing; 2015.
- Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. Bayesian model selection for group studies. Neuroimage. 2009;46:1004–17.
- Danwitz L, Mathar D, Smith E, Tuzsus D, Peters J. Parameter and model recovery of reinforcement learning models for restless bandit problems. Comput Brain Behav. 2022;5:547–63.
- Suzuki S, Katahira K. Applying reinforcement learning to the psychopathology of obsessive-compulsive and gambling disorders: practices and pitfalls in computational model fitting. PsyArXiv. 2024. https://doi.org/10. 31234/osf.io/vj2wp.
- Zhang L, Lengersdorff L, Mikus N, Glascher J, Lamm C. Using reinforcement learning models in social neuroscience: frameworks, pitfalls and suggestions of best practices. Soc Cogn Affect Neurosci. 2020;15:695–707.
- Sumiya M, Katahira K. Commentary: altered learning under uncertainty in unmedicated mood and anxiety disorders. Front Hum Neurosci. 2020;14: 561770.
- Palminteri S, Kilford EJ, Coricelli G, Blakemore SJ. The computational development of reinforcement learning during adolescence. PLoS Comput Biol. 2016;12: e1004953.
- Chierchia G, Soukupova M, Kilford EJ, Griffin C, Leung J, Palminteri S, et al. Confirmatory reinforcement learning changes with age during adolescence. Dev Sci. 2023;26: e13330.
- Nussenbaum K, Hartley CA. Reinforcement learning across development: what insights can we draw from a decade of research? Dev Cogn Neurosci. 2019;40: 100733.
- Sumiya M, Senju A. Brief reports: influence of friendship on loneliness among adolescents with autism spectrum disorders in Japan. J Autism Dev Disord. 2023. https://doi.org/10.1007/s10803-023-05958-z.

- Sumiya M, Igarashi K, Miyahara M. Emotions surrounding friendships of adolescents with autism spectrum disorder in Japan: a qualitative interview study. PLoS ONE. 2018;13: e0191538.
- 47. Picci G, Scherf KS. A two-hit model of autism: adolescence as the second hit. Clin Psychol Sci. 2015;3:349–71.
- Uddin LQ. Brain mechanisms supporting flexible cognition and behavior in adolescents with autism spectrum disorder. Biol Psychiatry. 2021;89:172–83.
- South M, Dana J, White SE, Crowley MJ. Failure is not an option: risk-taking is moderated by anxiety and also by cognitive ability in children and adolescents diagnosed with an autism spectrum disorder. J Autism Dev Disord. 2011;41:55–65.
- Buechel EC, Li R. Mysterious consumption: preference for horizontal (vs. vertical) uncertainty and the role of surprise. J Consum Res. 2023;49:987–1013.
- Jenkinson R, Milne E, Thompson A. The relationship between intolerance of uncertainty and anxiety in autism: a systematic literature review and meta-analysis. Autism. 2020;24:1933–44.
- Buhr K, Dugas MJ. The intolerance of uncertainty scale: psychometric properties of the English version. Behav Res Ther. 2002;40:931–45.
- Bervoets J, Milton D, Van de Cruys S. Autism and intolerance of uncertainty: an ill-fitting pair. Trends Cogn Sci. 2021;25:1009–10.
- Sandhu TR, Xiao B, Lawson RP. Transdiagnostic computations of uncertainty: towards a new lens on intolerance of uncertainty. Neurosci Biobehav Rev. 2023;148: 105123.
- Poli F, Koolen M, Velázquez-Vargas CA, Ramos-Sanchez J, Meyer M, Mars RB, Rommelse N. Hunnius S. Autistic traits foster effective curiosity-driven exploration. PLOS Comput Biol 2024;20:e1012453.
- Harden KP, Tucker-Drob EM. Individual differences in the development of sensation seeking and impulsivity during adolescence: further evidence for a dual systems model. Dev Psychol. 2011;47(3):739.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.