

RESEARCH ARTICLE

Decrease in decision noise from adolescence into adulthood mediates an increase in more sophisticated choice behaviors and performance gain

Vanessa Scholz^{1,2*}, Maria Waltmann^{1,3}, Nadine Herzog^{3,4}, Annette Horstmann^{3,4,5}, Lorenz Deserno^{1,3,4,6*}

1 Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Centre of Mental Health, University of Würzburg, Würzburg, Germany, **2** Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, the Netherlands, **3** Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, **4** IFB Adiposity Diseases, Leipzig University Medical Center, Leipzig, Germany, **5** Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland, **6** Department of Psychiatry and Psychotherapy, Technical University Dresden, Dresden, Germany

* scholz_v@ukw.de (VS); deserno_l@ukw.de (LD)



OPEN ACCESS

Citation: Scholz V, Waltmann M, Herzog N, Horstmann A, Deserno L (2024) Decrease in decision noise from adolescence into adulthood mediates an increase in more sophisticated choice behaviors and performance gain. *PLoS Biol* 22(11): e3002877. <https://doi.org/10.1371/journal.pbio.3002877>

Academic Editor: Matthew F. S. Rushworth, Oxford University, UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND

Received: June 26, 2024

Accepted: October 2, 2024

Published: November 14, 2024

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pbio.3002877>

Copyright: © 2024 Scholz et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Learning and decision-making undergo substantial developmental changes, with adolescence being a particular vulnerable window of opportunity. In adolescents, developmental changes in specific choice behaviors have been observed (e.g., goal-directed behavior, motivational influences over choice). Elevated levels of decision noise, i.e., choosing suboptimal options, were reported consistently in adolescents. However, it remains unknown whether these observations, the development of specific and more sophisticated choice processes and higher decision noise, are independent or related. It is conceivable, but has not yet been investigated, that the development of specific choice processes might be impacted by age-dependent changes in decision noise. To answer this, we examined 93 participants (12 to 42 years) who completed 3 reinforcement learning (RL) tasks: a motivational Go/NoGo task assessing motivational influences over choices, a reversal learning task capturing adaptive decision-making in response to environmental changes, and a sequential choice task measuring goal-directed behavior. This allowed testing of (1) cross-task generalization of computational parameters focusing on decision noise; and (2) assessment of mediation effects of noise on specific choice behaviors. Firstly, we found only noise levels to be strongly correlated across RL tasks. Second, and critically, noise levels mediated age-dependent increases in more sophisticated choice behaviors and performance gain. Our findings provide novel insights into the computational processes underlying developmental changes in decision-making: namely a vital role of seemingly unspecific changes in noise in the specific development of more complex choice components. Studying the neurocomputational mechanisms of how varying levels of noise impact distinct aspects of learning and decision processes may also be key to better understand the developmental onset of psychiatric diseases.

Data Availability Statement: Data and code for analyses and figures included in this manuscript are available publicly via <https://osf.io/mcx36/>.

Funding: This work was directly funded by a grant to L.D. and A.H. from the IFB Adiposity Diseases, Federal Ministry of Education and Research (BMBF: <https://www.bmbf.de>), Germany, GN: 01EO150. LD also receives funding from the German Research Foundation (DFG: <https://www.dfg.de/>) as part of the Collaborative Research Centre 265 (Project A02, Project Number: 402170461) and on ADHD (DE 2509/3-1, Project number 533682086) as well as by the BMBF on the computational foundations of internalizing versus externalizing symptoms (01GQ2302B), which partially supported this work. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: ADHD, attention-deficit hyperactivity disorder; HBI, hierarchical Bayesian inference; MB, model-based; PXP, protected exceedance probability; RL, reinforcement learning.

Introduction

Learning and decision-making change considerably from adolescence into adulthood [1–3]. Adolescents have to learn how to navigate a constantly changing environment and to strike a balance between exploring new outlets and sticking with already known “good” choices [4]. They do so while their brain undergoes substantial maturational changes that particularly affect cognitive control, value-based learning, and choice [3,5,6]. As yet, the neurocomputational processes accompanying these developmental challenges remain insufficiently understood [4]. Here, reinforcement learning (RL) models provide a computational framework to test hypotheses about latent processes underlying the development of learning and decision-making [7], e.g., learning rates or sensitivity to different outcomes, which are not accessible by the analysis of overt behavior.

A focus in the RL literature has been on the development of specific learning and choice signatures [3,8–18], especially during adolescence, a period frequently considered to be a “window of opportunity” for exploring new choice options and outlets [19]. For example, it was demonstrated that goal-directed behavior (model-based (MB) control over choice), increases from adolescence to adulthood [8–15]. Reward-based cognitive flexibility, as measured by reversal learning paradigms, improves with age. Another example is the influence of motivation on choices [17,20]. Individuals can be strongly biased to respond with behavioral activation to obtain reward, while punishment tends to facilitate behavioral inhibition [21–24]. This so-called Pavlovian choice bias has been found to be reduced in adolescence relative to childhood and adulthood (i.e., weaker behavioral activation to rewards and lower behavioral inhibition to punishments), which was suggested to result from an elevated exploration in adolescents [17].

In the recent past, it was noted that such specific developmental changes in learning and choice tend to be inconsistent across studies [25] and generalize rather poorly [4]. Likewise, some key parameters of RL models, e.g., the learning rate, did not generalize across tasks, which may reflect specific and necessary adaptation to experimental environments [4,26]. Slight differences in experimental environments (e.g., frequency of changes in reward contingencies) impact results, such that performance may be improved or hampered in youths as compared to older adults (compare results from reversal learning in [16] to [26]). On the other hand, one very consistent developmental finding across different tasks is that adolescents show increased levels of decision noise (or choice stochasticity [16,27], for review [4,28]). Decision noise describes a decoupling of learned values from action selection and leads to increased variability in choice behavior. This leads to the selection of less optimal or lower-valued options and usually does not optimize reward outcomes. However, such behavior could, in principle, be regarded as explorative [29]. Yet, in most tasks discussed so far, this noisy or random exploration cannot be clearly distinguished from directed information seeking. The latter aims at explicit information gain, a hallmark for more sophisticated styles of exploration [30]. Thus, decision noise has often been regarded as less interpretable, potentially even reflecting incomppliance to experimental designs.

One unexplored possibility is that the development of specific and more sophisticated choice behaviors may be related or even depend on individual levels of decision noise. To study this unaddressed question, the central aim of this study was to shed light on the role of developmentally elevated decision noise within and across 3 RL tasks. We collated and (re)analyzed data from a developmental sample that had completed the 3 RL tasks as part of a larger study protocol [8,16,23,31–33]. We set the stage by replicating specific developmental effects from a motivational Go/NoGo task capturing decision biases. Using a modified implementation of noise in an RL model of the motivational Go/NoGo task, we show that, in line with

previous work, decision noise may reflect a consistent (rather stable within-subject) signature characteristic for an individual [26]. Critically, we then set out to test our main novel proposal of mediating effects of such seemingly unspecific decision noise on developmental changes in specific choice signatures and corresponding performance readouts from all 3 RL tasks. We hypothesized that noise levels might mediate the relationship between age and specific developmental task effects. This expectation was based on previous findings of developmental changes in noise levels [26,34] as well as evidence suggesting that noise undermines core cognitive processes such as information updating and integration, both crucial for learning and decision-making [35].

Results

We analyzed data from 93 participants between 12 and 42 years of age (age mean [SD] = 22.65 [7.88], female: 45; male: 48) who completed 3 reinforcement learning tasks as part of a larger protocol [36]: (1) a modified motivational Go/NoGo task capturing Pavlovian choice and instrumental learning bias [23,31]; (2) a probabilistic reversal learning task measuring feedback-based cognitive flexibility [16] and a modified sequential (“two-step”) decision-making task to assess model-based control [8,32,37].

Specific age-dependent developmental changes

We employed a recent variant of a motivational Go/NoGo task designed to tease apart a non-selective Pavlovian influence (i.e., behavioral activation to reward as compared to inhibition to punishment) from selective instrumental influence (i.e., repetition of selective actions after reward and a selective avoidance after punishment [23,31] on choice behavior. For an in-depth task description, refer to the methods and supplement information [SI] (S1 Text and S2 Fig). Using mixed-effects models, we replicated an age-dependent increase of Pavlovian biases from adolescence into adulthood [17]: age predicted more activation (Go responses) to reward cues and more inhibition (NoGo responses) to punishment cues (age \times valence: $\beta = 0.107$, SE = 0.05, $\chi^2(1) = 5.25$, $p = 0.02$) (Fig 1D). This was accompanied by overall better performance with higher age ($r_s(91) = 0.24$, $p = 0.021$, overall task accuracy: mean [SD] = 66.24 [0.18], median [SD] = 69.69 [0.18]). Meanwhile, the selective impact of instrumental learning biases on choices did not relate to age (age \times action taken \times outcome valence \times outcome salience: $\beta = 0.033$, SE = 0.03, $\chi^2(1) = 1.73$, $p = 0.2$) (Fig 1E) (see S1 Text and S1–S5 Tables for more details, model statistics and control analyses regarding gender, which did not influence results).

Computational modeling

As in previous publications, all models included a learning rate. We iteratively added a Gobias parameter (capturing individual’s general tendency to make a Go response) and a Pavlovian bias parameter (capturing effects of reward versus punishment cues on activation Go versus inhibition NoGo). All models included either 1 noise parameter across all trials or 2 independent noise parameters for win versus punishment context. Partially in accordance with our previous work [16], we added a computational model with 2 separate feedback sensitivity parameters capturing the degree of noise (i.e., stochasticity) based on positive or negative outcome valence (ρ_{+FB} and ρ_{-FB}). Of note, feedback sensitivities are very similar to the inverse temperature parameter known from other RL model implementations [38–40], as they determine how deterministically or stochastically choices follow from learnt action values. They also determine how closely an agent follows a win-stay, lose-shift strategy. Also, higher feedback sensitivity can be interpreted as less decision noise and vice versa. The outcome-based

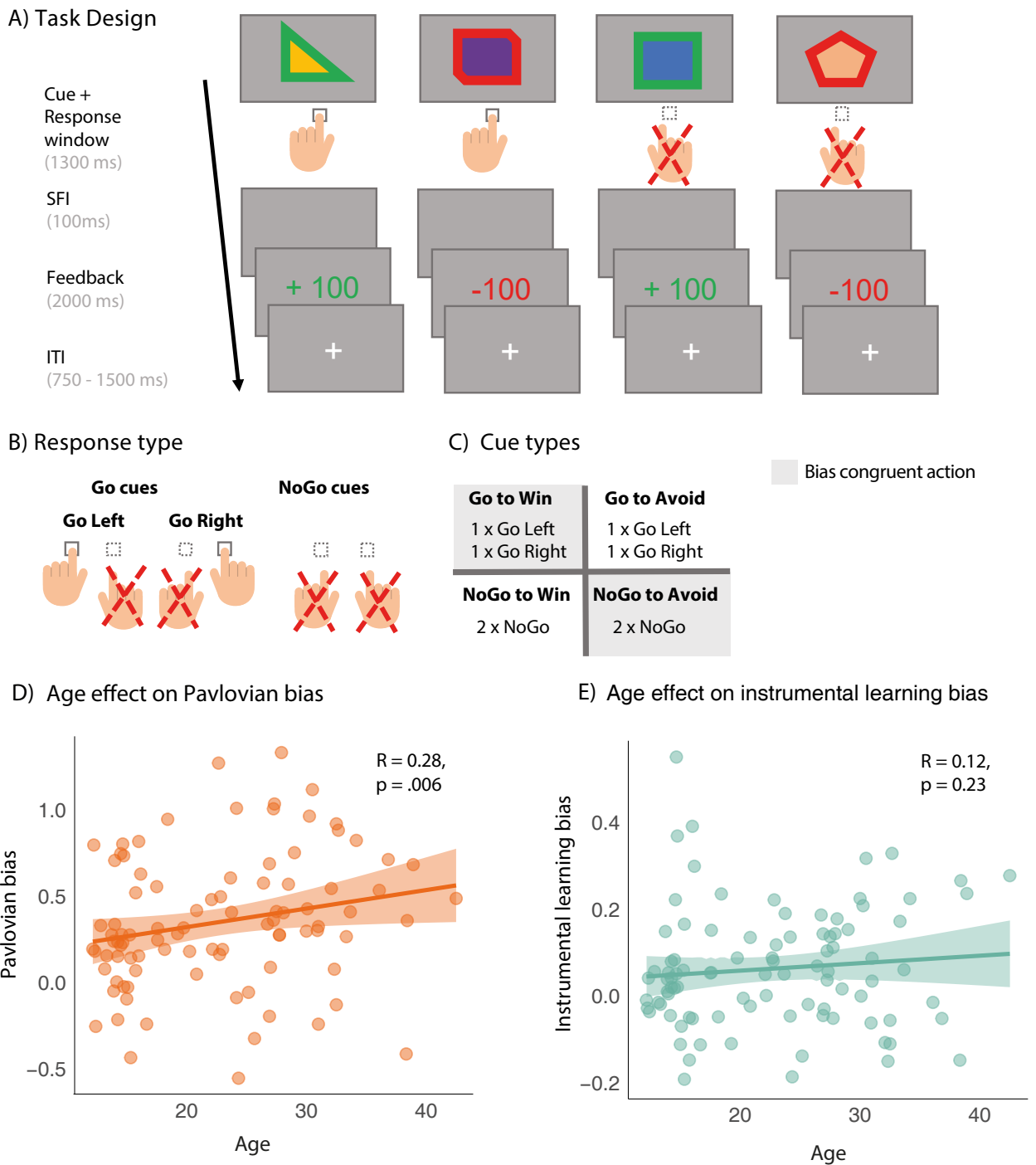


Fig 1. Task set-up motivational Go/NoGo task. (A) **Task trial sequence** displayed for all 4 cue categories, Go to Win, Go to Avoid, NoGo to Win, and NoGo to Avoid, but not split up for different Go responses (Go left and Go right). Go left/right responses for Win cues and NoGo responses for Avoid cues are considered bias-congruent, as the cue's respective action requirement matches with the stimulus-response coupling facilitated by the motivational bias. Meanwhile, Go responses for Avoid cues and NoGo for Win cues are considered bias-incongruent response-stimulus couplings. Each cue was presented for 1,300 milliseconds (ms) and participants had to decide whether to execute a Go Left or Go Right response by pressing the respective button or selecting a NoGo response by withholding responding. Subsequently, participants were shown a valid or invalid outcome (reward, neutral, punishment) for 2,000 ms based on the probabilistic feedback schedule (80:20% ratio) and cue valence. The inter-trial-interval (ITI) was 750–1,500 ms, in steps of 250 ms. (B) **Response type.** For the cues with a go response requirement, cues either required a left or a right button press. For the NoGo cues there was naturally no distinction. (C) **Cue types.** Depiction of which cues can be considered bias-congruent (gray shaded box), while the other cues (white box) are cues with bias-incongruent response requirements. (D) **Pavlovian Bias × Age.** Depiction of the

correlation between the individual slope for the valence term extracted from the mixed effects model capturing Pavlovian biases and age. The association shows a clear developmental change of Pavlovian biases across age. **(E) Instrumental Learning Bias \times Age.** Depiction of the correlation (spearman) between the individual slope for the interaction term (taken action \times outcome salience \times outcome valence) extracted from the second mixed effects model capturing instrumental learning biases and age. Here, no evidence for age-dependent changes of instrumental learning biases with age is discernible. Data for figure panels D and E can be found at <https://osf.io/mcx36/> together with code for reproducing those parts of the figure.

<https://doi.org/10.1371/journal.pbio.3002877.g001>

implementation of feedback sensitivity was indeed superior to the other tested models according to Bayesian model selection (protected exceedance probability (PXP) = 1; model expressed by 95.63%, see Table 1 for model comparison statistics and parameter estimates). Importantly, this model could also predict key characteristic behavioral patterns of the observed empirical data and parameter recovery was excellent (S6 Fig). In line with the mixed models on raw choice data, there was a positive correlation between age and Pavlovian biases ($r_s(91) = 0.23$, $p = 0.024$) and between age and feedback sensitivity to positive outcomes ($r_s(91) = 0.24$, $p = 0.018$), thus replicating previous reports of elevated noise in adolescents [4,16]. Other parameters did not correlate with age ($p > 0.2$) (Figs 2 and S4).

Noise generalizes across RL tasks

Next, we examined the cross-task generalizability versus context-specificity of noise parameters across several RL tasks: the Go/NoGo task, a probabilistic reversal learning task [16], and the two-step task [8]. For the reversal task, the winning model, as described in [16], comprised 4 feedback sensitivity parameters accounting for effects of motivational context (win reward versus avoid punishment) and feedback valence (positive /+FB versus negative/-FB). For simplicity, we averaged parameter estimates across motivational context, resulting in 2 estimates based on feedback valence. While $\rho_{\text{Go/NoGo} +\text{FB}}$ from the Go/NoGo task was positively associated with the reversal feedback sensitivity parameter for positive outcomes $\rho_{\text{Reversal} +\text{FB}}$ [$r_s(87) = 0.47$, $p < 0.001$], the reversal feedback sensitivity parameter

Table 1. Overview of transformed single hierarchical model parameter estimates (mean/standard error) and full (hierarchical) model comparison metrics and model frequencies. Protected exceedance probability and model frequency from the full model comparison identified model M7 as winning model. PXP = protected exceedance probability. The data and statistics presented in this table can be computed using code provided at <https://osf.io/mcx36/>.

	M1	M2	M3	M4	M5	M6	M7
ρ	25.73 [1.59]	24.17 [1.52]	15.37 [1.07]	21.10 [1.44]	15.71 [1.01]		
ρ_{win}						21.06 [1.22]	
ρ_{avoid}						12.60 [0.75]	
$\rho_{+\text{FB}}$							5.52 [0.24]
$\rho_{-\text{FB}}$							1.03 [0.10]
ϵ	0.04 [0.01]	0.05 [0.01]	0.08 [0.01]	0.06 [0.01]	0.06 [0.01]	0.06 [0.01]	0.18 [0.01]
b		-0.02 [0.06]	0.07 [0.06]	-0.001 [0.06]	0.06 [0.06]	0.02 [0.06]	-0.08 [0.03]
π			0.34 [0.09]		0.06 [0.07]	0.72 [0.08]	0.25 [0.03]
κ				2.61 [0.33]	1.68 [0.11]		
Full model comparison (M1 –M7)							
Model frequency (%)	0	0	4.37	0	0	0	96.63
PXP	0	0	0	0	0	0	1

<https://doi.org/10.1371/journal.pbio.3002877.t001>

Age effects on computational model parameters (winning model M7)

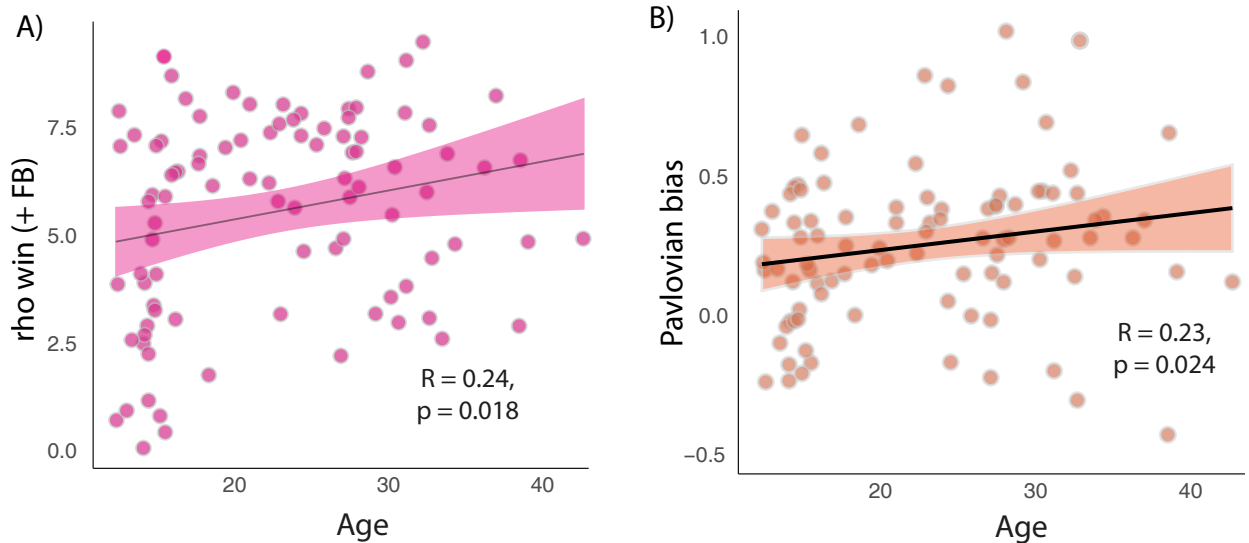


Fig 2. Significant age effects on model parameters extracted from the winning model M7. (A) The feedback sensitivity parameter for positive outcomes showed a significant age-dependency, such that older individuals were more sensitive to positive outcomes, as depicted in the scatterplots. (B) The Pavlovian bias parameter showed a significant age-dependency, such that older individuals' decisions were impacted more by motivational biases as captured by the Pavlovian bias parameter. Data underlying this figure and the code for reproducing it can be found at <https://osf.io/mcx36/>.

<https://doi.org/10.1371/journal.pbio.3002877.g002>

for negative outcomes correlated negatively with it [$\rho_{\text{Reversal -FB}}$: $r_s(87) = -0.3, p = 0.004$] (Fig 3A). Both survived multiple comparison correction with p -value < 0.0125 . No correlation between reversal task noise and feedback sensitivity for negative outcomes ($\rho_{\text{Go/NoGo -FB}}$) reached significance ($p > 0.1$).

For the two-step task, both noise parameters, β_1 and β_2 , were significantly correlated with Go/NoGo decision noise for positive outcomes (Fig 3B). More decision noise in the Go/NoGo task was associated with more decision noise from the two-step task (β_1 : $r_s(90) = 0.43, p$ -value < 0.001 ; β_2 : $r_s(90) = 0.52, p$ -value < 0.001). The association between second stage

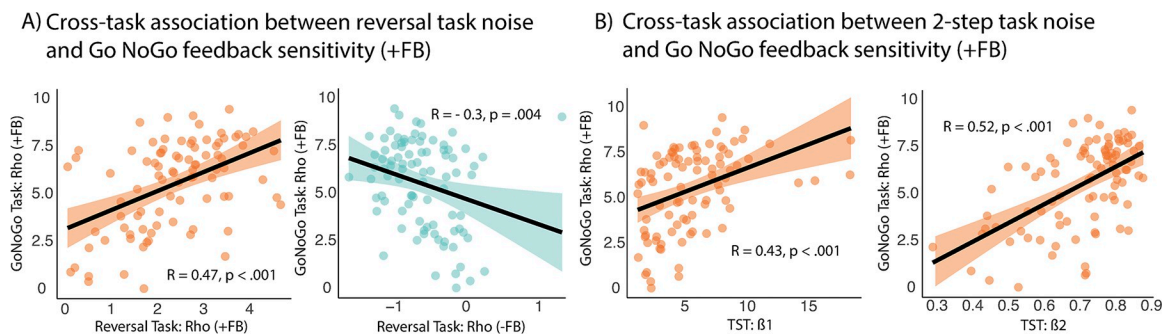


Fig 3. (A) Depiction of the association between feedback sensitivity for positive outcomes from the Go/NoGo Task and noise parameters from the reversal task. As expected, all noise parameters which captured noise related to positive outcomes were positively correlated, while negative correlations could be observed when correlating feedback sensitivity for positive outcomes with feedback sensitivity parameters from the reversal task capturing feedback sensitivity for negative outcome. (B) Significant cross-task correlation between both computationally derived noise parameters β_1 and β_2 from the two-step task with the feedback sensitivity parameter for positive outcomes from the Go/NoGo Task. Results were considered significant with $p < 0.05/4$ (0.0125), respectively, thereby correcting for multiple testing. The data underlying this figure and the code for reproducing it can be found at <https://osf.io/mcx36/>.

<https://doi.org/10.1371/journal.pbio.3002877.g003>

decision noise, β_2 , and Go/NoGo decision noise for negative outcomes was weaker and barely survived multiple comparison correction ($r_s(90) = -0.26$, p -value = 0.012). Cross-task correlation for learning rates from each task proved nonsignificant ($\epsilon_{\text{Go/NoGo}} \times \epsilon_{\text{Reversal}}$: $r_s(87) = -0.027$, $p = 0.8$; $\epsilon_{\text{Go/NoGo}} \times \epsilon_{\text{TST alpha1}}$: $r_s(90) = -0.008$, $p = 0.94$; $\epsilon_{\text{Go/NoGo}} \times \epsilon_{\text{TST alpha2}}$: $r_s(90) = -0.04$, $p = 0.74$), highlighting the task- or context-specificity of those parameters.

Results of additional correlational analyses of unspecific noise from the motivational Go/NoGo task with general and more specific task performance indices across all 3 tasks are reported in the SI (S1 Text and S8 Fig). In short, less decision noise for positive outcomes in the Go/NoGo task was associated with increased overall accuracy on the Go/NoGo and reversal task (pre-post reversal accuracy) as well as more goal-directed behavior and less switching after negative outcomes in the reversal task. Meanwhile, less decision noise for negative outcomes in the Go/NoGo task was associated with decreased task performance in the Go/NoGo task and the reversal task.

Noise mediates developmental changes in decision processes

A key interest of our study was to examine whether decision noise mediates specific developmental changes in decision processes, critically, across tasks. We assessed mediation effects of noise from the motivational Go/NoGo task (feedback sensitivity for positive feedback) on the association between age and (1) general task performance and non-selective Pavlovian Bias on the Go/NoGo task; (2) performance and loose shift behavior on the reversal task; and (3) model-based behavior on the 2-step task. Given that feedback sensitivity for negative feedback showed no association with age, this was done specifically for feedback sensitivity for positive feedback.

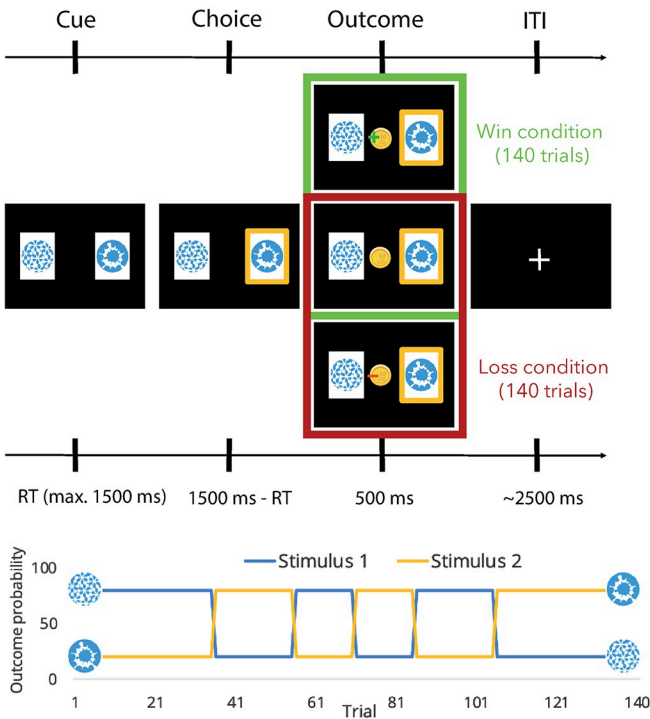
Within-task mediation: Motivational Go/NoGo task

We set up a mediation model with the age-dependent noise parameter (sensitivity to positive feedback) from the computational model of the motivational Go/NoGo task as mediator of the relationship between age and general performance in the task and Pavlovian bias. We found a significant partial mediation effect for ρ_{+FB} ($p = 0.02$), accounting for 83.6% of the total effect of the relationship between age and overall performance (Indirect effect = 0.004, CI [0.0005–0.01], $p = 0.03$; direct effect = 0.0009, CI [–0.001–0.003], $p = 0.5$; Total effect = 0.005, CI [0.0008–0.01], $p = 0.02$). Next, we examined whether the association between age and the score computed for overall Pavlovian bias [$P_{\text{corr}}(\text{Go2Win}) - P_{\text{corr}}(\text{Go2Avoid}) + P_{\text{corr}}(\text{NoGo2Win}) - P_{\text{corr}}(\text{NoGo2Avoid})$] or the computationally derived parameter capturing the Pavlovian bias was mediated by positive feedback sensitivity. This association between age and Pavlovian bias was not significantly mediated by positive feedback sensitivity, neither for the task score (p -value = 0.3) nor for the computational parameter capturing the effect of Pavlovian biases (p -value = 0.5).

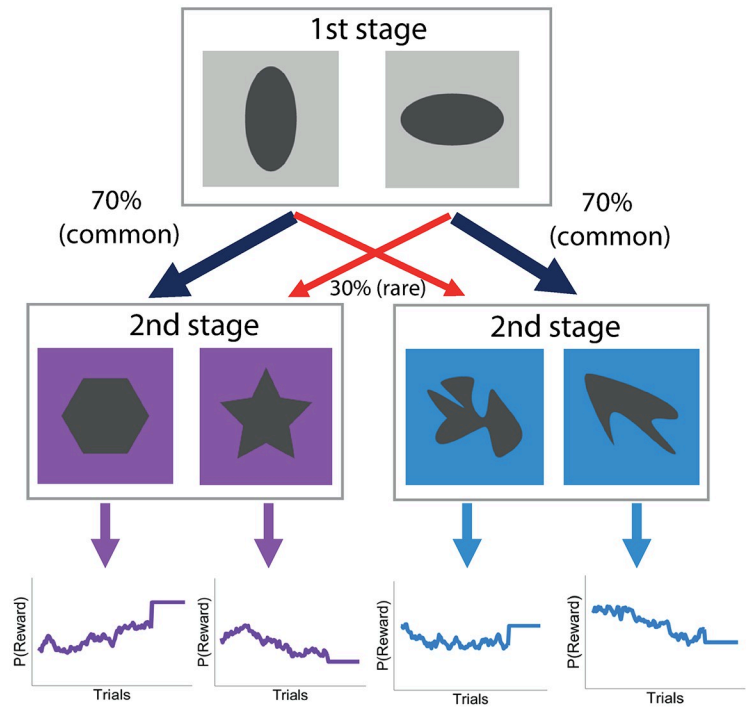
Cross-task mediation: Reversal task

Two mediation models were computed to determine the mediating effect of feedback sensitivity for positive outcomes (noise) in the Go/NoGo task on performance parameters derived from the reversal learning task, namely pre- minus post-reversal accuracy and switching following negative outcomes, both of which have previously been shown to correlate with age [16]. Assessment of the mediation effect of noise on the association between age and pre- minus post-reversal accuracy provided evidence for a partial mediation with ρ_{+FB} accounting for up to 37.1% of the total effect ($p = 0.03$) (Indirect effect = 0.0016, CI [0.0002–0.004], $p = 0.02$; direct effect = 0.003, CI [–0.0004–0.01], $p = 0.08$; Total effect = 0.005, CI [0.001–

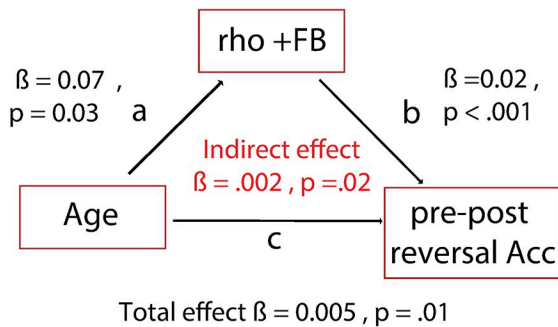
A) Reversal Task Design



B) Modified sequential (2-step) learning task



D) Mediation effects of noise on reversal



D) Mediation effects of noise on MB control

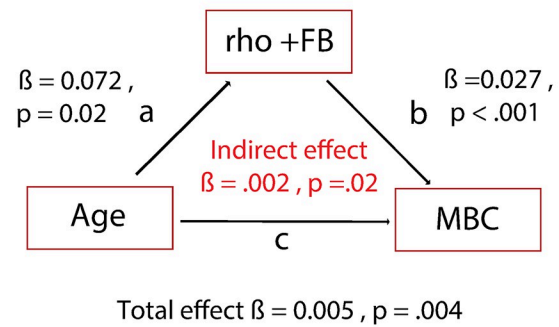


Fig 4. (A) Probabilistic reversal task design. In this task, participants had to adapt their behavior according to the changing outcome probabilities across the reversal learning task. Image of task design modified from Waltmann and colleagues [16]. **(B) Task design of sequential decision-making task.** In this 2-step task, a choice on the first stage led to one of 2 possible second stages. Here, participants had to make a second choice, upon which participants received a reward or neutral outcome (rewards were replaced by punishments in the punishment context). The probability of receiving a reward/punishment was determined by constantly changing probabilities, i.e., based on Gaussian random walks, while transition probabilities to transfer from stage 1 to stage 2 were fixed. They were either considered common transitions (70%) or rare (30%). Image modified from Scholz and colleagues [8]. **(C) Mediation analysis reversal task** indicating a significant mediation effect of noise from positive outcomes on the association between age and pre minus post reversal accuracy. **(D) Mediation analysis 2-step task** showing the significant mediation effect of feedback sensitivity for positive outcomes on the relationship between age and model-based control. *P*-values below 0.025 were considered significant thereby correcting for multiple testing.

<https://doi.org/10.1371/journal.pbio.3002877.g004>

0.01], $p = 0.01$) (Fig 4C). In the second mediation model, ρ_{+FB} accounted for up to 26,03% of the total effect between age and switching after negative feedback ($p = 0.02$) (Indirect effect = -0.002 , CI $[-0.003 \text{ -- } -0.0002]$, $p = 0.02$; direct effect = 0.005 , CI $[-0.007 \text{ -- } -0.001]$, $p = 0.005$; Total effect = -0.006 , CI $[-0.009 \text{ -- } -0.003]$, $p < 0.001$) (Fig 4A).

Cross-task mediation: 2-step task

Assessment of the mediation effect of feedback sensitivity for positive outcomes in the Go/NoGo task on the association between age and model-based control provided evidence of a partial mediation with ρ_{+FB} accounting for up to 44.01% of the total effect of the relationship of age and model-based control ($p = 0.02$) (Indirect effect = 0.002, CI [0.0004–0.004], $p = 0.02$; direct effect = 0.003, CI [–0.0003–0.01], $p = 0.08$; Total effect = 0.005, CI [0.002–0.01], $p = 0.004$) (Fig 4D). Rerunning the same mediation model using the computationally derived parameter omega (weight parameter capturing the balance between model-free and model-based control) indicated a similar, though somewhat weaker mediation effect (19.75%, p -value = 0.017) (see Fig 4B).

Discussion

In this study, we relied on computational modeling across 3 distinct RL tasks to assess a novel mediating role of decision noise—known to be elevated in adolescents—on age-dependent increases in more sophisticated choice behaviors and performance gain. This mediating role referred to model-based control, switching after negative outcomes on the reversal task and performance gain in both the Go/NoGo and reversal task. A choice heuristic like the Pavlovian bias was not mediated by decision noise. In line with this mediating role of decision noise, we also confirm previous findings indicating decision noise as a rather stable characteristic across tasks and contexts.

Using computational modeling, we assessed decision noise as latent feature of the decision process during development. We show an age-dependent increase of feedback sensitivity specifically for positive outcomes, as did [16] in a different task from the same sample. This is in line with well-known decreases of noise with higher age [27,41] and previous reports of more random “noisy” choice behavior in adolescents [4,16,27]. Consistent with previous work [4,26,42], we show stable cross-task generalization of decision noise by means of strong correlations of noise parameters across tasks. This reinforces the notion of decision noise as a task-independent feature with substantial interindividual differences. In contrast, cross-task correlations of learning rates indicated a lack of generalizability, much in accord with previous studies showing task-specificity of learning rates [4,26]. One tentative explanation for this is that learning per se, unlike decision noise, may be highly context dependent [26]. This might explain why specific developmental effects appear more inconsistent across (RL) tasks: while Rosenbaum and colleagues (2022) reported specific developmental effects on punishment learning rates in the absence of effects on reward learning rates, Pauli and colleagues (2023) described elevated reward learning rates [43]. Thus, even subtle difference in a task’s design may considerably impact the way adolescents learn and thus result in the detection of distinct effects, while individual differences in noise seem more robust against such differences in task design [4].

Critically, the central novel finding of our study indicated a mediating effect of this seemingly “unspecific” but stable noise on age-dependent increases in specific and more sophisticated choice signature across different task settings. Here, mediation analyses revealed that a substantial part of the variances between age and MB control, switching behavior after negative outcomes as well as overall task performance on the reversal and Go/NoGo task were accounted for by age-related developmental differences in noise levels. Thus, “unspecific” noise mediates the development of highly specific functions or strategies.

One reason for these mediation effects could be a limited availability of cognitive resources in adolescents [44–47] due to the ongoing development of brain areas related to cognitive control [1,48,49]. Having fewer cognitive resources might make adolescents more prone to rely on

computationally cheaper decision strategies, rendering them more susceptible to emotional, motivational, and social influences [28,50,51]. Still, other work has shown adolescents to employ more complex strategies relative to adults when mentalizing and processing social emotions [52]. The exertion of cognitive control as value-based choice, i.e., the willingness to allocate and exert control in certain situations [53–58] might be another possible explanation. Here, cognitive control is a choice, governed by a cost-benefit tradeoff, where people choose to exert control whenever this will result in a large enough increase in expected reward. Individuals can thus learn to selectively exert control, when this returns additional reward, but refrain from it if costs outweigh the benefits [59,60]. More noisy choice behavior, i.e., rather random-like exploration, may thus constitute a somewhat “rational” choice by adolescents to not mobilize control to reduce effort expenditure while achieving higher levels of control, may seem too costly. Alternatively or in addition, Ma and colleagues (2022) suggested that elevated explorative (or noisier) behavior might serve the development of central social behavior in adolescence [61], such that choice uncertainty (comparable to decision noise) predicts contagion effects of peer’s choices, which may be beneficial for social integration [62].

Of note, decision noise can reflect distinct underlying processes, such as random or more sophisticated directed exploration [30]. For example, directed exploration leads individuals to occasionally stray from selecting the optimal choice to deliberately choose less known options to gather information [44,45] to maximize long-term outcomes. Meanwhile, random exploration refers to randomly choosing options, a pattern resulting in more frequent choices of the non-optimal option, rather than deliberately choosing the worse option [45]. According to Findling and colleagues (2019) [45], the majority of choices that seemingly do not optimize reward values, appear to originate from so called learning noise. The locus-coeruleus-norepinephrine system has been proposed as potential neural correlate underlying learning noise [44,45] and using psychopharmacological manipulation, has been implicated in computationally “cheaper,” value-free exploration strategies [44,45]. Meanwhile, more elaborate exploration strategies appear to rely on other systems, like the dopaminergic system [63,64]. Future studies may disentangle such distinct noise components and its neurochemical correlates more precisely.

Interestingly, in our data no age effect was evident for decision noise for negative outcomes. Speculatively, this could be linked to relatively stronger effects of development on the reward versus punishment domain [8,16,43]. Alternatively, the parameter for decision noise for negative outcomes might capture different aspects of noise that are less affected by developmental changes, though we cannot make any more specific claims about this, as our Go/NoGo task cannot dissect distinct subcomponents of noise such as decision from learning noise.

One avenue of future research is distinct patterns of decision noise in developmental psychopathology. Indeed, elevated decision noise was reported across a wide range of psychiatric disorders [65–69]. Developmentally, a particular interesting condition associated with noisy behavior is attention-deficit hyperactivity disorder (ADHD). Some reports show elevated noise and exploration in ADHD patients [69–71] and non-clinical samples reporting ADHD symptoms [72,73]. Further assessment of the computational phenotype of ADHD leading to the characteristic profile of undirected explorative behavior, rapid task switching, and inattention might be key [30,74]. Such assessment would be particularly exciting given the implication of dopaminergic and noradrenergic pathways not only in decision noise and the exploration/exploitation trade-off [29,63,64,75] but also in the psychostimulant treatment of ADHD symptoms [76–81]. It remains unknown whether decision noise could serve as predictor for clinical outcomes like individual differences in response to psychostimulant treatment.

With respect to Pavlovian biases, we replicate a reduction during adolescence [17,20] but alongside increasing task performance from adolescence into adulthood. The latter finding on

performance conflicts with previous work, where lower Pavlovian biases presented with improved task performance, especially when participants had to actively “overcome” inherent Pavlovian responding [17]. This discrepancy might be explained by different task versions such as different number of Go responses, outcome types, and cued valence. We extend previous work [17,20] by showing that instrumental learning biases in this type of task do not undergo significant developmental changes.

As for limitations, a sample including younger children would have been superior by making our sample more comparable to previous work [17]. As data was collected as part of a larger study, this was not feasible. Still, we were able to partially replicate findings by showing a linear trend based on the particular age range available to us. While our study had a cross-sectional between-subject design, a longitudinal, within-subject design with multiple measurement points could have better captured developmental changes in the examined RL processes.

In sum, using computational modeling and mediation analysis, we showed that decision noise had a significant mediating effect on age-dependent increases in higher-level cognitive processes such as model-based control, switching after negative outcomes in the reversal task and overall performance in the motivational Go/NoGo and the reversal task. Cross-task analyses also emphasized decision noise as representing an interindividually more stable parameter, maybe even a trait-like feature. Future work may unravel the neural basis as well as the developmental and clinical real-life relevance of decision noise for neurodevelopmental disorders such as ADHD to bridge the gap between observed symptom-level behavior and neurobiological mechanisms. Moreover, given that many of the cognitive processes we measure appear to be at least in part impacted by noise, future studies should attempt to quantify the degree of noise relative to the central cognitive processes under investigation.

Material and methods

Sample

We recruited 103 participants as part of a larger developmental study, all of whom were screened for current psychiatric diagnosis. Participants completed several RL tasks, such as a reversal learning task capturing behavioral flexibility [16], a 2-step task measuring model-based control [8] and a well-established motivational Go/NoGo task assessing motivational biases in decision-making [23,31] (for a more detailed study description refer to the preregistered study protocol at <https://doi.org/10.17605/OSF.IO/FYN6Q> [36]). Of those 103, 99 participants completed the Go/NoGo task, of which 93 were subsequently analyzed ($n = 93$: age mean [SD] = 22.65 [7.88], age range 12 to 42 years, female: 45; male: 48), as $n = 5$ participants did not meet a rudimentary performance check (see SI) and one was an age outlier; 40% ($n = 37$) of our final sample were adolescents (18 years of age or below). The age distribution was not uniform but right-skewed (S1 Fig), as participants had been initially recruited as matched controls for a clinical sample in terms of age and gender. Participant reimbursement was 9 Euro per hour for study participation. Study proceedings were in agreement with the declaration of Helsinki and approved by the ethics board of the medical faculty at the University of Leipzig (385/17-ek). All participants were informed about the study proceedings and provided informed written consent before participating in the study.

Motivational Go/NoGo task

For this study, we used a well-established probabilistic reinforcement learning task known to experimentally measure motivational biases by examining the impact of valence (gain win versus avoid punishment contexts signaled by cues) on behavioral activation or inhibition (Go versus NoGo action) [23,31]. Here, on each trial, study participants were shown a cue for

which they had to decide whether to execute one of 2 Go responses (make a right or left button press) or abstain from it (NoGo, no button press) to either win a reward (Win cues) or avoid a punishment (Avoid cues) (Fig 1). Importantly, participants were aware whether they were playing for rewards or avoiding punishments, as cue valence was cued using a colored frame around each cue (green = Win cue; red = Avoid Cue). However, despite showing the correct response for a cue, participants could still receive invalid feedback 20% of the time (versus 80% valid feedback) according to probabilistic feedback. Participants could show a correct left Go response for a Win cue requiring a left Go response and still receive neutral feedback, the non-favorable outcome for a Win cue, on this trial. They heard a specific sound for receiving a reward, neutral feedback, or punishment. The task comprised 320 trials in total and each one of the 8 cues was presented 40 times. Before starting the main task, participants completed practice trials to guarantee that participants had understood task requirements such as the possibility of a NoGo response.

The impact of choice biases is operationalized by how well participants learn to show or omit a response when facing a reward or punishment cue requiring a Go or NoGo choice as optimal response. Consecutive choices represent learning of optimal choices and are guided by probabilistic feedback (rewards/neutral feedback for reward cues; punishment/neutral feedback for punishment cues). Importantly, bias-congruent responding (Go response for a Go-to-Win cue; NoGo for Avoid Punishment cue) should be facilitated, i.e., participants show better performance on these trials, while bias-incongruent performance should be impeded.

To tease apart differences in Pavlovian choice from instrumental learning biases, this task version had 2 Go responses, Go Left and Right [23,31,82]. This manipulation discerns whether, for example, a previously rewarded optimal Go response (e.g., Go Left) will be specifically reinforced and repeated or omitted more frequently in subsequent trials with the same cue [23,31]. This specificity for the optimal (Go) response reflects an instrumental learning bias, which will increase throughout the task while participants learn the optimal cue response, while the impact of Pavlovian choice bias recedes across the task [23]. Overall, instrumental learning biases are more selective when compared with Pavlovian biases.

Unlike the instrumental learning bias, the Pavlovian bias does not distinguish between the type of Go response, here Go Left or Go Right, in its facilitation effect, so any Go response followed by a reward will increase the likelihood that a Go response will be selected on the next trial this cue is presented. Also, Pavlovian biases usually do so as early as the first trial [23]. They are also characterized by a nonspecific tendency to show more Go responses for Win versus Avoid Punishment cues.

Mixed model analysis of Go/NoGo task

To examine Pavlovian choice biases, we assessed whether the probability of making a Go response, subsequently termed P(Go), was impacted by the within-subject factors required action (Go versus NoGo) and valence (Win versus Avoid Punishment) while also including age (z-standardized) as additional covariate of interest. We expected a linear age effect for Pavlovian biases based on previous reports, such that adolescents would display lower levels of Pavlovian biases relative to adults [17,20]. We were therefore particularly interested in the 2-way interaction valence \times age indicating whether Pavlovian biases change with age, alongside the main effects of (1) required action representing whether individuals actually learn to make the correct response; (2) the main effect of valence capturing the presence of a motivational bias on choice behavior.

$$P(\text{Go}) \sim \text{required action}^* \text{cue valence}^* \text{age} + (\text{required action}^* \text{cue valence} + 1|\text{Subject})$$

To examine developmental effects on instrumental learning biases, we set up a second model based on previous work [31]. Here, we tested whether the probability of repeating the previous response $P(\text{repeat})$, changed depending on 3 within-subject factors, namely the response selected on the previous trial (Go versus NoGo response), the outcome valence (positive: reward for reward cues; neutral feedback for avoidance cues versus negative: punishment for avoidance cues; neutral feedback for reward cues) and outcome salience (salient: reward/punishment feedback, non-salient = neutral feedback). Importantly, in this model, the presence of an instrumental learning bias is indicated by a significant 3-way interaction: action taken \times outcome valence \times outcome salience. As we were specifically interested in the presence of age-dependent effects, the model also included a linear age term (z-standardized) as covariate of interest (also see S1 Text for an alternative model implementation). Prior behavioral and neural work [83–85] has reported age-dependent differences in feedback-based, instrumental learning processes, for instance heightened negative (relative to positive) learning rates during adolescence [85]. Consequently, we expected to observe reduced selective response facilitation of the correct response for a reward cue during adolescence.

$$P(\text{repeat}) \sim \text{action taken}_{t-1} * \text{outcome valence}_{t-1} * \text{outcome salience}_{t-1} * \text{age} + (\text{action taken}_{t-1} * \text{outcome valence}_{t-1} * \text{outcome salience}_{t-1} + 1 | \text{Subject})$$

Given previous evidence of differences in age of onset of puberty [86], we also ran models including gender as additional control variable, thereby assessing potential effects of gender on choice behavior and biases.

All generalized logistic mixed effects models were computed using the lme4 package, version 1.1–31 in R 4.2.2 with the optimizer bobyqa and the maximal number of iterations set to $n = 1e+9$. Statistical significance was determined using p -values with $\alpha < 0.05$, two-sided.

Computational modeling of Go/NoGo task and age-dependent changes

To dissect the computational mechanisms sub-serving Pavlovian action biases and instrumental learning, we fitted several hierarchically nested RL (M1–M7). For this, we relied on the cbm toolbox implemented in matlab [87], which is based on hierarchical Bayesian inference (HBI) and treats the model itself as a random effect [88,89]. Models M1–M5 have been previously employed and outlined in much detail by Swart and colleagues [23]. In brief, model M1 represented a Rescorla Wagner model [90] comprising a learning rate (ϵ) and a second parameter capturing feedback sensitivity, to learn value of each respective action (a : Go left, Go right, NoGo) for each stimulus (s) on each trial t :

$$Q_t(a_t, s_t) = Q_{t-1}(a_t, s_t) + \epsilon(\rho r_t - Q_{t-1}(a_t, s_t)) \quad (1)$$

M2 is an extension of M1 with an additional “Gobias” parameter b capturing an overall tendency to give a Go response. Model M3 extended model M2 with another parameter operationalizing the Pavlovian tendency π to show more Go responses for Win relative to Avoid Punishment cues. Both bias parameters were integrated with the learnt Q values into the action weights w :

$$w_t(a_t, s_t) = \begin{cases} Q_t(a_t, s_t) + b + V\pi(s) & \text{if } a = \text{Go} \\ Q_t(a_t, s_t) & \text{else} \end{cases} \quad (2)$$

Model M4 included an instrumental learning bias parameter κ instead of the Pavlovian parameter π to assess whether the choice behavior could have been solely produced by a learning bias. For this, κ was included as a modification of the learning as follows:

$$\epsilon = \begin{cases} \epsilon_o + \kappa & \text{if } r_t = 1 \text{ and } a = \text{Go} \\ \epsilon_o - \kappa & \text{if } r_t = -1 \text{ and } a = \text{NoGo} \\ \epsilon_o & \text{else} \end{cases} \quad (3)$$

Importantly, to ensure a symmetric impact of κ , the following requirements were implemented depending on the size of the learning rate

$$\epsilon = \begin{cases} \epsilon_o = \text{inv.logit}(\epsilon) \\ \epsilon_{\text{punished NoGo}} = \text{inv.logit}(\epsilon - \kappa) & \text{if } \epsilon_o < .5 \\ \epsilon_{\text{rewarded Go}} = \epsilon_o + (\epsilon_o - \epsilon_{\text{punished NoGo}}) & \text{if } \epsilon_o < .5 \\ \epsilon_{\text{rewarded Go}} = \text{inv.logit}(\epsilon + \kappa) & \text{if } \epsilon_o > .5 \\ \epsilon_{\text{punished NoGo}} = \epsilon_o + (\epsilon_o - \epsilon_{\text{rewarded Go}}) & \text{if } \epsilon_o > .5 \end{cases} \quad (4)$$

For all models, V denoted the cued valence ($V_{\text{win}} = +0.5$; $V_{\text{avoid}} = -0.5$). Consequently, a positive π lead to an increased action weight for Go responses for Win cues, while resulting in a reduced action weight for Go responses on Avoid cues. Action weights were transformed to action probabilities using a softmax function:

$$p(a_t | s_t) = \frac{\exp(w(a_t, s_t))}{\sum_a \exp(w(a', s_t))} \quad (5)$$

Model M5 included both bias parameters κ and π . Due to previous reports of distinct learning and processing of positive and negative feedback in adolescents [8,16], we aimed to examine whether feedback sensitivity parameters for each motivational context (ρ_{win} and ρ_{avoid}) or feedback sensitivity parameters for positive versus negative feedback ($\rho_{\text{+FB}}$ or $\rho_{\text{-FB}}$; present in win and avoid motivational context) would provide a better account of the data (see [Table 1](#) for an overview of the model space). This was motivated by rather noisy behavior seen in adolescents in similar RL tasks [4,16,27]. Information of parameter transformation can be found in the SI.

We performed extensive model simulation based on the established models to rigorously compare the observed relative to the simulated data ([S5 Fig](#)). This revealed that the model only including an instrumental learning bias parameter κ captured the observed behavioral data very poorly. In the same vein, model M5 including both motivational bias parameters, π and κ , did not perform better compared to model M3, which only included the Pavlovian bias parameter π (see [S1 Text](#), [S5](#) and [S6 Figs](#) for details on simulations, model validation, and parameter recovery). Hence, for the purpose of parsimony, we focused on 2 additional model extensions, model M6 and M7, which implemented 2 feedback sensitivity parameters for cue valence (Win and Avoid Punishment) or for positive and negative outcome valence together with a single learning rate, a Gobias and a Pavlovian bias parameter (see [S1 Text](#) for full model space).

We concluded this analysis using random effects model comparison [88,89], which computes the Laplace approximation of model evidence based on the individual level [38,91], from which group model evidence is derived to establish which model best captured the behavioral data. Model evidence was examined by comparing the PXP [87]. PXP assesses the most frequently expressed model [88] while accounting for the possibility of chance results. We then

extracted hierarchical model parameter estimates from the winning model and examined age-dependent effects using spearman correlations.

As we were particularly interested in the developmental changes underlying decision noise and biases in choice behavior [16,17,26,92], we evaluated these associations for the respective computational parameters from the winning model with age using spearman correlation coefficients. Given our specific hypotheses regarding the developmental pattern of those three parameters, we considered these confirmatory analyses for which we did not apply multiple comparison correction. For completeness, we also report associations for age with the remaining 2 parameters, namely the learning rate and the Gobias in the supplement.

Cross task generalizability of unspecific decision noise

Given our primary interest in the cross-task generalizability of decision noise, we subsequently assessed whether feedback sensitivity parameters from the motivational Go/NoGo task would be associated with related parameters from 2 other RL tasks, detailed results for which have been published elsewhere [8,16]. Apart from the motivational Go/NoGo task, we had access to data from a probabilistic reversal learning task capturing cognitive flexibility and a sequential probabilistic decision-making task assessing model-based control (see [S1 Text](#) for task details).

We first examined the association between computationally modeled noise parameters across the Go/NoGo task (ρ_{+FB} and ρ_{-FB}) and the probabilistic reversal learning task using spearman correlations. For the reversal task, noise parameters were extracted from the winning computational model that had contained 4 noise parameters accounting for both outcome and cue valence, namely, $\rho_{Win +FB}$, $\rho_{Win -FB}$, $\rho_{Loss +FB}$ and $\rho_{Loss -FB}$ [16]. Those estimates were further simplified by averaging parameters on the dimension of cue valence, thereby creating 2 parameter estimates, namely, $\rho_{Reversal +FB}$, $\rho_{Reversal -FB}$. In total, this meant 4 correlations were computed across both tasks. Second, we assessed cross-task associations for noise parameters from the Go/NoGo task (ρ_{+FB} and ρ_{-FB}) and the 2-step task, for which noise parameters β_1 and β_2 were extracted from a well-established hybrid model (see [38] for an extensive model description).

Associations between unspecific noise with specific cognitive functions

Next, we also assessed the relationship between the 2 feedback sensitivity parameters (decision noise) extracted from the winning model for the Go/NoGo task and the index capturing MB control from the modified 2-step RL task (reported in [8]). Given that we did not find significant valence differences for MB control in our previous developmental work [8], we only assessed associations for noise parameters and overall MB control. Furthermore, we computed correlation between the Go/NoGo task noise parameters with the behavioral indices from the reversal learning task, namely pre-post reversal accuracy and switching after negative outcomes.

To determine the specificity of findings, we evaluated the cross-task association between learning rates from the Go/NoGo task, the 2-step and reversal learning task. Here, based on prior work suggesting considerable variation in learning rates based on situational context [4,26], we did not expect cross-task learning rates to show significant correlations.

Results were considered significant with a p -value < 0.0125 for associations between Go/NoGo and the reversal task parameters (0.05/4) and a p -value < 0.025 for the associations between the Go/NoGo task and the 2-step task parameters (0.05/2) to correct for multiple testing. This was done separately for each task, as both tasks were considered independently.

Decision noise as a mediator for higher-level cognitive processes?

To address our second aim for this paper, we also computed several mediation analyses. All included feedback sensitivity as mediator variable and assessed its significance for age related changes on specific cognitive functions or decision processes, namely MB control, decision biases, and markers of cognitive flexibility. The mediation package in R was used for all mediation analysis. Results are reported based on nonparametric bootstrap confidence intervals based on the percentile method (simulations $n = 10,000$).

The impact of decision noise on Go/NoGo task performance and Pavlovian biases

We first assessed whether the noise parameter for positive outcomes extracted from our winning computational model M7 (ρ_{+FB}) exerted a mediating effect on the relationship between age and overall task performance ($P(\text{correct})$), on a previously employed score computed for the overall impact of Pavlovian biases on choice behavior [$P_{\text{corr}}(\text{Go2Win}) - P_{\text{corr}}(\text{Go2Avoid}) + P_{\text{corr}}(\text{NoGo2Win}) - P_{\text{corr}}(\text{NoGo2Avoid})$] [23] as well as the computational parameter capturing Pavlovian biases.

The role of decision noise for the association between age and cognitive flexibility

Based on [16], the difference between pre and post reversal accuracy as well as the degree of switching behavior after negative feedback showed significant correlations with age. Hence, we examined whether these parameters were mediated by the noise parameters for positive outcomes from the motivational Go/NoGo task.

The impact of decision noise on the association between age and MB control

Finally, given previous reports linking age and MB control [8], we also assessed whether this association might be (partially) mediated by the extracted noise parameter for positive outcomes from our winning model M7, namely ρ_{+FB} . MB control was operationalized in 2 different ways, namely, the computationally derived parameter ω extracted from a well-established hybrid model [38] as well as effect estimates extracted from a mixed-effects model characterizing MB control (details reported in [8]).

Supporting information

S1 Text. Supplemental material. File providing additional information on data analysis and results including mixed and computational models as well as correlational analysis and the employed tasks.
(PDF)

S1 Table. Age-dependent changes in Pavlovian biases. Table displaying the β estimates, the standard error (SE) as well as statistics for the main and interaction effects from the mixed-effects model computed to assess the impact of age on Pavlovian biases as well as general learning of the task. Here, the dependent variable was the probability of making a go response P (Go). Data and code to compute the statistics presented in this table is available at <https://osf.io/mcx36/>.
(PDF)

S2 Table. Age-dependent changes in accurate task performance. Table displaying the β estimates, the standard error (SE) as well as statistics from the mixed-effects model computed to assess the impact of age on Pavlovian biases as well as general learning of the task. Here, the dependent variable was the probability of making a correct response $P(\text{Correct})$. Data and code to compute the statistics presented in this table is available at <https://osf.io/mcx36/>. (PDF)

S3 Table. Age-dependent changes in instrumental learning biases. Table displays the β estimates, standard errors (SE) as well as statistics from the mixed-effects model computed to assess the impact of age on instrumental learning biases. Here, the dependent variable was the probability of repeating the same response for a given cue $P(\text{repeat})$. Data and code to compute the statistics presented in this table is available at <https://osf.io/mcx36/>. (PDF)

S4 Table. Effects of gender on Pavlovian biases. Table providing an overview of the β estimates, standard errors (SE) as well as statistics from the mixed-effects model computed assessing the impact of gender on the measured task effects. Here, the dependent variable was the probability of making a Go response $P(\text{Go})$. Data and code to compute the statistics presented in this table is available at <https://osf.io/mcx36/>. (PDF)

S5 Table. Effects of gender on instrumental learning biases. Table providing an overview of the β estimates, standard errors (SE) as well as statistics from the mixed-effects model computed assessing the impact of gender on the instrumental learning bias. Here, the dependent variable was the probability of repeating the same response for a given cue $P(\text{repeat})$. Data and code to compute the statistics presented in this table is available at <https://osf.io/mcx36/>. (PDF)

S1 Fig. Distribution of Age and gender. (A) Age distribution (age between 12 and 43 years) with the solid line indicating the mean age of 22.65 for this sample ($n = 93$). (B) Gender distribution of sample. (Female = 45, Male = 48). (EPS)

S2 Fig. (A) Conceptualization of motivational choice biases. (B) **Probability of $P(\text{Go})$ as a function of required action and cue valence.** Learning is apparent from the increased frequency of Go responses for Go cues. The impact of motivational biases is evident from the decreased probability of Go responses for Avoid cues and the increased frequency of Go responses towards Win cues independent of the actually required action. (C) **Probability of $P(\text{Repeat})$ as a function of the action taken and outcome valence.** Here, outcome valence is additionally split up by salience, i.e., whether participants received an actual reward or punishment or neutral feedback. Probability of repeating an action depicted as a function of outcome received on the previous trial and its salience. Cue categories are abbreviated as follows: G2W = Go to Win, G2A = Go to Avoid Punishment, N2W = NoGo to Win, N2A = NoGo to Avoid Punishment. The data underlying the figure panels B and C as well as the code for plotting it can be found at <https://osf.io/mcx36/>. (EPS)

S3 Fig. Depiction of age effects on the interaction action \times valence. (A) Scatterplot depicting the association between age and model estimates for interaction term valence \times action extracted from the mixed-effect model with PGo as dependent variable. (B) Scatterplot showing the correlation between age and the approach bias computed based on the raw data PGo (G2W - G2A). (C) Scatterplot showing the correlation between age and the avoid component

of the Pavlovian bias computed based on the raw data PGo (NG2W - NG2A). Data and code necessary to replicate this figure can be found at <https://osf.io/mcx36/>.
(EPS)

S4 Fig. Age effects on computational model parameters. Scatterplots depicting significant age effects on feedback sensitivity for positive outcomes and Pavlovian biases, while showing age independent trajectories for the remaining computational parameters, namely, feedback sensitivity for negative outcomes, the learning rate parameter and the go bias, for which (spearman) correlation coefficients were all non-significant (p -value > 0.5). The code with which the data underlying this figure was produced and for plotting this figure can be found at <https://osf.io/mcx36/>.
(EPS)

S5 Fig. Posterior predictive model simulations. (A–C) Panels A–C depict the aggregated results of model simulations computed for models M3–M5 (colored lines) compared to the actually observed data (gray colored lines) to determine whether those models can capture the basic patterns seen for the actual observed behavioral data. Here, simulations ($n = 100$) computed new choices and outcomes according to response probabilities that were based on the optimal parameter estimates generated for the respective model and which were subsequently averaged across all subjects. Here, key elements of the behavioral pattern are for instance whether participants learnt the task (i.e., showed the correct Go and NoGo responses for the respective Go vs. NoGo cues) and whether a strong valence effect characteristic for the influence of Pavlovian biases, such as an increased frequency of Go responses for Win cues relative to Avoid Punishment cues, could be detected. (Panel A–C) Trial by trial estimates of the probability of showing a Go response for models M3–M5. Here, M3 only contains a Pavlovian bias parameter alongside one overall feedback sensitivity parameter, a learning rate and a gobias parameter, M4 exchanged the Pavlovian bias parameter (π) with a learning bias parameter, while model M5 contained both bias parameters. It becomes evident that only model M3 containing a Pavlovian bias parameter only is able to capture the observed behavioral patterns, while both M4 and M5 show major divergence from the behavioral data indicating strong effects of over- and underestimation of the observed data. (Panel D–F) Depiction of the difference score between the probability of repeating the same choice shown on the previous trial for the same cue on the next trial [P (Repeat)] with the predicted choices based on the simulation runs being subtracted from the originally observed choices and subsequently averaged across participants. Here, choice patterns are split up by valence (Win and Avoid Punishment) and Salience (reward vs. punishment, no reward vs. no punishment). Again, model M3 including only 1 parameter capturing the impact of Pavlovian biases appears to outperform the other 2 models M4 and M5 given the smaller rate of over- and underestimation of actual performance indicated by the smaller difference bars overall. The data underlying this figure and the plotting code can be found at <https://osf.io/mcx36/>.
(EPS)

S6 Fig. Overview of parameter recovery. (A) Scatterplots depicting the correlation (spearman coefficient) between the parameter estimates of the winning model M7 comprising 2 feedback sensitivity parameters for positive and negative outcomes, a learning rate as well as a Go and a Pavlovian bias computed for observed data as well as simulated data (parameter estimates averaged across $n = 100$ simulations). (B) Density distributions of parameter estimates of model M7 for observed data (green) and recovered data (blue). Dashed line indicates the mean of the distribution. (C) Scatterplot depicting the correlation (spearman) between the observed and averaged, simulated probability of making a go response, P(Go), and the probability of making a

correct response (left Go response for a Go cue requiring a left button press, a right Go response for cues with a right button press as optimal response, and NoGo response for NoGo cues). The data underlying this figure and the plotting code can be found at <https://osf.io/mcx36/>.

(EPS)

S7 Fig. Mediation effect of positive feedback noise on task performance. Mediation analysis indicated a significant partial mediation effect of feedback sensitivity for positive outcomes (ρ +FB) on P(correct), i.e., the overall response accuracy when performing the task. Data and code to complete the mediation analysis are available at <https://osf.io/mcx36/>.

(EPS)

S8 Fig. Cross-task association for noise parameters from the Go NoGo and Reversal learning task. The scatterplots depict the association between feedback sensitivity for negative outcomes from the motivational Go NoGo task and the 2 noise parameters for positive and negative outcomes from the reversal task. Both correlation coefficients were nonsignificant (p -value > 0.1). The data underlying this figure and the plotting code can be found at <https://osf.io/mcx36/>.

(EPS)

Author Contributions

Conceptualization: Annette Horstmann, Lorenz Deserno.

Data curation: Maria Waltmann, Nadine Herzog.

Formal analysis: Vanessa Scholz, Lorenz Deserno.

Funding acquisition: Annette Horstmann, Lorenz Deserno.

Investigation: Maria Waltmann, Nadine Herzog.

Methodology: Vanessa Scholz, Lorenz Deserno.

Project administration: Maria Waltmann, Nadine Herzog.

Supervision: Lorenz Deserno.

Validation: Lorenz Deserno.

Visualization: Vanessa Scholz.

Writing – original draft: Vanessa Scholz, Lorenz Deserno.

Writing – review & editing: Vanessa Scholz, Maria Waltmann, Nadine Herzog, Annette Horstmann, Lorenz Deserno.

References

1. Blakemore SJ. Imaging brain development: The adolescent brain. *Neuroimage*. 2012; 61. <https://doi.org/10.1016/j.neuroimage.2011.11.080> PMID: 22178817
2. Bolenz F, Reiter AMF, Eppinger B. Developmental changes in learning: Computational mechanisms and social influences. *Front Psychol*. 2017; 8. <https://doi.org/10.3389/fpsyg.2017.02048> PMID: 29250006
3. Steinbeis N, Crone EA. The link between cognitive control and decision-making across child and adolescent development. *Curr Opin Behav Sci*. 2016; 10. <https://doi.org/10.1016/j.cobeha.2016.04.009>
4. Nussenbaum K, Hartley CA. Reinforcement learning across development: What insights can we draw from a decade of research? *Dev Cogn Neurosci*. 2019; 40. <https://doi.org/10.1016/j.dcn.2019.100733> PMID: 31770715

5. DePasque S, Galván A. Frontostriatal development and probabilistic reinforcement learning during adolescence. *Neurobiol Learn Mem.* 2017; 143. <https://doi.org/10.1016/j.nlm.2017.04.009> PMID: [28450078](https://pubmed.ncbi.nlm.nih.gov/28450078/)
6. Fuhrmann D, Knoll LJ, Blakemore SJ. Adolescence as a Sensitive Period of Brain Development. *Trends Cogn Sci.* 2015; 19. <https://doi.org/10.1016/j.tics.2015.07.008> PMID: [26419496](https://pubmed.ncbi.nlm.nih.gov/26419496/)
7. Sutton RS, Barto AG. Reinforcement learning: an introduction. 1998; 322. <https://doi.org/10.1109/TNN.1998.712192>
8. Scholz V, Waltmann M, Herzog N, Reiter A, Horstmann A, Deserno L. Cortical grey matter mediates increases in model-based control and learning from positive feedback from adolescence to adulthood. *J Neurosci.* 2023. <https://doi.org/10.1523/JNEUROSCI.1418-22.2023> PMID: [36823039](https://pubmed.ncbi.nlm.nih.gov/36823039/)
9. Smid CR, Kool W, Hauser TU, Steinbeis N. Computational and behavioral markers of model-based decision making in childhood. *Dev Sci.* 2022;1–36. <https://doi.org/10.1111/desc.13295> PMID: [35689563](https://pubmed.ncbi.nlm.nih.gov/35689563/)
10. Vaghi MM, Moutoussis M, Váša F, Kievit RA, Hauser TU, Vértes PE, et al. Compulsivity is linked to reduced adolescent development of goal-directed control and frontostriatal functional connectivity. *Proc Natl Acad Sci U S A.* 2020; 117. <https://doi.org/10.1073/pnas.1922273117> PMID: [32989168](https://pubmed.ncbi.nlm.nih.gov/32989168/)
11. Decker JH, Otto AR, Daw ND, Hartley CA. From Creatures of Habit to Goal-Directed Learners. *Psychol Sci.* 2016; 27. <https://doi.org/10.1177/0956797616639301> PMID: [27084852](https://pubmed.ncbi.nlm.nih.gov/27084852/)
12. Nussenbaum K, Scheuplein M, Phaneuf CV, Evans MD, Hartley CA. Moving developmental research online: Comparing in-lab and web-based studies of model-based reinforcement learning. *Collabra Psychol.* 2020; 6. <https://doi.org/10.1525/collabra.17213>
13. Potter TCS, Bryce NV, Hartley CA. Cognitive components underpinning the development of model-based learning. *Dev Cogn Neurosci.* 2017; 25. <https://doi.org/10.1016/j.dcn.2016.10.005> PMID: [27825732](https://pubmed.ncbi.nlm.nih.gov/27825732/)
14. Palminteri S, Kilford EJ, Coricelli G, Blakemore SJ. The Computational Development of Reinforcement Learning during Adolescence. *PLoS Comput Biol.* 2016; 12. <https://doi.org/10.1371/journal.pcbi.1004953> PMID: [27322574](https://pubmed.ncbi.nlm.nih.gov/27322574/)
15. Decker JH, Lourenco FS, Doll BB, Hartley CA. Experiential reward learning outweighs instruction prior to adulthood. *Cogn Affect Behav Neurosci.* 2015; 15. <https://doi.org/10.3758/s13415-014-0332-5> PMID: [25582607](https://pubmed.ncbi.nlm.nih.gov/25582607/)
16. Waltmann M, Herzog N, Reiter AMF, Villringer A, Horstmann A, Deserno L. Diminished reinforcement sensitivity in adolescence is associated with enhanced response switching and reduced coding of choice probability in the medial frontal pole. *Dev Cogn Neurosci.* 2023; 60. <https://doi.org/10.1016/j.dcn.2023.101226> PMID: [36905874](https://pubmed.ncbi.nlm.nih.gov/36905874/)
17. Raab HA, Hartley CA. Adolescents exhibit reduced Pavlovian biases on instrumental learning. *Sci Rep.* 2020; 10. <https://doi.org/10.1038/s41598-020-72628-w> PMID: [32978451](https://pubmed.ncbi.nlm.nih.gov/32978451/)
18. Raab HA, Hartley CA. The development of goal-directed decision-making. *Goal-Directed Decision Making: Computations and Neural Circuits.* 2018. <https://doi.org/10.1016/B978-0-12-812098-9.00013-9>
19. Telzer EH, Dai J, Capella JJ, Sobrino M, Garrett SL. Challenging stereotypes of teens: Reframing adolescence as window of opportunity. *Am Psychol.* 2022; 77:1067–1081. <https://doi.org/10.1037/amp0001109> PMID: [36595405](https://pubmed.ncbi.nlm.nih.gov/36595405/)
20. Betts MJ, Richter A, de Boer L, Tegelbeckers J, Perosa V, Baumann V, et al. Learning in anticipation of reward and punishment: perspectives across the human lifespan. *Neurobiol Aging.* 2020; 96. <https://doi.org/10.1016/j.neurobiolaging.2020.08.011> PMID: [32937209](https://pubmed.ncbi.nlm.nih.gov/32937209/)
21. Dayan P, Niv Y, Seymour B, Daw ND. The misbehavior of value and the discipline of the will. *Neural Netw.* 2006; 19. <https://doi.org/10.1016/j.neunet.2006.03.002> PMID: [16938432](https://pubmed.ncbi.nlm.nih.gov/16938432/)
22. Guitart-Masip M, Duzel E, Dolan R, Dayan P. Action versus valence in decision making. *Trends Cogn Sci.* 2014; 18:194–202. <https://doi.org/10.1016/j.tics.2014.01.003> U6 - <http://www.cell.com/article/S1364661314000205/fulltext> M4—Citavi PMID: [24581556](https://pubmed.ncbi.nlm.nih.gov/24581556/)
23. Swart JC, Froböse MI, Cook JL, Geurts DEM, Frank MJ, Cools R, et al. Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. *Elife.* 2017; 6:1–36. <https://doi.org/10.7554/eLife.22169> PMID: [28504638](https://pubmed.ncbi.nlm.nih.gov/28504638/)
24. Boureau YL, Dayan P. Opponency revisited: Competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology.* 2011; 36:74–97. <https://doi.org/10.1038/npp.2010.151> PMID: [20881948](https://pubmed.ncbi.nlm.nih.gov/20881948/)
25. Moutoussis M, Bullmore ET, Goodyer IM, Fonagy P, Jones PB, Dolan RJ, et al. Change, stability, and instability in the Pavlovian guidance of behaviour from adolescence to young adulthood. *PLoS Comput Biol.* 2018; 14. <https://doi.org/10.1371/journal.pcbi.1006679> PMID: [30596638](https://pubmed.ncbi.nlm.nih.gov/30596638/)

26. Eckstein MK, Master SL, Dahl RE, Wilbrecht L, Collins AGE. Reinforcement learning and Bayesian inference provide complementary models for the unique advantage of adolescents in stochastic reversal. *Dev Cogn Neurosci*. 2022; 55. <https://doi.org/10.1016/j.dcn.2022.101106> PMID: 35537273
27. Chierchia G, Soukupová M, Kilford EJ, Griffin C, Leung J, Palminteri S, et al. Confirmatory reinforcement learning changes with age during adolescence. *Dev Sci*. 2021. <https://doi.org/10.1111/desc.13330> PMID: 36194156
28. Reiter A, Heinz A, Deserno L. Linking social context and addiction neuroscience: a computational psychiatry approach. *Nat Publ Group*. 2017; 18. <https://doi.org/10.1038/nrn.2017.67> PMID: 28626229
29. Daw NDO'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. *Nature*. 2006; 441. <https://doi.org/10.1038/nature04766> PMID: 16778890
30. Wilson RC, Geana A, White JM, Ludvig EA, Cohen JD. Humans use directed and random exploration to solve the explore-exploit dilemma. *J Exp Psychol Gen*. 2014; 143:2074–2081. <https://doi.org/10.1037/a0038199> PMID: 25347535
31. Algermissen J, Swart JC, Scheeringa R, Cools R, den Ouden HEM. Prefrontal signals precede striatal signals for biased credit assignment in motivational learning biases. *Nat Commun*. 2024; 15:19. <https://doi.org/10.1038/s41467-023-44632-x> PMID: 38168089
32. Doll BB, Bath KG, Daw ND, Frank MJ. Variability in dopamine genes dissociates model-based and model-free reinforcement learning. *J Neurosci*. 2016; 36:1211–1222. <https://doi.org/10.1523/JNEUROSCI.1901-15.2016> PMID: 26818509
33. Voon V, Derbyshire K, Ruck C, Irvine MA, Worbe Y, Enander J, et al. Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry*. 2015; 20:345–352. <https://doi.org/10.1038/mp.2014.44> PM—PMID: 24840709 M4—Citavi
34. Master SL, Eckstein MK, Gotlieb N, Dahl R, Wilbrecht L, Collins AGE. Distinguishing the systems contributing to changes in learning during adolescence. *Dev Cogn Neurosci*. 2020; 41. <https://doi.org/10.1016/j.dcn.2019.100732> PMID: 31826837
35. Sheng Y, Dong D, He G, Zhang J. How Noise Can Influence Experience-Based Decision-Making under Different Types of the Provided Information. *Int J Environ Res Public Health*. 2022; 19:10445. <https://doi.org/10.3390/ijerph191610445> PMID: 36012080
36. Herzog N, Waltmann M, Deserno L, Janssen L, Horstmann A. Shared and Differential Neurocognitive Mechanisms in Obesity and Binge Eating Disorder from Adolescence to Adulthood: An Attempt to Improve Prediction of Clinical Outcome. 2020.
37. Voon V, Baek K, Enander J, Worbe Y, Morris LS, Harrison NA, et al. Motivation and value influences in the relative balance of goal-directed and habitual behaviours in obsessive-compulsive disorder. *Transl Psychiatry*. 2015; 5:e670. <https://doi.org/10.1038/tp.2015.165> PM—PMID: 26529423 M4—Citavi
38. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 2011; 69:1204–1215. <https://doi.org/10.1016/j.neuron.2011.02.027> PM—PMID: 21435563 M4—Citavi
39. Guitart-Masip M, Huys QJM, Fuentemilla L, Dayan P, Duzel E, Dolan RJ. Go and no-go learning in reward and punishment: interactions between affect and effect. *Neuroimage*. 2012; 62:154–166. <https://doi.org/10.1016/j.neuroimage.2012.04.024> PM—PMID: 22548809 M4—Citavi
40. Kool W, Cushman FA, Gershman SJ. When Does Model-Based Control Pay Off? *PLoS Comput Biol*. 2016; 12. <https://doi.org/10.1371/journal.pcbi.1005090> PMID: 27564094
41. Giron AP, Ciranka S, Schulz E, Van Den Bos W, Ruggeri A, Meder B, et al. Developmental changes in exploration resemble stochastic optimization. *Nat Hum Behav*. 2023; 7:1955–1967. <https://doi.org/10.1038/s41562-023-01662-1> PMID: 37591981
42. Somerville LH, Sasse SF, Garrad MC, Drysdale AT, Akar NA, Insel C, et al. Charting the expansion of strategic exploratory behavior during adolescence. *J Exp Psychol Gen*. 2017; 146. <https://doi.org/10.1037/xge0000250> PMID: 27977227
43. Pauli R, Brazil IA, Kohls G, Klein-Flügge MC, Rogers JC, Dikeos D, et al. Action initiation and punishment learning differ from childhood to adolescence while reward learning remains stable. *Nat Commun*. 2023; 14:5689. <https://doi.org/10.1038/s41467-023-41124-w> PMID: 37709750
44. Findling C, Wyart V. Computation noise in human learning and decision-making: origin, impact, function. *Curr Opin Behav Sci*. 2021; 38. <https://doi.org/10.1016/j.cobeha.2021.02.018>
45. Findling C, Skvortsova V, Dromnelle R, Palminteri S, Wyart V. Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nat Neurosci*. 2019; 22. <https://doi.org/10.1038/s41593-019-0518-9> PMID: 31659343
46. Juechems K, Balaguer J, Spitzer B, Summerfield C. Optimal utility and probability functions for agents with finite computational precision. *Proc Natl Acad Sci U S A*. 2021; 118. <https://doi.org/10.1073/pnas.2002232118> PMID: 33380453

47. Polanía R, Woodford M, Ruff CC. Efficient coding of subjective value. *Nat Neurosci*. 2019; 22. <https://doi.org/10.1038/s41593-018-0292-0> PMID: 30559477
48. Tamnes CK, Walhovd KB, Grydeland H, Holland D, Østby Y, Dale AM, et al. Longitudinal working memory development is related to structural maturation of frontal and parietal cortices. *J Cogn Neurosci*. 2013; 25. https://doi.org/10.1162/jocn_a_00434 PMID: 23767921
49. Ziegler G, Hauser TU, Moutoussis M, Bullmore ET, Goodyer IM, Fonagy P, et al. Compulsivity and impulsivity traits linked to attenuated developmental frontostriatal myelination trajectories. *Nat Neurosci*. 2019; 22. <https://doi.org/10.1038/s41593-019-0394-3> PMID: 31086316
50. Gregorova K, Eldar E, Deserno L, Reiter AMF. A cognitive-computational account of mood swings in adolescence. 2022. <https://doi.org/10.31234/osf.io/u62vr>
51. Somerville LH, Casey BJ. Developmental neurobiology of cognitive control and motivational systems. *Curr Opin Neurobiol*. 2010; 20. <https://doi.org/10.1016/j.conb.2010.01.006> PMID: 20167473
52. Burnett S, Bird G, Moll J, Frith C, Blakemore S-J. Development during adolescence of the neural processing of social emotion. *J Cogn Neurosci*. 2009; 21:1736–1750. <https://doi.org/10.1162/jocn.2009.21121> PMID: 18823226
53. Shenhav A, Botvinick MM, Cohen JD. The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*. 2013; 79:217–240. <https://doi.org/10.1016/j.neuron.2013.07.007> PMID: 23889930
54. Lieder F, Shenhav A, Musslick S, Griffiths TL. Rational metareasoning and the plasticity of cognitive control. *PLoS Comput Biol*. 2018; 14:1–27. <https://doi.org/10.1371/journal.pcbi.1006043> PMID: 29694347
55. Cools R. The costs and benefits of brain dopamine for cognitive control. *Wiley Interdiscip Rev Cogn Sci*. 2016; 7:317–329. <https://doi.org/10.1002/wcs.1401> PMID: 27507774
56. Cools R. Chemistry of the Adaptive Mind: Lessons from Dopamine. *Neuron*. 2019; 104:113–131. <https://doi.org/10.1016/j.neuron.2019.09.035> PMID: 31600509
57. Kool W, McGuire JT, Rosen ZB, Botvinick MM. Decision Making and the Avoidance of Cognitive Demand. *J Exp Psychol Gen*. 2010; 139:665–682. <https://doi.org/10.1037/a0020198> PMID: 20853993
58. Westbrook A, Kester D, Braver TS. What Is the Subjective Cost of Cognitive Effort? Load, Trait, and Aging Effects Revealed by Economic Preference. *PLoS ONE*. 2013; 8. <https://doi.org/10.1371/journal.pone.0068210> PMID: 23894295
59. Manohar SG, Chong TTJ, Apps MAJ, Batla A, Stamelou M, Jarman PR, et al. Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control. *Curr Biol*. 2015; 25:1707–1716. <https://doi.org/10.1016/j.cub.2015.05.038> PMID: 26096975
60. Chong TTJ, Apps M, Giehl K, Sillence A, Grima LL, Husain M. Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS Biol*. 2017; 15. <https://doi.org/10.1371/journal.pbio.1002598> PMID: 28234892
61. Ma I, Westhoff B, van Duijvenvoorde ACK. Uncertainty about others' trustworthiness increases during adolescence and guides social information sampling. *Sci Rep*. 2022; 12:7634. <https://doi.org/10.1038/s41598-022-09477-2> PMID: 35538170
62. Reiter AMF, Moutoussis M, Vanes L, Kievit R, Bullmore ET, Goodyer IM, et al. Preference uncertainty accounts for developmental effects on susceptibility to peer influence in adolescence. *Nat Commun*. 2021; 12:3823. <https://doi.org/10.1038/s41467-021-23671-2> PMID: 34158482
63. Cremer A, Kalbe F, Müller JC, Wiedemann K, Schwabe L. Disentangling the roles of dopamine and noradrenaline in the exploration-exploitation tradeoff during human decision-making. *Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol*. 2023; 48:1078–1086. <https://doi.org/10.1038/s41386-022-01517-9> PMID: 36522404
64. Dubois M, Habicht J, Michely J, Moran R, Dolan RJ, Hauser TU. Human complex exploration strategies are enriched by noradrenaline-modulated heuristics. *Elife*. 2021; 10. <https://doi.org/10.7554/eLife.59907> PMID: 33393461
65. Deserno L, Boehme R, Mathys C, Katthagen T, Kaminski J, Stephan KE, et al. Volatility Estimates Increase Choice Switching and Relate to Prefrontal Activity in Schizophrenia. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2020; 5:173–183. <https://doi.org/10.1016/j.bpsc.2019.10.007> PMID: 31937449
66. Reiter AMF, Heinze HJ, Schlagenhaut F, Deserno L. Impaired Flexible Reward-Based Decision-Making in Binge Eating Disorder: Evidence from Computational Modeling and Functional Neuroimaging. *Neuropsychopharmacology*. 2017; 42:628–637. <https://doi.org/10.1038/npp.2016.95> PMID: 27301429
67. Schlagenhaut F, Huys QJM, Deserno L, Rapp MA, Beck A, Heinze HJ, et al. Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. *Neuroimage*. 2014; 89. <https://doi.org/10.1016/j.neuroimage.2013.11.034> PMID: 24291614

68. Moutoussis M, Garzón B, Neufeld S, Bach DR, Rigoli F, Goodyer I, et al. Decision-making ability, psychopathology, and brain connectivity. *Neuron*. 2021; 109:2025–2040.e7. <https://doi.org/10.1016/j.neuron.2021.04.019> PMID: 34019810
69. Hauser TU, Iannaccone R, Ball J, Mathys C, Brandeis D, Walitza S, et al. Role of the medial prefrontal cortex in impaired decision making in juvenile attention-deficit/hyperactivity disorder. *JAMA Psychiatry*. 2014; 71. <https://doi.org/10.1001/jamapsychiatry.2014.1093> PMID: 25142296
70. Addicott MA, Pearson JM, Schechter JC, Sapyta JJ, Weiss MD, Kollins SH. Attention-deficit/hyperactivity disorder and the explore/exploit trade-off. *Neuropsychopharmacology*. 2021; 46:614–621. <https://doi.org/10.1038/s41386-020-00881-8> PMID: 33040092
71. Aster H-C, Waltmann M, Busch A, Romanos M, Gamer M, Maria van Noort B, et al. Impaired flexible reward learning in ADHD patients is associated with blunted reinforcement sensitivity and neural signals in ventral striatum and parietal cortex. *NeuroImage Clin*. 2024; 42:103588. <https://doi.org/10.1016/j.nicl.2024.103588> PMID: 38471434
72. Van den Driessche C, Chevrier F, Cleeremans A, Sackur J. Lower Attentional Skills predict increased exploratory foraging patterns. *Sci Rep*. 2019; 9:10948. <https://doi.org/10.1038/s41598-019-46761-0> PMID: 31358789
73. Dubois M, Bowler A, Moses-Payne ME, Habicht J, Moran R, Steinbeis N, et al. Exploration heuristics decrease during youth. *Cogn Affect Behav Neurosci*. 2022; 22:969–983. <https://doi.org/10.3758/s13415-022-01009-9> PMID: 35589910
74. Meder B, Wu CM, Schulz E, Ruggeri A. Development of directed and random exploration in children. *Dev Sci*. 2021; 24:e13095. <https://doi.org/10.1111/desc.13095> PMID: 33539647
75. Chakroun K, Mathar D, Wiehler A, Ganzer F, Peters J. Dopaminergic modulation of the exploration/exploitation trade-off in human decision-making. Gershman SJ, Frank MJ, Gershman SJ, Averbeck BB, Pearson J, editors. *Elife*. 2020; 9:e51260. <https://doi.org/10.7554/eLife.51260> PMID: 32484779
76. Engert V, Pruessner J. Dopaminergic and Noradrenergic Contributions to Functionality in ADHD: The Role of Methylphenidate. *Curr Neuropharmacol*. 2009; 6. <https://doi.org/10.2174/157015908787386069> PMID: 19587853
77. Frank MJ, Santamaria A, O'Reilly RC, Willcutt E. Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology*. 2007; 32. <https://doi.org/10.1038/sj.npp.1301278> PMID: 17164816
78. Hauser TU, Fiore VG, Moutoussis M, Dolan RJ. Computational Psychiatry of ADHD: Neural Gain Impairments across Marrian Levels of Analysis. *Trends Neurosci*. 2016; 39:63–73. <https://doi.org/10.1016/j.tins.2015.12.009> PMID: 26787097
79. Volkow ND, Wang G-J, Kollins SH, Wigal TL, Newcorn JH, Telang F, et al. Evaluating Dopamine Reward Pathway in ADHD: Clinical Implications. *JAMA*. 2009; 302:1084–1091. <https://doi.org/10.1001/jama.2009.1308> PMID: 19738093
80. Yokokura M, Takebasashi K, Takao A, Nakaizumi K, Yoshikawa E, Futatsubashi M, et al. In vivo imaging of dopamine D1 receptor and activated microglia in attention-deficit/hyperactivity disorder: a positron emission tomography study. *Mol Psychiatry*. 2021; 26. <https://doi.org/10.1038/s41380-020-0784-7> PMID: 32439845
81. Cortese S, Adamo N, Giovane CD, Mohr-Jensen C, Hayes AJ, Carucci S, et al. Comparative efficacy and tolerability of medications for attention-deficit hyperactivity disorder in children, adolescents, and adults: a systematic review and network meta-analysis. *Lancet Psychiatry*. 2018; 5:727–738. [https://doi.org/10.1016/S2215-0366\(18\)30269-4](https://doi.org/10.1016/S2215-0366(18)30269-4) PMID: 30097390
82. Swart JC, Frank MJ, Määttä JI, Jensen O, Cools R, den Ouden HEM. Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. *PLoS Biol*. 2018; 16:e2005979. <https://doi.org/10.1371/journal.pbio.2005979> PMID: 30335745
83. Christakou A, Gershman SJ, Niv Y, Simmons A, Brammer M, Rubia K. Neural and psychological maturation of decision-making in adolescence and young adulthood. *J Cogn Neurosci*. 2013; 25. https://doi.org/10.1162/jocn_a_00447 PMID: 23859647
84. Hauser TU, Iannaccone R, Walitza S, Brandeis D, Brem S. Cognitive flexibility in adolescence: Neural and behavioral mechanisms of reward prediction error processing in adaptive decision making during development. *Neuroimage*. 2015; 104:347–354. <https://doi.org/10.1016/j.neuroimage.2014.09.018> PMID: 25234119
85. Rosenbaum GM, Grassie HL, Hartley CA. Valence biases in reinforcement learning shift across adolescence and modulate subsequent memory. *Elife*. 2022; 11. <https://doi.org/10.7554/eLife.64620> PMID: 35072624
86. d'Acremont M, Van der Linden M. Gender differences in two decision-making tasks in a community sample of adolescents. *Int J Behav Dev*. 2006; 30:352–358. <https://doi.org/10.1177/0165025406066740>

87. Piray P, Dezfouli A, Heskes T, Frank MJ, Daw ND. Hierarchical Bayesian inference for concurrent model fitting and comparison for group studies. *PLoS Comput Biol*. 2019; 15:e1007043. <https://doi.org/10.1371/journal.pcbi.1007043> PMID: 31211783
88. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. Bayesian model selection for group studies. *Neuroimage*. 2009; 46:1004–1017. <https://doi.org/10.1016/j.neuroimage.2009.03.025> PMID: 19306932
89. Rigoux L, Stephan KE, Friston KJ, Daunizeau J. Bayesian model selection for group studies—Revisited. *Neuroimage*. 2014; 84:971–985. <https://doi.org/10.1016/j.neuroimage.2013.08.065> PMID: 24018303
90. Rescorla RA, Wagner AR. A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. *Classical Cond II Curr Res Theory*. 1972.
91. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc*. 1995; 90. <https://doi.org/10.1080/01621459.1995.10476572>
92. Xia L, Master SL, Eckstein MK, Baribault B, Dahl RE, Wilbrecht L, et al. Modeling changes in probabilistic reinforcement learning during adolescence. *PLoS Comput Biol*. 2021; 17. <https://doi.org/10.1371/journal.pcbi.1008524> PMID: 34197447