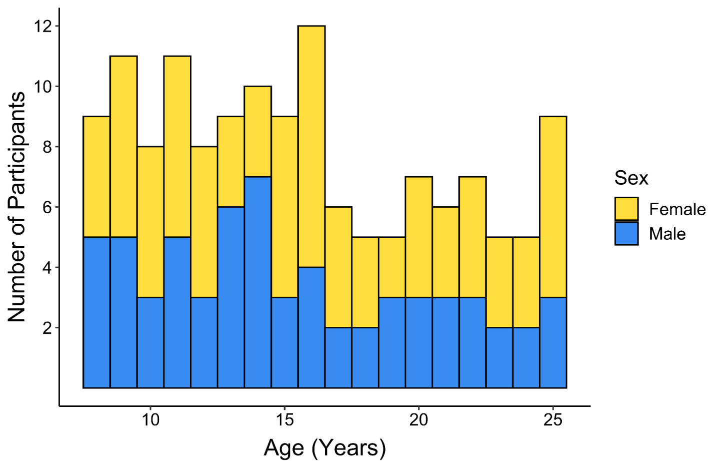
Supplemental Materials for:

Flexibility in valenced reinforcement learning computations across development

Participant demographics and exclusions

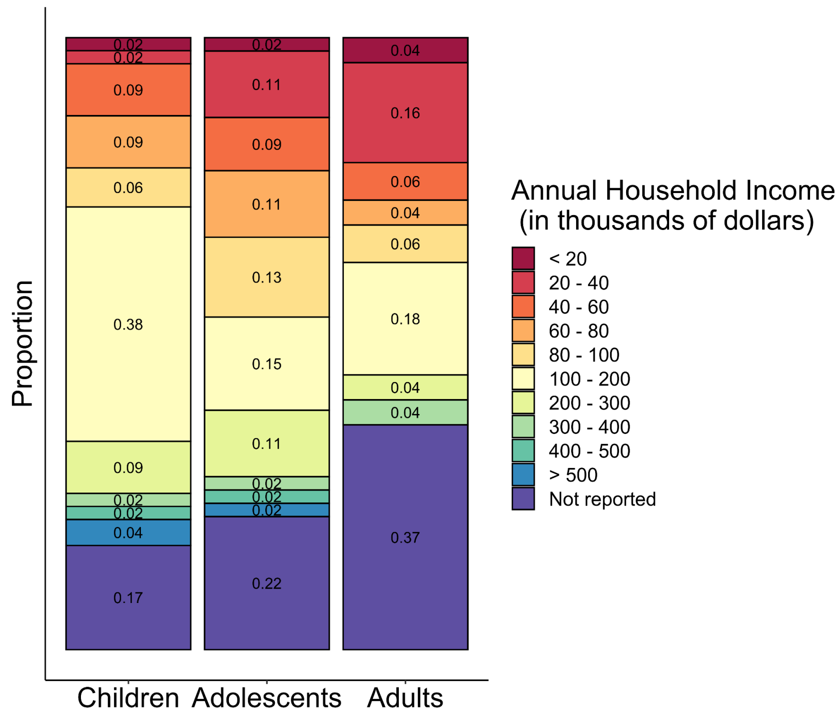
Participant age and sex distribution



Supplemental Figure 1. Participant age and sex distribution. 142 participants between the ages of 8 and 25 years were included in analyses.

Participant annual household income distribution

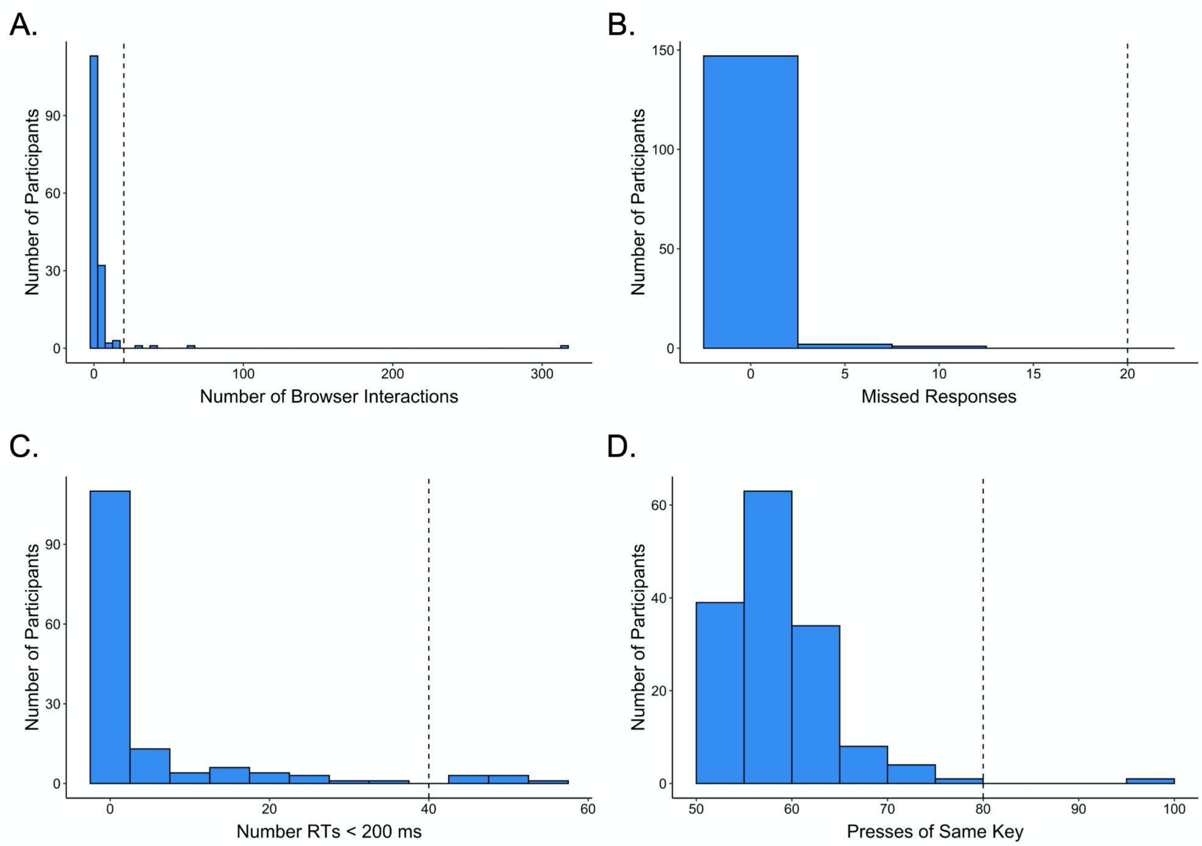
As described in the main text of the manuscript, participants in the study were primarily recruited from ads on Facebook and Instagram (n = 40), via word-of-mouth (n = 28), and through our database for in-lab studies (n = 30), for which we solicit sign-ups at local science fairs and events and through fliers on New York University's campus. Participants' annual household incomes ranged from less than $20,000 to more than $500,000 (Supplemental Figure 2).



Supplemental Figure 2. Participants' annual household incomes. The median household income for child and adolescent participants in our study was between $100,000 and $200,000, and for adult participants was between $60,000 and $80,000. For comparison, in the U.S. in 2020, the median income for family households was $86,372, and for non-family households was $40,464 (U.S. Census).

Participant exclusions

As described in the main text of the manuscript, 154 participants aged 8 - 25 years completed the experiment. Participants were excluded from all analyses if they: a) interacted with their browser window (minimized, maximized, or clicked outside the window) more than 20 times throughout the learning task (n = 4), b) failed to respond on more than 10% of (20) choice trials (n = 0), c) pressed the same key on more than 40% of (80) choice trials (n = 1), or d) responded in less than 200 ms on more than 20% of (40) choice trials (n = 7) (Supplemental Figure 3).



Supplemental Figure 3. Distribution of data quality metrics. The majority of participants in the online study (A) made fewer than 20 'browser interactions', (B) missed fewer than 20 responses, (C) responded in less than 200 ms on fewer than 40 trials, and (D) repeated the same key press fewer than 80 times. Participants who did not meet these data quality thresholds (indicated by dashed vertical lines) were excluded from all analyses.

Explicit Reports of Deck 'Luckiness' and 'Value'

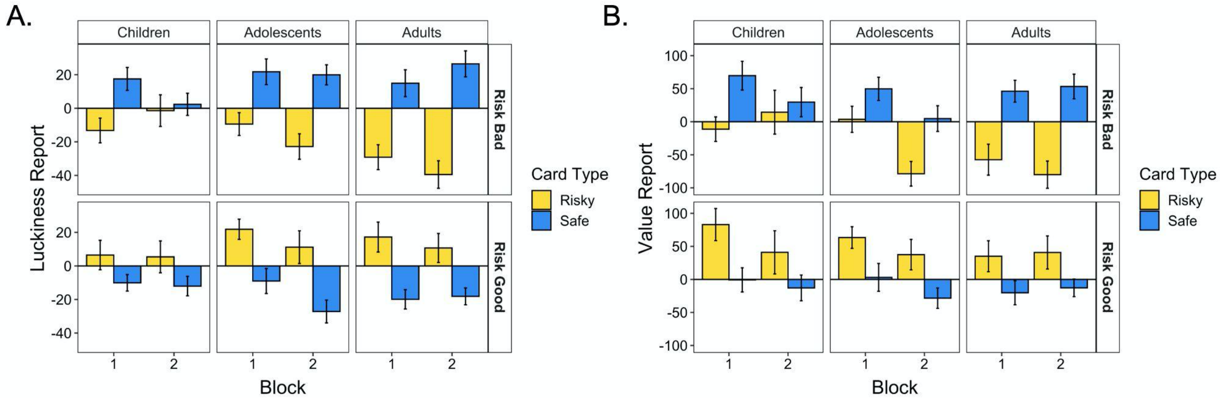
Explicit report methods

After completing each block of 100 choice trials, participants were asked two explicit questions about the 'luckiness' and 'value' value of each deck. Participants were shown each deck and first asked to report how 'lucky' they believed it was by moving a slider. They were asked, "Think about the cards you picked from this deck. You may have picked some bad cards that made you lose tokens and some good cards that made you win tokens. Overall, would you say this is an unlucky deck or a lucky deck?" The far left of the sliding scale was labeled 'very unlucky,' the middle was labeled 'right in the middle' and the far right was labeled 'very lucky.' Participants could position the slider with the mouse and then click to submit their rating, which was converted into a number between -100 (far left) and 100 (far right) based on its position, though these numbers were not visible to participants. After rating the deck's 'luckiness,'' participants were also asked, "On average, how many tokens do you think you received from each card in this deck?" Here, the sliding scale ranged from -300 to 300, with explicit markers every 50 values (e.g., -300, -250, -200, etc.). Participants answered the 'luckiness' and 'value' question for each card, which were presented in a random order.

Explicit report results

We examined how participants transformed their experiential learning into explicit beliefs about the value structure of their learning environments by analyzing their explicit reports. Specifically, we ran linear mixed-effects models with random participant intercepts examining how age, block type, block order, and deck type (risky versus safe) influenced their reports of both the 'luckiness' and 'value' of each card deck. For both dependent variables, we observed a deck type x block type interaction, (*Luckiness:* $F(1, 979.6) = 132.51$, $p < .001$, *Value*: $F(1, 979.4) = 94.07$, $p < .001$) indicating that participants' reports about the decks aligned with the structure of the contexts in which they were encountered; participants rated risky decks as luckier and more valuable in the risk good context and safe decks as luckier and more valuable in the risk bad context (Supplemental Figure 4). These effects, however, were qualified by a significant interaction with age on luckiness reports and a marginal interaction with age on value reports (*Luckiness:* $F(1, 979.7) = 18.46$, $p < .001$, *Value*: $F(1, 979.4) = 3.74$, $p = .054$), such that older participants better differentiated risky and safe decks across contexts. For value reports, we also observed a main effect of age, $F(1, 138.1) = 3.99$, $p = .048$, with younger participants providing higher estimates overall. This effect interacted with deck type, $F(1, 979.4) = 5.44$, $p = .02$, such that differences in explicit ratings across age were greater for risky decks.

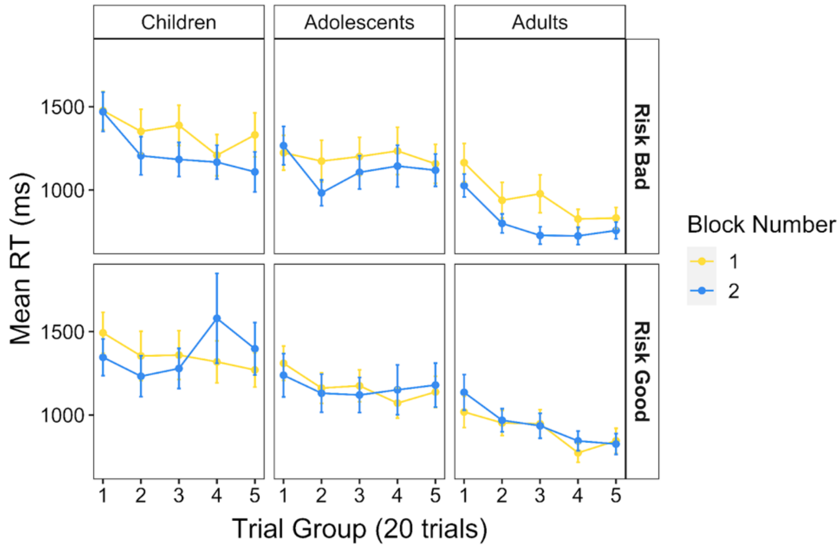
In line with our choice findings, we also observed effects of both block type and block number. Specifically, for luckiness reports, we observed an age x deck type x block number interaction effect, $F(1, 979.7) = 4.65$, $p = .031$, such that younger participants tended to rate safe decks as less 'lucky' in the second block. This further suggests that younger participants' had difficulty learning that safer decks were advantageous after learning in the environment in which riskier decks were advantageous. For value reports, we observed main effects of both block type and block number, with participants providing lower value ratings for the risk bad block, $F(1, 979.4) = 4.72$, $p = .03$, and in the second block, $F(1, 979.4) = 8.14$, $p = .004$.



Supplemental Figure 4. Participants' explicit 'luckiness' (A) and 'value' (B) reports. Participants' reports of each decks' 'luckiness' (A) and 'value' (B) aligned with the structure of the contexts in which they were encountered. Participants rated risky decks as luckier and more valuable in the risk good context and safe decks as luckier and more valuable in the risk bad context, as evidenced by deck type x block type interaction effects (*Luckiness:* $F(1, 979.6) = 132.51$, $p < .001$, *Value*: $F(1, 979.4) = 94.07$, $p < .001$).

Choice Response Times

We examined developmental differences in the speed with which participants chose between the different card deck options throughout the task. Specifically, we ran a linear mixed-effects model examining how response times varied across age, trial, block type, and block number. Our model included random intercepts for each participant and random slopes across trial, block number, and their interactions. As is common in many developmental studies, we observed an age-related decrease in response times, $F(1, 138) = 20.4$, $p < .001$ (Supplemental Figure 5). In addition, response times also decreased across the task — both across trials, $F(1, 138) = 27.03$, $p < .001$, and across blocks, such that participants responded more quickly in the second versus the first block of the task, $F(1, 137.9) = 16.47$, $p <.001$. We did not observe a significant effect of block type on reaction times ($p = .733$). We did, however, observe a significant age x trial x block type x block number interaction effect, $F(1, 138) = 4.62$, $p = .033$ (Supplemental Figure 5). No other effects or interactions were significant (ps > .05).



Supplemental Figure 5. Participant response times. Older participants selected choice options more rapidly than younger participants (p < .001). All participants demonstrated a decrease in response times across trials and across task blocks (ps < .001). The lines show the average response time within each trial group for each age group. Error bars show the standard error across the group means.

Recent work has also suggested that a strong confirmatory bias — learning more from the positive outcomes of one's actions than the negative outcomes — may lead to exaggerated value differences between choice options that promote more rapid decision-making (Lefebvre, Summerfield, & Bogacz, 2022). As such, we tested whether there was a relation between participants' asymmetry indices in each block and their average response time for that block of the task. We ran a linear-mixed effects model with random participant intercepts examining how participant age, asymmetry indices, block type, block number, and their interactions influenced mean response times. We did not, however, observe a significant relation between response times and asymmetry indices, $F(1, 166.9) = 3.37$, $p = .068$.

Effects of Participant Sex on Choice Behavior

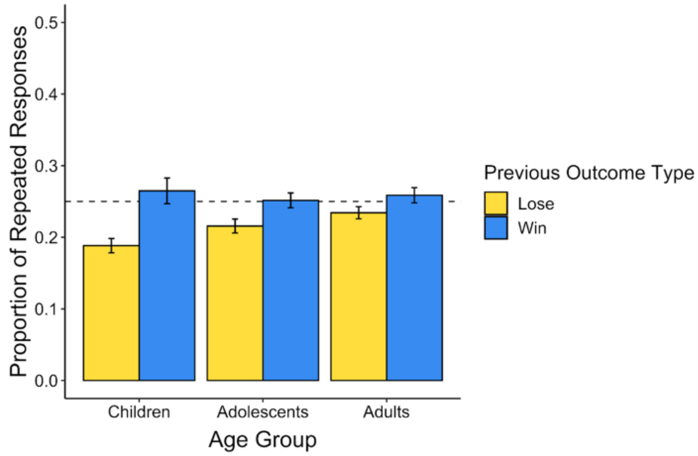
To examine whether choice behavior varied across sex, we first re-ran our mixed-effects model examining optimal choices across trials, with sex as an additional interacting fixed effect. Specifically, we modeled the influence of continuous age, trial number, block type, block number (risk good vs. risk bad), sex, and their interactions on trial-wise optimal choices. We included random participant intercepts and random slopes across trials, block types, and their interaction. We did not observe a significant relation between participant sex and optimal choices, $X^2(1) = 1.44$, $p = .230$. We also did not observe any significant interactions between participant sex and any other predictor variable on choice behavior (ps > .18). The inclusion of sex in the model did not change the pattern of significant effects that we observed; we continued to observe significant main effects of age, trial, and block type (ps < .018), as well as significant two-way interactions between age and trial and age and block number (*p*s < .02), as well as significant age x trial x block number and age x block type x block number interaction effects (*p*s < .02).

We further examined whether there was a relation between participant sex and learning rate asymmetry indices. We ran a mixed-effects model examining asymmetry indices with continuous age, block type, block number, sex, and their interactions as fixed effects, and random participant intercepts. Participant sex was not significantly related to AI, $F(1, 268) = 0.50$, $p = .482$, nor did it significantly interact with any other predictor variables (*p*s > .35). As before, we continued to observe a significant effect of block type on AI ($p = .009$).

Accounting for Motor Perseveration Effects

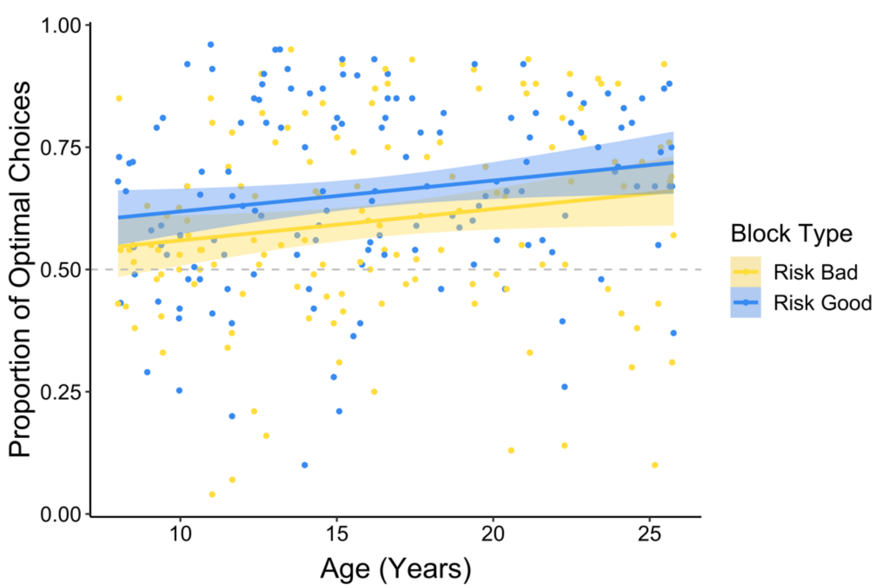
Participants were explicitly instructed that card deck positions did not influence outcome probabilities or magnitudes. However, it is still possible that participant decisions were influenced by the positions of the decks and their corresponding motor responses. To test for an influence of motor perseveration, we ran a logistic mixed-effects model examining how the binary outcome type on the previous trial (coded as 1 for a previous gain outcome and 0 for a previous loss outcome), age, and their interaction influenced whether the participant made the same motor response as they did on the immediately preceding trial. Despite the deck positions being randomized, we found that participants were more likely to repeat the same motor response after a previous win than after a previous loss, *2*(1) = 32.9, $p < .001$. Further, younger participants showed stronger motor perseveration effects, as evidenced by a previous outcome type x age interaction effect, *2*(1) = 10.21, $p = .001$ (Supplemental Figure 6).

To account for these effects in our analysis of asymmetry indices across blocks, we computed a ‘motor perseveration score’ for each participant by subtracting the proportion of trials in which they repeated a motor response following a loss outcome from the proportion of trials in which they repeated a motor response following a win outcome. Participants with higher perseveration scores were more likely to repeat responses following win versus loss outcomes. We then re-ran our analysis of asymmetry indices across blocks, including this control variable. Specifically, we ran a mixed-effects model examining the effects of age, block type, block number, motor perseveration scores and their interactions on asymmetry indices. Critically, while controlling for motor perseveration effects, we continued to observe a robust effect of block type on asymmetry indices, $F(1, 268) = 7.41$, $p = .007$.

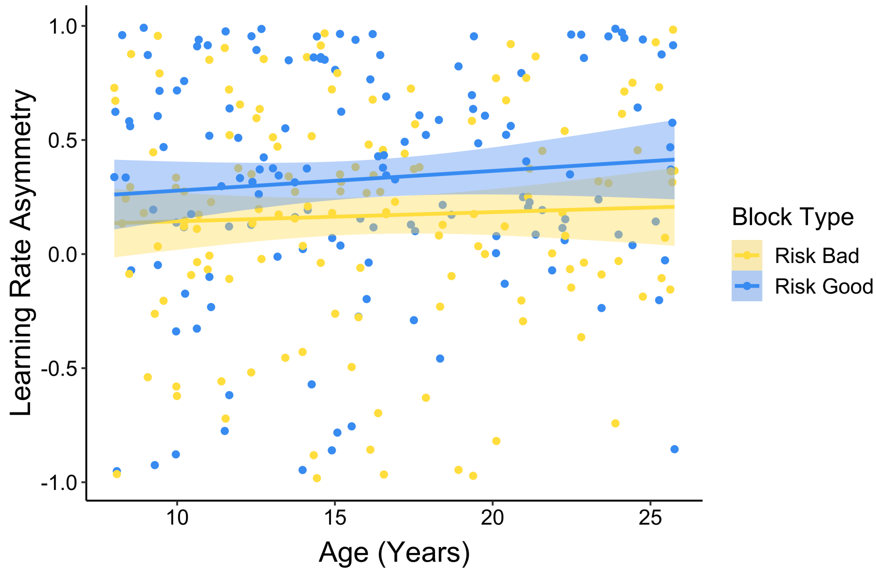


Supplemental Figure 6. Motor perseveration effects. Though deck positions were randomized on every trial and did not relate to outcome probabilities or magnitudes, participants were more likely to repeat position selections following a win outcome on the previous trial than following a loss outcome ($p < .001$). Younger participants showed stronger motor perseveration effects relative to older participants ($p = .001$).

Supplemental Figures with Continuous Age



Supplemental Figure 7. Optimal choices by continuous age. Across age, participants made more optimal choices in the risk good block relative to the risk bad block ($p = .022$). Older participants made a higher proportion of optimal choices relative to younger participants ($p = .011$). Points on the plot indicate individual participants' average proportion of optimal choices in each block of the task. The lines indicate the best-fitting linear regression lines for each block, and the shaded regions around them represent the 95% confidence intervals.

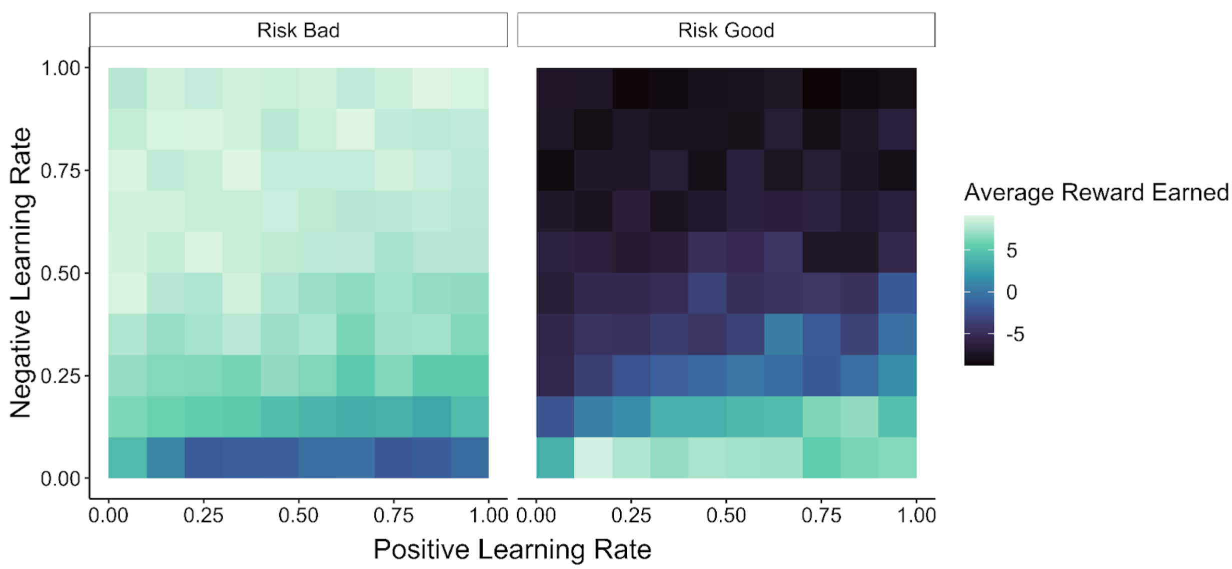


Supplemental Figure 8. Asymmetry indices by continuous age. Across age, participants demonstrated more positive learning rate asymmetry indices in the risk good relative to the risk bad block of the task (*p* = .007). Points on the plot indicate individual participants' asymmetry indices for each block of the task. The lines indicate the best-fitting linear regression lines for each block, and the shaded regions around them represent the 95% confidence intervals.

Modeling methods and results

Simulated performance across different learning rates

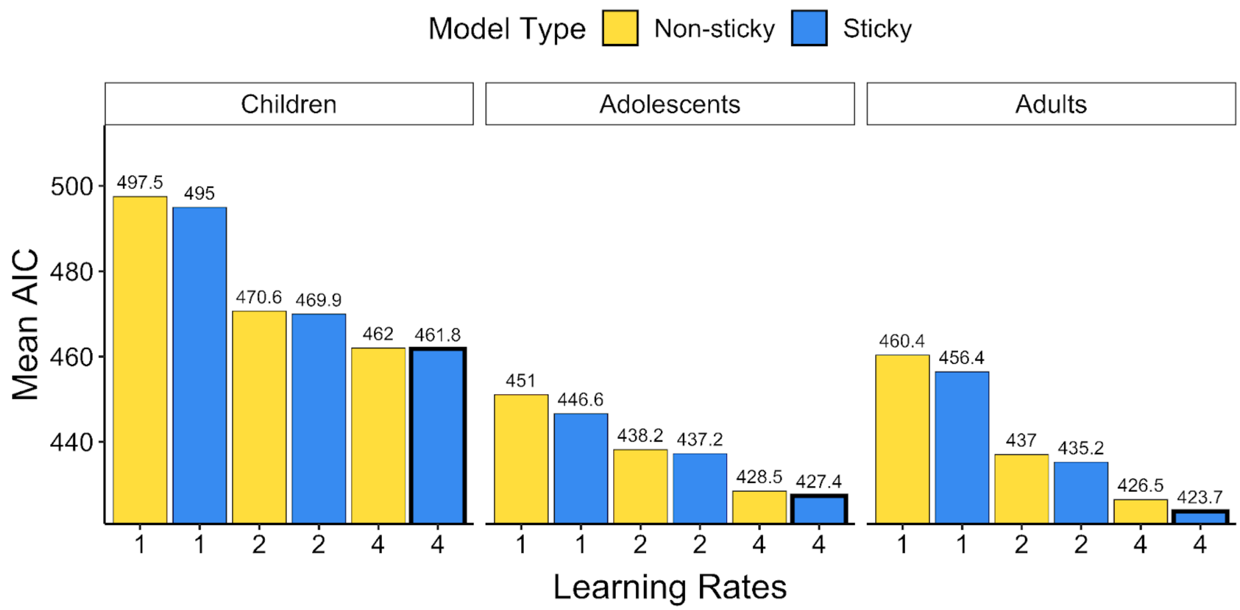
We designed our task such that higher positive versus negative learning rates would be beneficial in the risk good block and higher negative versus positive learning rates would be beneficial in the risk bad block. To confirm that this was in fact the case, we simulated the performance of 100 agents with 100 different combinations of learning rates (10,000 agents total). In these simulations, positive and negative learning rates were drawn from 10 evenly spaced values between .05 and .95, and inverse temperatures were drawn randomly for each agent from the uniform distribution [2, 30]. For simplicity, we set the stickiness parameter to 0. We then computed the average number of points earned by agents with each combination of learning rates (Supplemental Figure 9). These simulations confirmed that indeed it was more advantageous to have a more negative learning rate asymmetry in the risk bad block and a more positive learning rate asymmetry in the risk good block.



Supplemental Figure 9. Simulated performance of agents with different learning rates. Simulations of 10,000 agents with 100 different learning rate combinations confirmed that participants with higher negative learning rates performed better in the risk bad context and participants with lower negative learning rates and higher positive learning rates performed better in the risk good context.

Model comparison: sticky versus non-sticky models

As mentioned in the main text of the manuscript, we tested whether the inclusion of a stickiness parameter improved the fit of the one, two, and four learning rate models. In all cases, the inclusion of the stickiness parameter improved model fit both across the entire sample and within each age group (Supplemental Figure 10).

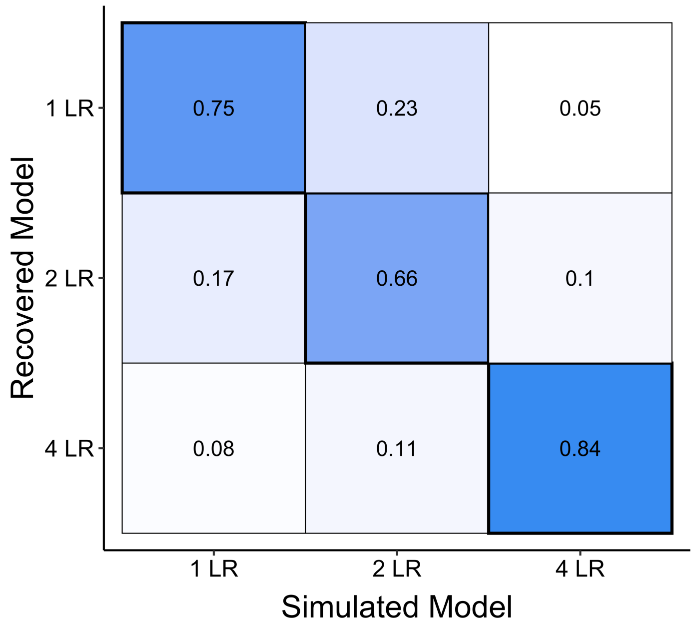


Supplemental Figure 10. Across age groups, mean AIC values were lower for models that included a stickiness parameter, indicating that accounting for perseverative choices improved model fit.

Model recovery

To confirm that the one, two, and four learning rate models were distinguishable from one another, we simulated data from 500 participants for each model. For each model, we randomly drew parameter values from distributions covering the full range of observed parameter values (β: U ~ (.15, 30), ϕ ~ $t_{50,}$ all variants of α ~ U(0, 1)) that we obtained when we fit the models to our real data (Wilson & Collins, 2019). We then fit each of these datasets with each of the three models and examined the proportion of participants from each generating model best fit by each model, according to AIC values (Supplemental Figure 11).

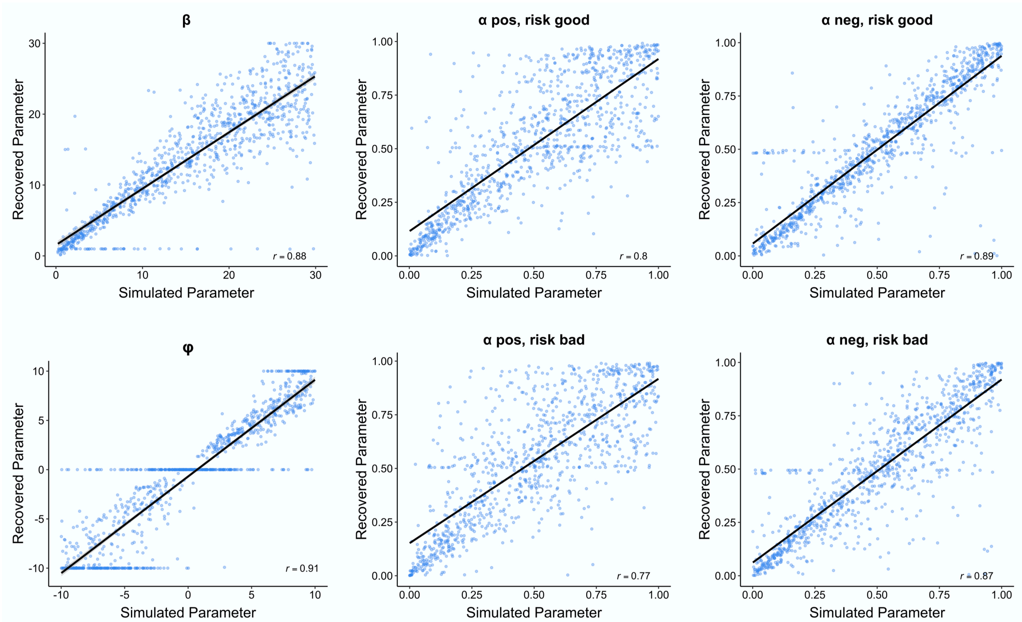
In general, we recovered the 'correct' model for the majority of simulations. Some participants simulated with the two learning rate models were best fit by the one learning rate model. Because we generated parameter values randomly, many of these participants were simulated with similar positive and negative learning rates. Thus, the additional parameter required to capture different positive and negative learning rates was not needed to capture these participants' choices.



Supplemental Figure 11. Model recovery results. The majority of the 1,500 simulated participants were best fit by the same model that was used to generate their choice data.

Parameter recovery

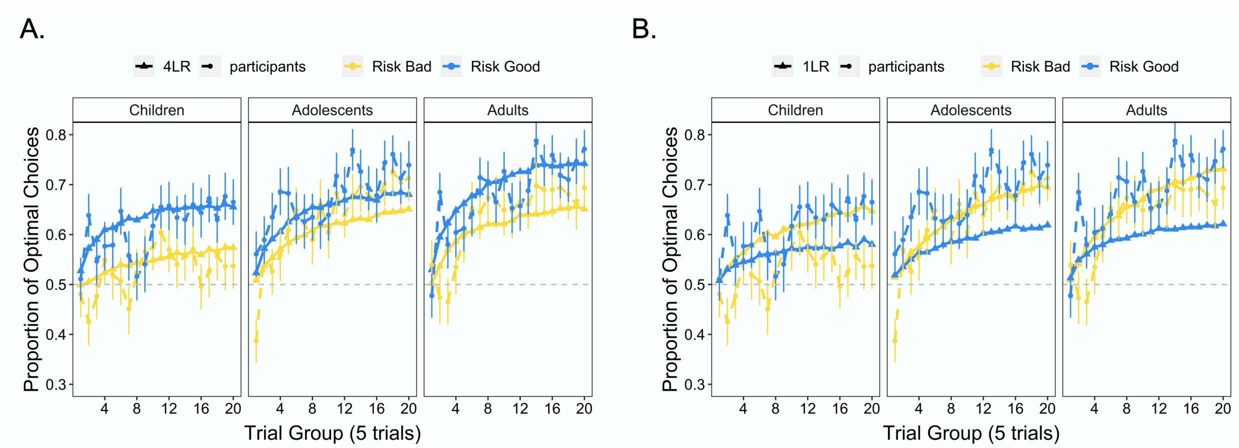
We simulated 1,000 participants with the four learning rate model, with parameters randomly drawn from uniform distributions: β: U ~ (.15, 30), ϕ ~ U(-10, 10),, all variants of α ~ U(0, 1). We examined the correlation between the 'true' generating parameters that we used to simulate the data and the fitted parameter values (Supplemental Figure 12). For all parameters, the correlation between the 'true' simulated parameter value and the recovered parameter value was high (>= .77). This suggests that the fitted parameter values from this model are interpretable, as they are largely reflective of those that governed the learning process.



Supplemental Figure 12. Parameter recovery results for the four learning rate model. Correlations between simulated and recovered parameter values for the four learning rate model ranged from .77 to .91.

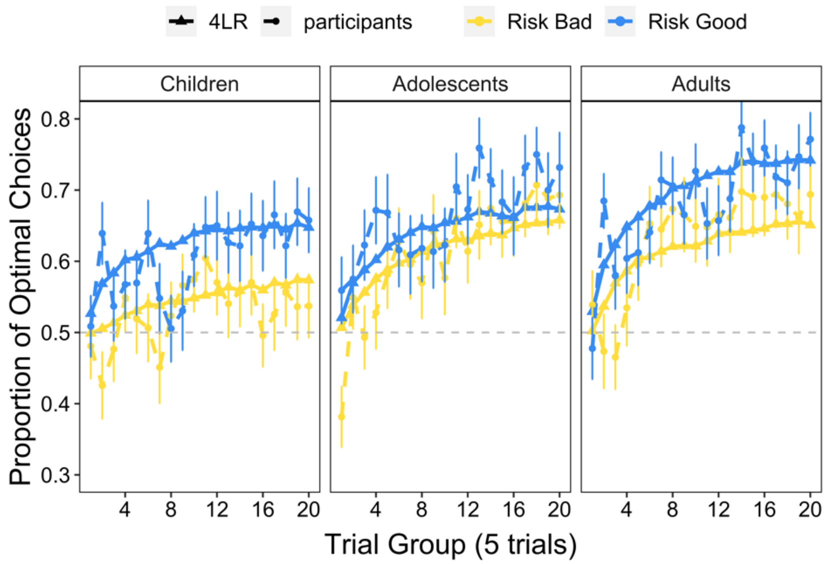
Posterior predictive check

We also examined whether the best-fitting model, the four learning rate model, captured qualitative features of participants' real choices. To do so, we simulated 100 repetitions of each participant's parameter combination and block order. Thus, we simulated data from 14,200 participants (142 parameter combinations x 100 repetitions of each). We then examined the proportion of optimal choices generated by each simulation, broken down into the age group from which the parameter combination was drawn (Supplemental Figure 13). We compared these proportions to those from our real data set. In general, the four learning rate model captured many, but not all, of the qualitative features of real participants' choices, particularly in contrast to the one learning rate model. Specifically, the four learning rate model reproduced the effects we observed of age and block type, with older participants' outperforming younger participants, and all participants performing better in the risk good relative to the risk bad block.



Supplemental Figure 13. Four and one learning rate model posterior predictive checks. A.) Data generated from the four learning rate model using participants' best-fitting parameters, plotted with the solid lines, captured many of the qualitative features of participants' real choice data, plotted with the dashed lines. B.) Data from the one learning rate model, also generated using participants' best fitting parameters, did not capture many of the qualitative features of participants' choice data. One learning rate model simulations revealed better performance in the risk bad versus the risk good block, whereas participants tended to show the opposite pattern of performance.

Though the four learning rate model simulations captured many of the qualitative features of participants' patterns of choice behavior, the model predictions in the 'risk bad' block systematically underestimated adolescents' and adults' proportion of optimal choices. To better understand the discrepancy between the empirical data and the model predictions, we examined each individuals' pattern of choice behavior as well as the model simulations that were generated with their fitted parameters. In doing so, we found that five participants (1 child and 4 adolescents) selected the same choice option on 90 or more of the 100 trials within at least one of the two task blocks. Because the positions of the choice options were randomly shuffled on each trial, this pattern of behavior is likely *not* a sign of inattention but rather indicative of a particular decision-making strategy. However, due to the lack of variance in their choices, the model predictions generated with these participants' parameters did not well-capture their behavior. When we exclude these participants from our posterior predictive check, we find that the adolescent model predictions more closely match the adolescent behavioral data (Supplemental Figure 14). Importantly, when we exclude these five participants from our analysis of asymmetry indices, we continue to observe a robust effect of block type on AI, $F(1, 266) = 7.35$, $p = .007$, indicating that our findings hold when we do not include participants whose behavior may be poorly captured by the four learning rate model.



Supplemental Figure 14. Four learning rate model posterior predictive check with participant exclusions. Data generated from the four learning rate model using participants' best-fitting parameters, plotted with the solid lines, captured many of the qualitative features of participants' real choice data, plotted with the dashed lines. Here, participants who selected the same choice option on 90 or more trials within a block were excluded from both the empirical and simulated data.

Finally, we note that even with these participant exclusions, the model simulations continue to systematically underestimate the adults' performance in the risk bad task block. One possibility is that adults' behavior is better captured by a model with a dynamic learning rate that *changes* throughout the task block. We discuss this possibility in much more detail in the next section of the supplement.

Dynamic learning rate models

In addition to the three static learning rate models described in the main text of the manuscript, we also fit nine variants of a decaying learning rate model to the data, to account for changes in the extent to which individuals updated their value estimates following positive and negative prediction errors over the course of each task block.

Dynamic model specification

The simplest instantiation of the decaying learning rate model includes two different learning rate parameters — an initial learning rate (α_{init}) and a decay rate (η). On every trial (*t*) within each block, the learning rate (α_t) is determined by:

$$\alpha_t = \frac{\alpha_{init}}{1 + \eta * (t - 1)}$$

where η is a free parameter that ranges from 0 to 1 and determines the rate at which the learning rate decays.

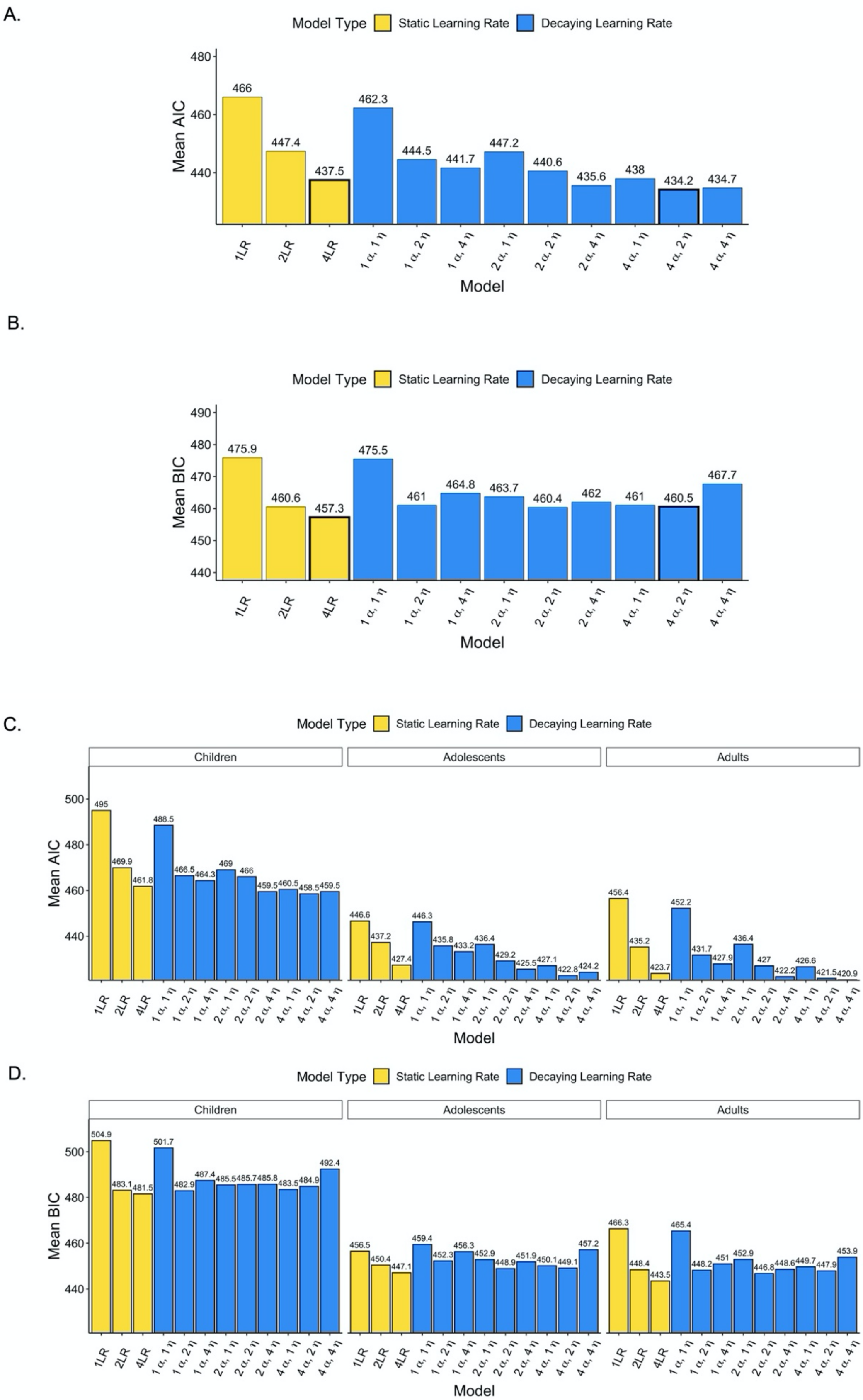
To account for potential differences in positive and negative initial learning rates and decay rates across task contexts, we fit nine variants of this model with one, two (positive and negative), and four (positive and negative for each block) different α_{init} parameters, and one, two, and four different η parameters. In total, we tested nine variants of the decaying learning rate model, which included every combination of different numbers of α_{init} and η.

As with the static learning rate models described in the main text of the manuscript, value estimates were converted to choice probabilities via a softmax function with an inverse temperature parameter and a stickiness parameter.

We applied the following bounds and priors to all variants of α_{init} and η: bounds: [0, 1]; prior: beta(1.1, 1.1).

Static and dynamic model comparison results

We then compared the fits of these dynamic learning rate models to those of our static learning rate models across our entire sample and within each age group (Supplemental Figure 15). Across the entire sample, the best-fitting model was the dynamic learning rate model with four initial learning rates and two decay rates.

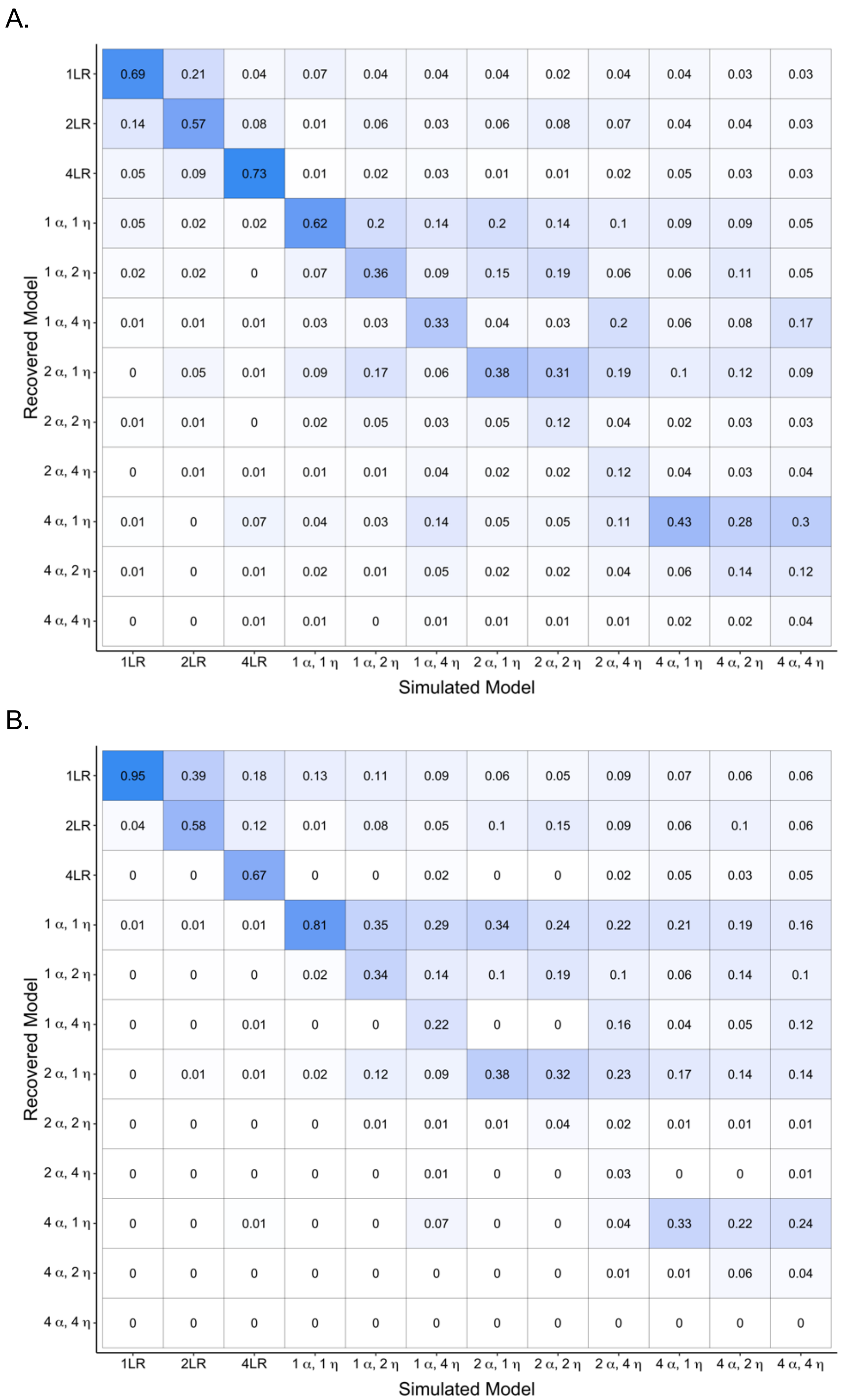

A.
Model Type
Static Learning Rate
Decaying Learning Rate
Mean AIC
480
460
440
466
447.4
437.5
462.3
444.5
441.7
447.2
440.6
435.6
438
434.2
434.7
1LR
2LR
4LR
1 α, 1 η
1 α, 2 η
1 α, 4 η
2 α, 1 η
2 α, 2 η
2 α, 4 η
4 α, 1 η
4 α, 2 η
4 α, 4 η
Model
B.
Model Type
Static Learning Rate
Decaying Learning Rate
Mean BIC
490
480
470
460
450
440
475.9
460.6
457.3
475.5
461
464.8
463.7
460.4
462
461
460.5
467.7
Model
C.
Model Type
Static Learning Rate
Decaying Learning Rate
Children
Adolescents
Adults
Mean AIC
500
480
460
440
495
469.9
461.8
488.5
466.5
464.3
469
466
459.5
460.5
458.5
459.5
446.6
437.2
427.4
446.3
435.8
433.2
436.4
429.2
425.5
427.1
422.8
424.2
456.4
435.2
423.7
452.2
431.7
427.9
436.4
427
422.2
426.6
421.5
420.9
Model
D.
Model Type
Static Learning Rate
Decaying Learning Rate
Children
Adolescents
Adults
Mean BIC
500
480
460
440
504.9
483.1
481.5
501.7
482.9
487.4
485.5
485.7
485.8
483.5
484.9
492.4
456.5
450.4
447.1
459.4
452.3
456.3
452.9
448.9
451.9
450.1
449.1
457.2
466.3
448.4
443.5
465.4
448.2
451
452.9
446.8
448.6
449.7
447.9
453.9
Model

Supplemental Figure 15. Model comparison across static and dynamic models. (A) Across participants, average AIC values were lowest for the four-alpha-two-eta model, indicating that participants used different initial learning rates for both better-than-expected and worse-than-expected outcomes and across task blocks, and that these initial learning rates decayed at different rates for better- and worse-than-expected outcomes. (B) However, average BIC values were lowest for the static four learning rate model, suggesting that the winning dynamic model may be overly complex for the data. (C) Average AIC values within each age group indicated that children and adolescents were best fit by the four-alpha-two-eta model and adults were best fit by a four-alpha-four-eta model. (D) Average BIC values within each age group suggested that participants across ages were best fit by the static four learning rate model.

Static and dynamic model recovery

As with our static learning rate models, to confirm that these models were distinguishable from one another, we simulated data from 500 participants for each model. For each model, we randomly drew parameter values from distributions covering the full range of observed parameter values (β: U ~ (.15, 30), ϕ ~ $t_{50,}$ all variants of α ~ U(0, 1), all variants of η ~ U(0, 1)) that we obtained when we fit the models to our real data (Wilson & Collins, 2019). We then fit each of these twelve datasets with each of the twelve models and examined the proportion of participants from each generating model best fit by each model, according to AIC and BIC values (Supplemental Figure 16).

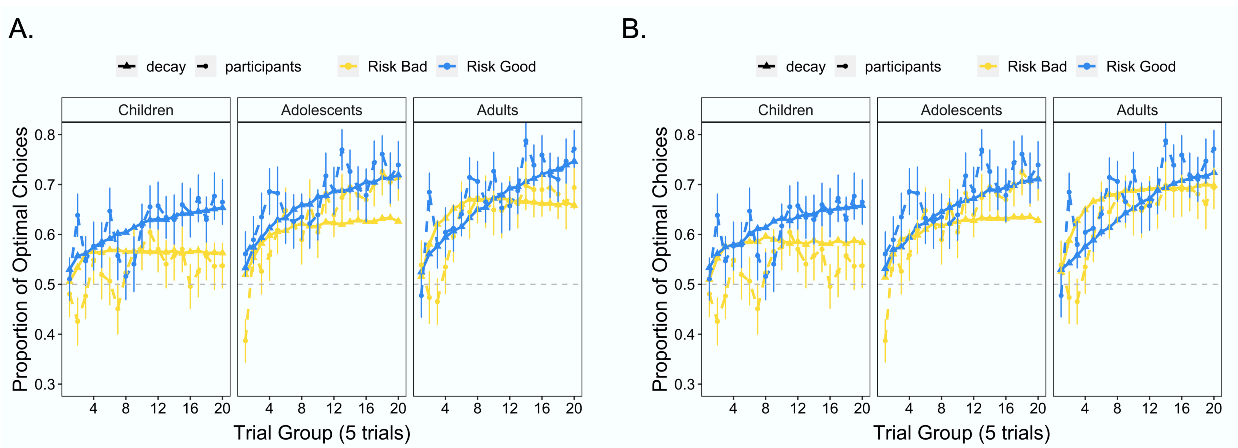
In general, we were able to recover the ‘correct’ model for the majority of static learning rate model simulations, but *not* for the majority of dynamic learning rate model simulations, suggesting that most of these models were overparameterized for the data we collected. A different task design is likely needed to distinguish the nine dynamic learning rate models we tested. Due to the poor recoverability of these models, we focused the main text of our manuscript on the static learning rate models. However, for completeness, we include posterior predictive checks and analyses of the parameter estimates derived from the two best-fitting dynamic models (four-alpha-two-eta and four-alpha-four-eta) in the subsequent sections of our supplement.



Supplemental Figure 16. Model recovery results for full set of models. The majority of dynamic-learning-rate model simulated participants were not best fit by the same model that generated their data, both when we used minimum AIC values (A) and BIC values (B) to determine the best-fitting model for each participant, suggesting that these models were likely overparameterized for the data that we collected.

Dynamic learning rate models posterior predictive checks

We examined whether the two best-fitting decaying learning rate models, the four-alpha-two-eta model and the four-alpha-four-eta model, captured qualitative features of participants' real choices. To do so, we simulated 100 repetitions of each participant's parameter combination and block order. Thus, we simulated data from 14,200 participants (142 parameter combinations x 100 repetitions of each) for each model. We then examined the proportion of optimal choices generated by each simulation, broken down into the age group from which the parameter combination was drawn (Supplemental Figure 17). We compared these proportions to those from our real data set. In general, the decaying learning rate models well-captured the choices of children and adults, though systematically underestimated adolescents' performance in the risk bad block, to an even greater extent than the static learning rate models.



Supplemental Figure 17. Decaying learning rate models posterior predictive checks. A.) Data generated from the four-alpha-two-eta model using participants' best-fitting parameters, plotted with the solid lines, well captured child and adult choices. B.) Data generated from the four-alpha-four-eta model provided an even better approximation of adult behavior, particularly in the risk bad block.

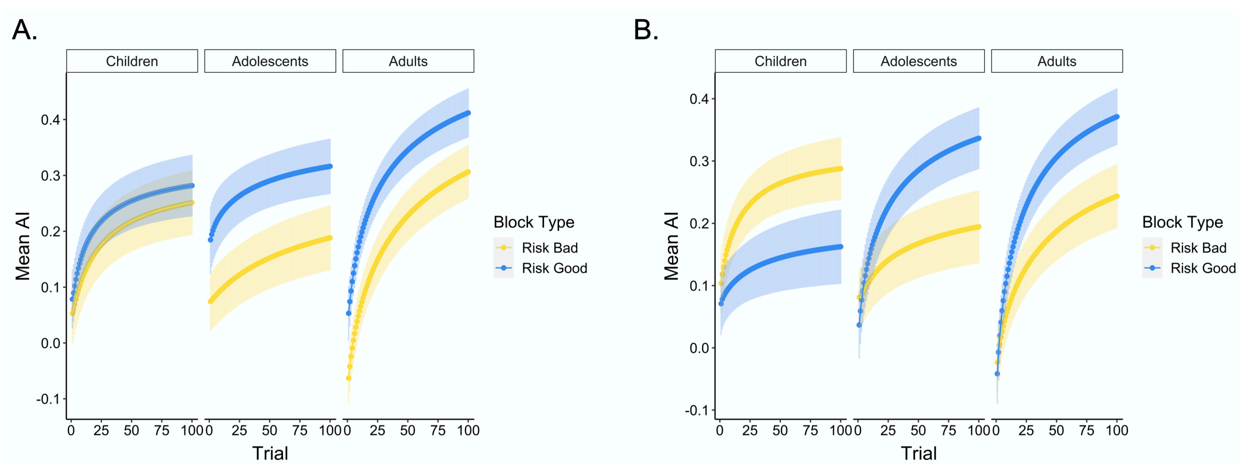
Dynamic learning rate asymmetry findings

The primary goal of our experiment was to determine whether participants across a wide age range adapted their learning rates to the demands of different task contexts. As we note in the discussion of our manuscript, though we focused our analysis on models with static learning rates, this process likely unfolded dynamically across each task block. Thus, we also examined participant learning rate asymmetries using learning rates derived from parameter estimates from the two best-fitting dynamic learning rate models.

Four-alpha-two-eta results. We first examined how asymmetry indices derived from the four-alpha-two-eta model varied across age, trial, block type, and block number. We ran a mixed-effects model predicting trial-wise asymmetry indices from these four predictors, as well as their interactions. Our model included random intercepts for each participant, as well as random slopes across trial, block type and block number. We observed a main effect of trial, $F(1, 138) = 25$, $p < .001$, such that asymmetry indices became more positive over the course of each task block (Supplemental Figure 18). This effect was modulated by age, $F(1, 138) = 3.92$, $p = .05$, such that older participants showed greater increases in asymmetry indices across trials relative to younger participants. In line with our hypothesis, we observed a significant block type x trial interaction effect, $F(1, 56370) = 32.4$, $p < .001$, such that asymmetry indices became more positive in the risk good block relative to the risk bad block. Here, we also observed an age x block type x trial interaction effect, $F(1, 56370) = 26.6$, $p < .001$, such that older participants demonstrated larger differences in AI between the two task blocks across trials. Finally, we also observed effects of block *number* — specifically, we observed a significant trial x block number interaction effect, $F(1, 56370) = 181.1$, $p < .001$, as well as an age x trial x block number interaction effect, $F(1, 56370) = 15.1$, $p < .001$, with younger participants showing greater increases in asymmetry indices across trials in the second relative to the first block of the task.

Four-alpha-four-eta results. We repeated the same analysis with asymmetry indices derived from the four-alpha-four-eta model and observed identical patterns of significant and non-significant effects (Supplemental Figure 18).

These findings largely support the conclusions we drew from our static learning rate models: Though participants showed an overall bias to update their beliefs to a greater extent following better-than-expected versus worse-than-expected outcomes, they adapted the extent to which they learned from valenced feedback based on what promotes optimal decision-making in different contexts. Further, older participants demonstrated greater adaptability, adjusting their learning rates more optimally based on the statistics of the two learning environments.



Supplemental Figure 18. Asymmetry indices derived from decaying learning rate models. A.) Asymmetry indices derived from the four-alpha-two-eta model changed differently across trials within each block ($p < .001$), becoming more positive in the risk good block, particularly for older participants ($p < .001$). Points represent age group averages within each block. The shaded region around each group of points shows the standard error of the mean. B.) Asymmetry indices derived from the four-alpha-four-eta model showed similar patterns.