NeuroImage: Clinical 42 (2024) 103588

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Impaired flexible reward learning in ADHD patients is associated with blunted reinforcement sensitivity and neural signals in ventral striatum and parietal cortex

Hans-Christoph Aster [a,*], Maria Waltmann [a,b], Anika Busch [a], Marcel Romanos [a], Matthias Gamer [c], Betteke Maria van Noort [d,e], Anne Beck [f,g], Viola Kappel [d], Lorenz Deserno [a,b,h]

[a] Department of Child and Adolescent Psychiatry, Psychotherapy and Psychosomatics, University Hospital Würzburg, Würzburg, Germany
[b] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[c] Department of Psychology, University of Würzburg, Würzburg, Germany
[d] Department of Child and Adolescent Psychiatry, Charité University Medicine, Campus Virchow Klinikum, Berlin, Germany
[e] MSB Medical School Berlin, Department of Psychology, Germany
[f] Department of Psychiatry and Neurosciences, Charité University Medicine, Berlin, Germany
[g] Department of Psychology, Faculty of Health, Health and Medical University, Potsdam, Germany
[h] Department of Psychiatry and Psychotherapy, Technische Universität Dresden, Dresden, Germany

ARTICLE INFO

Keywords:
ADHD
Reinforcement learning
Functional MRI
Dopamine
Prediction errors

ABSTRACT

Reward-based learning and decision-making are prime candidates to understand symptoms of attention deficit hyperactivity disorder (ADHD). However, only limited evidence is available regarding the neurocomputational underpinnings of the alterations seen in ADHD. This concerns flexible behavioral adaption in dynamically changing environments, which is challenging for individuals with ADHD. One previous study points to elevated choice switching in adolescent ADHD, which was accompanied by disrupted learning signals in medial prefrontal cortex.

Here, we investigated young adults with ADHD (n = 17) as compared to age- and sex-matched controls (n = 17) using a probabilistic reversal learning experiment during functional magnetic resonance imaging (fMRI). The task requires continuous learning to guide flexible behavioral adaptation to changing reward contingencies. To disentangle the neurocomputational underpinnings of the behavioral data, we used reinforcement learning (RL) models, which informed the analysis of fMRI data.

ADHD patients performed worse than controls particularly in trials before reversals, i.e., when reward contingencies were stable. This pattern resulted from 'noisy' choice switching regardless of previous feedback. RL modelling showed decreased reinforcement sensitivity and enhanced learning rates for negative feedback in ADHD patients. At the neural level, this was reflected in a diminished representation of choice probability in the left posterior parietal cortex in ADHD. Moreover, modelling showed a marginal reduction of learning about the unchosen option, which was paralleled by a marginal reduction in learning signals incorporating the unchosen option in the left ventral striatum.

Taken together, we show that impaired flexible behavior in ADHD is due to excessive choice switching ('hyper-flexibility'), which can be detrimental or beneficial depending on the learning environment. Computationally, this resulted from blunted sensitivity to reinforcement of which we detected neural correlates in the attention-control network, specifically in the parietal cortex. These neurocomputational findings remain preliminary due to the relatively small sample size.

* Corresponding author at: Department of Child and Adolescent Psychiatry, Psychotherapy and Psychosomatics, University Hospital Würzburg, Margarete-Höppel-Platz 1, Würzburg 97080.
E-mail address: Aster_H@ukw.de (H.-C. Aster).

https://doi.org/10.1016/j.nicl.2024.103588
Received 2 June 2023; Received in revised form 6 February 2024; Accepted 28 February 2024
Available online 1 March 2024

1. Introduction

Attention deficit hyperactivity disorder (ADHD), a common child and adolescent psychiatric disorder (Faraone et al., 2015), is characterized by its core symptoms of hyperactivity, inattention, and impulsivity. Reward-based learning and decision-making are prime candidates that may underlie symptoms, as alterations were reported in some of these domains (Mowinckel et al., 2015; Marx et al., 2021). However, only limited evidence is available regarding the neurocomputational underpinnings of reward learning and decision-making in ADHD. This is particularly true with respect to flexible behavioral adaption in dynamically changing environments, which may be challenging for individuals with ADHD (Humphreys et al., 2018) due to attentional and learning deficits. Fig. 1. Example of a sequence from the reversal learning task (adapted from Schlagenhauf et al., 2014.

How individuals learn from positive and negative reward feedback and guide decisions accordingly can be formalized by computational models of reinforcement learning (Sutton and Barto, 1998). At the core of RL models are reward prediction errors (RPEs), which reflect the differences between delivered and expected reward. Neurally, prediction errors are signaled by phasic release of midbrain dopamine (Hollerman and Schultz, 1998, Schultz, 2013), with corresponding echoes of neural activity in the striatum as well as other brain regions (Pine et al., 2018). Human functional neuroimaging studies reported correlates of RPEs in the midbrain, striatum and several cortical regions (O'Doherty et al., 2004; D'Ardenne et al., 2008; Daw et al., 2011; Deserno et al., 2015b). Individual differences in neurobehavioral correlates of RL have been indeed linked to a variety of dopamine measures available in humans, including pharmacological manipulations (Pessiglione et al., 2006; Westbrook et al., 2020; Rostami Kandroodi et al., 2021; Deserno et al., 2021), neurochemical positron emission tomography (PET) (Deserno et al., 2015b; Westbrook et al., 2020; Calabro et al., 2023) and specific genotypes (Frank et al., 2007; Dreher et al., 2009).

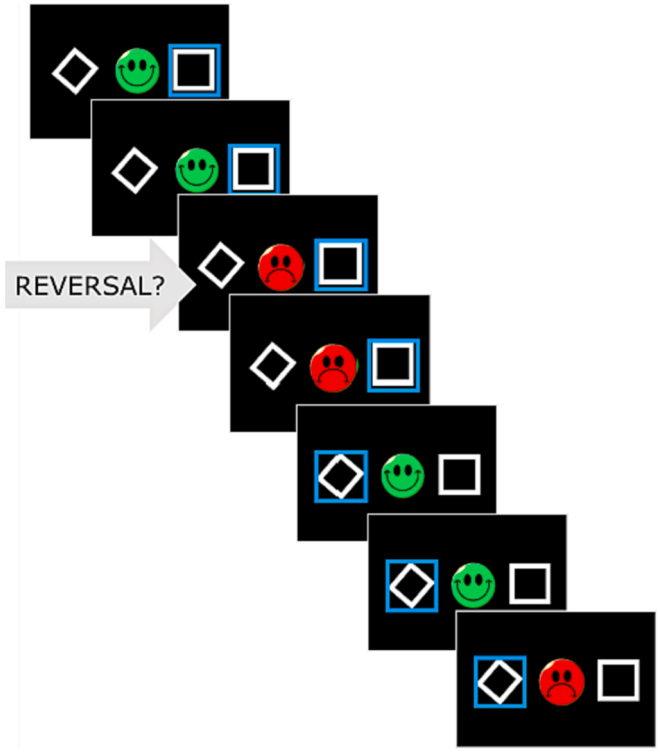

Fig. 1. Example of a sequence from the reversal learning task (adapted from Schlagenhauf et al., 2014, Neuroimage). The subjects need to decide between two geometric figures and receive feedback in the form of a smiley face. The probability of which figure is most likely to elicit positive feedback changes over the course of the experiment (reversal).

In patients with ADHD, neurochemical studies reported altered dopamine neurotransmission and presumably lower baseline dopamine levels (Fusar-Poli et al., 2012). Brain activation, measured with fMRI during the anticipation and delivery of rewards, was reported to be disrupted (Plichta and Scheres, 2014; von Rhein et al., 2015), in particular in the ventral striatum during reward anticipation (Plichta and Scheres, 2014). This line of work supports hypothetical alterations in RL and its neural underpinnings. However, evidence based on studies that directly test learning and apply computational modeling (Ziegler et al., 2016; Véronneau-Veilleux et al., 2022) is missing. A particular challenging scenario for individuals with ADHD is not only to learn from reward to guide decision-making but also to strike a balance between exploration and exploitation when action-outcome contingencies change dynamically. This capacity can be examined using reversal learning (Reiter et al., 2017; Waltmann et al., 2023). Reinforcement learning has been shown to undergo substantial neurodevelopmental changes (Nussenbaum and Hartley, 2019; Weiss et al., 2021; Scholz et al., 2023; Waltmann et al., 2023), and has been used to study a wide range of psychiatric disorders (Chantiluke et al., 2015; Geisler et al., 2017; Reiter et al., 2017). Yet, there is only one study available that directly examined RL in adolescent ADHD patients during fMRI (Hauser et al., 2014). This study revealed noisy switching behavior in ADHD patients, which may computationally arise from enhanced levels of decision noise, an impairment in distinctly representing values of alternative choice options. In the study by Hauser et al., (2014) this was accompanied by reduced activation to RPEs in the medial prefrontal cortex. Our study aimed to extend these findings by investigating RL in adult ADHD patients. Furthermore, this study explored for the first time explicitly whether impaired learning of the selected action or impaired simultaneous learning of the unselected action caused the difficulties in RL.

In n = 17 patients and n = 17 controls, we closely followed the study by Hauser et al., (2014) by examining reversal learning during fMRI with extended RL modelling and more detailed computational fMRI analysis. We hypothesized that altered task performance would be driven by noisy choice switching, computationally accounted for by enhanced decision noise. In our RL models, we addressed not only learning from the chosen option (single-update learning) but also learning from the option that was not chosen (double-updating). Thus, we explored whether differences in these types of learning also contributed to the observed behavioral alterations seen in ADHD. On the neural level, we dissociated correlates of RPE with respect to single- and double-update learning. We further analyzed the neural correlates of choice probability, which closely reflects decision noise. We focused these analyses on the ventromedial prefrontal cortex and the ventral striatum, which were previously reported to be altered in ADHD, and where we expected reduced correlates of RPEs and choice probability.

2. Methods

2.1. Study protocol

Before participation, all participants provided written informed consent. Ethical approval was obtained through the ethics committee of the German Psychological Society (DGPs registration number: HSAAS04082008DGPS). Data were collected between 2008 and 2011.

All participants completed several diagnostic and neuropsychological assessments before the MRI acquisition. All patients fulfilled the DSM-IV-TR criteria for ADHD combined subtype as assessed by clinical experts with the structured assessment scale 'ADHD Diagnostic Checklist' (Rösler et al., 2005). ADHD symptomatology in childhood was assessed retrospectively with the 'Wender Utah Rating Scale – German short form' (WURS, (Retz-Junginger et al., 2002)). The current severity of ADHD symptomatology was examined via the 'Conner's Adult ADHD Rating Scale' (CAARS, (Christiansen et al., 2013)). To exclude the

presence of other Axis I or Axis II disorders, subjects were interviewed using the SCID-I and –II (Wittchen and Pfister, 1997). Furthermore, the specific presence of substance abuse, amongst others due to its role in reward processing, was examined via the Composite International Diagnostic Interview (Wittchen and Pfister, 1997). To rule out previous or current ADHD symptoms or other psychiatric disorders in the control group, we used the same diagnostic assessments. Finally, handedness was assessed via the 'Handedness Questionnaire' (Coren, 1993). Forty-eight hours before the study appointment, the ADHD patients discontinued their intake of psychostimulants.

The neuropsychological assessment consisted of a language-independent measure for IQ, the Culture Fair Test (CFT-20-R (Weiß and Weiß 2008)), a Digit Span task (Von Aster et al., 2006) and the Trail-Making-Test (TMT) Part A and B (Reitan, 1958). The Digit Span task measures verbal working memory capacity. It consists of two conditions: forward and backward. A span of 6–7 is considered an average score. The TMT assesses visual attention and processing speed in Part A and executive control and flexibility in Part B. The TMT Part B and A difference score provides a more precise measure of task-switching ability. Depending on the homogeneity of variances and normal distribution of the data, the neuropsychological data were compared between groups via independent samples t-tests, Mann-Whitney-U tests or Welch tests using the Jasp Toolbox (JASP Team (2022). JASP (Version 0.16.4) [Computer software]). Alpha was set at 0.05.

Exclusion criteria for HCs were 1) left-handedness, 2) current psychiatric diagnosis according to ICD-10 or DSM-IV-TR, except alcohol abuse, 3) the presence of neurological disorders, 4) a first-degree family member suffering from a neurological or psychiatric disorder or 5) currently taking psychotropic medication. HCs were recruited via advertisements in the community.

2.2. Reversal learning paradigm

During functional MRI (fMRI) acquisition, participants performed a reversal learning task (Schlagenhauf et al., 2013; Schlagenhauf et al., 2014; Deserno et al., 2015a). The task required participants to choose between one of two geometric figures with different reward probabilities. After each choice, they received positive (green smiley) or negative feedback (red frowning face) (Fig. 1). The chosen stimulus and the feedback remained visible for 1 s. If participants did not choose a target within 2 s, the trial was rated as incorrect. A fixation cross was shown between the trials. The interval had a varying duration of 1 to 6.5 s (exponentially distributed). The task consisted of two runs of 100 trials each with a short break in between. During each run, the participants were exposed to three types of blocks in which the reward probabilities for a correct choice (right figure vs. left figure) were 80:20, 20:80 and 50:50. The block changed based on performance, i.e., after a minimum of 10 subsequent trials if participants reached 70 % correct choices, or automatically after a maximum of 16 trials. Thus, each block was encountered between twice and four times during each run. Learning can only take place between in the 80:20 and 20:80 blocks and is not possible in 50:50 conditions. Hence, only the former block types were included in the initial analysis of choice behavior.

2.3. Analysis of behavioral data

Trials with correct choices (regardless of whether positive feedback was actually received as a result of the 80/20 probability), coded as 1 vs. 0, as well as trials with a different chosen response as in the previous trial (Switching, coded as 1 vs. 0) were extracted for each trial. An analysis of reaction times can be found in the supplementary information (Supplementary Figs. S2 and S3).

As the task has phases with constant and changing probabilities of outcomes, we investigated their impact on predicting accuracy in our supplementary analysis. This was achieved by integrating diverse interpretations of task dynamics into the model. Following Waltmann et al.'s methodology (Waltmann et al., 2023), we evaluated four distinct strategies for this integration. The model that differentiated between pre-reversal and post-reversal trials provided the most optimal fit, and the results derived from this model are presented in our manuscript (Supplementary Fig. S1).

Generalized linear mixed models were used to analyze behavior. We used a full random structure (random intercepts, random slopes, correlation of slopes) (Barr et al., 2013). A binomial family was chosen and the logit link function was used:

1. [Correct choices ~ Group * Phase + (1 + Phase |subject)].

To analyze accuracy, we use a mixed effects logistic regression predicting correct choices from group (referring to ADHD or control) and task phase (indicating pre- or post-reversal trial), as well as their interaction. The full random structure allowed individual slopes and intercepts per subject.

2. [Switching ~ Group * Previous feedback + (1 + Previous feedback | subject)].

To analyze switching, we use a mixed effects logistic regression predicting trials where subjects switched the chosen response option compared to the previous trials from the factors group (referring to ADHD or control) and previous feedback (indicating a win or loss in the previous trial). Again, the full random structure allowed for individual slopes and intercepts per subject.

2.4. Computational modelling of reinforcement learning

We analyzed the behavioral data using Q-learning models of reinforcement learning (Watkins and Dayan, 1992). Thus, for each model, we identified the parameters which best accounted for each individual's observed history of choices and outcomes. The model fitting was conducted on the data of all the subjects from both groups. Our initial model was a single-update model, which only updates (or learns) the value of the chosen action $Q_{a,t}$: $Q_{a,t} + 1 = Q_{a,t} + \alpha(r - Q_{a,t})$. This value is updated in each trial by the prediction error $\delta = r - Q_{a,t}$. The rate to which prediction errors influence the update of Q value is captured by the learning rate α. Because reward probabilities of the two available actions in the reversal learning task are perfectly anti-correlated, the feedback of the chosen action could also influence the Q-value of the non-chosen action. We therefore additionally defined a double-update model, which updates both actions simultaneously to opposite directions: $Q_{a(unchosen),t} + 1 = Q_{a(unchosen),t} + \alpha((-r) - Q_{a(unchosen),t})$. It is conceivable that individuals vary in the degree of using double updating and thus we included a weighting parameter (κ) that quantifies the degree of double-updating in some models: $Q_{a(unchosen),t} + 1 = Q_{a(unchosen),t} + \kappa\alpha((-r) - Q_{a(unchosen),t})$.

Further, since there could be inter-individual differences in the extent to which the current prediction error impacts updating depending on positive or negative feedback (Eppinger and Kray, 2011, Cazé and van der Meer, 2013), different learning rates for wins and losses were implemented in some models ($Q_{a,t+1} = Q_{a,t} + \alpha_{win/loss}(r - Q_{a,t})$).

Lastly, we examined decision noise, which is determined by the degree to which values of choice options are represented distinctly. Typically, this is determined by passing values to a sigmoid softmax function with an individually varying steepness parameter, which scales the differences between values and thus determines choice probabilities. Here, we used a slightly different but largely equivalent approach by implementing a reinforcement sensitivity parameter (ρ): $Q_{a,t} + 1 = Q_{a,t} + \alpha(\rho r - Q_{a,t})$. ρ is a free parameter that determines the maximum difference between values by defining the upper bound of the Q-values. These values scaled by ρ are then entered into a softmax function with steepness fixed to 1. While it was shown that this has an equivalent effect on choice probability as a steepness parameter, it is straightforward to

implement differences in reinforcement sensitivity to positive and negative outcomes (Huys et al., 2013). Further, reinforcement sensitivities have improved estimation properties (Huys et al., 2013, Katahira, 2015) and higher reliability (Waltmann et al., 2022). Thus, in some of our models, reinforcement sensitivity was again distinguished for sensitivity to positive and negative feedback.

To summarize, by combining different learning rates and levels of reinforcement sensitivity, four models were created. These models were then estimated for single-update, double-update, and variable double-update scenarios, resulting in a total of 12 fitted models. To compare these models, the integrated Bayesian Information Criterion (iBIC) was used.

For hierarchical model estimation, we used the emfit toolbox in MATLAB R2020b (Huys, 2017). Model estimation was performed to obtain maximum a posteriori estimation with empirical priors based on the trial-by-trial data of all participants. We have previously shown that this hierarchical estimation leads to improved reliability (Waltmann et al., 2022). An expectation maximization procedure was used (Huys et al., 2012). Since the model with only one reinforcement sensitivity can logically only assign positive values for the sensitivity, its value was transformed exponentially to ensure positive values. To keep the learning rate (α) and weighting parameter (κ) between 0 and 1, these parameters were inverse logit transformed.

The modeling parameters reinforcement sensitivity and learning rate (both for positive and negative feedback) were analyzed with linear mixed models using the Jasp Toolbox (JASP Team (2022). JASP (Version 0.16.4) [Computer software]. This resulted in two models with full random structure (1. Reinforcement sensitivity ~ group * feedback + (1 + feedback |subject), 2. Learning rate ~ group * feedback + (1 + feedback |subject)). The parameter kappa, which expresses the weighting between single and double updating, was compared between groups with a *t*-test after testing for normal distribution and equality of variance. Pearson correlation coefficients were calculated to discover possible associations between symptom expression and modeling parameters in an exploratory analysis. All variables were z-standardized before the correlation analysis.

2.5. Functional MRI data acquisition

Imaging was conducted using a 3 Tesla GE Sigma Scanner with an eight channel head coil to acquire gradient echo T2*-weighted echo-planar images with blood oxygenation level-dependent (BOLD) contrast. Twenty-nine slices were acquired, covering the whole brain, with 4 mm thickness, 2×2 mm^2 in-plane voxel resolution, repetition time (TR) = 2.3 ms, echo time (TE) = 27 ms and a flip angle α = 90°. T1-weighted structural images were acquired with TR = 7.8 ms, TE = 3.2 ms, α = 20°, matrix size = 256×256, slice thickness = 1 mm, voxel size = 1×1×1 mm. Right before the MRI acquisition, all participants were vigilant as assessed by the Stanford Sleepiness Scale (M_{ADHD} = 2.12 ±0.60, M_{HC} = 2.35±0.70; p = 0.301, d = −0.35).

2.6. Functional MRI data preprocessing

fMRI data were analysed with SPM8 (Wellcome Department of Imaging Neuroscience). ArtRepair was used to remove noise spikes and to repair bad slices within a particular scan and bad slices were repaired by interpolation between adjacent slices (Mazaika et al., 2005). Data was then corrected for delay of slice time acquisition and was motion corrected using realignment. The images were then registered into the Montreal Neurological Institute (MNI) space by using the normalised parameters generated during the segmentation of each participant's anatomical T1-image (Ashburner and Friston, 2005). Spatial smoothing with an isotropic Gaussian kernel of 8 mm full width at half-maximum (FWHM) kernel was applied to the images.

2.7. Model informed fMRI Analysis

In the first-level general linear model, onsets of feedback, cue and missing trials were convolved with the hemodynamic response function and the 6 motion parameters were added as regressors of no interest. As orthogonalized parametric modulators on the feedback regressor, we added, for each person, the trial-by-trial prediction errors (PEs) from the best fitting RL model. This included, first, the single update (SU) PEs from the best SU model and the double update (DU) PEs from the best-fitting model. Due to high collinearity between PEs and to isolate unique variance of the DU PEs, we subtracted SU PEs from DU PEs for each trial (Daw et al., 2011). This approach has already been applied successfully in previous studies (Reiter et al., 2016, Reiter et al., 2017, Waltmann et al., 2023). As orthogonalized parametric modulators to the cue onset, we added two model-derived trial-by-trial regressors. The choice probability maps the individual expected values of the choices per trial which are drawn from the best fitting DU model. The larger the difference in expected values between the two choices, the more likely an individual will choose one of the two options. From the choice probabilities, we constructed a regressor reflecting trial-by-trial model-fit, where choices predicted with below-chance accuracy (<50 %) were coded as 1 (noisy or explorative behavior) and 0 otherwise. This regressor addresses brain activation associated with noisy or explorative behavior and removes variance solely associated with poor model fit (Waltmann et al., 2023).

At the second level, a full factorial model was used on SU PEs and DU PEs with group and type of RPE as predictors. Separate between-group t-tests were calculated for choice probability and exploratory trials. Results were adjusted at the peak level for multiple comparisons using the family-wise error control. Small volume correction was performed using the following *a priori* regions of interests (ROIs): 1) the ventral striatum, using an anatomical definition of the nucleus accumbens (as obtained in the IBASPM atlas as part of the WFU Pick Atlas) with respect to SU and DU PEs; 2) the ventromedial prefrontal cortex (vmPFC) because of its central role in choice value, which is closely linked to DU PEs and choice probability. The vmPFC ROI was defined using a functional ROI of the effects of DU RPE and choice probability, respectively, published by a previous independent study on development of reversal learning (Waltmann et al., 2023); 3) a functional ROI from the same previous study (Waltmann et al., 2023) reflecting brain activation to noisy/explorative behavior, covering parts of the insula, thalamus, vmPFC and parietal cortex (Supplementary Fig. S4).

3. Results

3.1. Descriptive statistics

17 ADHD patients and 17 age- and gender-matched healthy controls were included. Except for two left-handed ADHD patients, all participants were right-handed. One female participant was included in each group. According to the CIDI DIA-X screening interview, two subjects in each group fulfilled the diagnostic criteria of alcohol abuse (F10.1). Ten subjects in the ADHD group reported nicotine use, of which two subjects met criteria for nicotine dependence (F17.2). Four subjects in the control group reported nicotine use. In the ADHD group, two subjects had not previously been treated with stimulants, seven had been treated with methylphenidate in the past, and nine were still taking methylphenidate (but discontinued the medication 48 h prior to the study appointment). As expected, compared with healthy controls, the ADHD group reported stronger ADHD symptom ratings in the CAARS and WURS-K questionnaires, but no differences in other psychiatric symptom ratings according to the Symptom Checklist (SCL-90) (Derogatis and Savitz, 1999). Descriptive group statistics are presented in Table 1.

A detailed summary of the neuropsychological testing is presented in Table 2. The ADHD group had a lower intelligence in comparison to controls and performed worse than controls in working memory (digit span) and processing speed (TMT) domains.

Table 1
Sample Description of Age, Clinical Symptoms and Handedness.

	ADHD	HC				
	M ± *SD*	*M* ± *SD*	*t-value*	*df*	*p*	*Effect Size*
Age (in years)	22.14 ± 4.07	23.58 ± 3.47	189.000[c]		0.131	0.31
CAARS (t-score)						
Inattention	57.53 ± 12.25	46.25 ± 6.94[a]	3.278[b]	26	0.003*	1.13
Hyperactivity	57.47 ± 8.52	42.94 ± 6.32[a]	5.536	31	<0.001**	1.94
Impulsivity	54.12 ± 10.59	42.56 ± 7.07[a]	3.661	31	0.001**	1.28
Self-Concept	52.76 ± 13.98	43.88 ± 5.03[a]	2.457[b]	20	0.023	0.85
ADHD Index	60.65 ± 11.77	43.13 ± 7.43[a]	5.075	31	<0.001**	1.78
GSI t-score (SCL-90-R)	53.06 ± 9.50	50.88 ± 7.33	0.748	32	0.460	0.26
WURS-K (raw score)	42.76 ± 14.93	22.18 ± 9.14	4.850	32	<0.001**	1.66
Handedness (raw score)	32.06 ± 7.81	35.12 ± 1.58	1.584[b]	17	0.131	−0.54

Note. CAARS = Conners Adult ADHD Rating Scale (Conners et al., 1998), GSI = Global Severity Index, SCL-90-R = Symptom Checklist-90-R (Franke, 2002), WURS-K, Wender Utah Rating Scale' (Rösler et al., 2008). For the Student *t*-test and the Welch *t*-test, effect size is given by Coheńs *d*. For the Mann-Whitney *U* test effect size is given by the rank biserial correlation.
[a] $n = 16$. [b] Welch T-Test as equal variances were not assumed. [c] Mann-Whitney U-Test as not normally distributed. * $p < 0.01$; ** $p < 0.001$.

Table 2
Neuropsychological Performance of ADHD Patients versus Healthy Controls.

	ADHD	HC				
	M ± *SD*	*M* ± *SD*	*t-value*	*df*	*p*	*d*
IQ (CFT-20-R)	98.18 ± 15.96	108.82 ± 7.86	2.468[a]	23	0.021*	-0.85
Digit Span						
Forward	6.65 ± 1.84	8.47 ± 2.00	2.767	32	0.009**	-0.95
Backward	6.00 ± 1.73	7.53 ± 1.91	2.447	32	0.020*	-0.84
Total	12.65 ± 3.26	16.06 ± 3.53	2.930	32	0.006**	−1.00
TMT						
Part A (in sec.)	28.09 ± 5.16	23.61 ± 7.65	2.005	32	0.053	0.69
Part B (in sec.)	79.12 ± 20.92	60.12 ± 18.77	2.787	32	0.009**	0.96
Part B minus Part A	51.02 ± 21.09	36.51 ± 16.39	2.240	32	0.032*	0.82

Note. IQ = intelligence quotient, CFT-20-R = Culture Fair Test 20, revised (Weiß & Weiß, 2008), Digit Span (von Aster et al., 2006), TMT = Trail-making-test (Reitan, 1958). [a] Welch T-Test as equal variances were not assumed. * $p < 0.05$; ** $p < 0.01$.

3.2. Behavorial data

Accuracy differed only marginally between phases ($t = 1.75$ ($df = 7$), $p = 0.08$) and not between groups ($t = 0.08$ ($df = 7$), $p = 0.94$). However, there was a significant group*phase interaction effect, as ADHD patients performed better in the post-reversal phase and worse in the pre-reversal phase ($t = 4.70$ ($df = 7$), $p < 0.001$, Fig. 2a).

All subjects were more likely to switch after previous negative feedback (feedback effect: $t = 8.14$ ($df = 7$), $p < 0.001$). ADHD patients were more likely to switch (group effect: $t = 4.12$ ($df = 7$), $p < 0.001$), irrespectively of previous feedback (group*feedback effect: $t = 0.97$ ($df = 7$), $p = 0.33$, see Fig. 2b).

3.3. Computational modeling of behavior: model comparison

We compared a total of twelve RL models with respect to their evidence to account for the data based on the integrated Bayesian Information Criterion (Huys et al., 2012). The double update model with separate learning rates and reinforcement sensitivities, as well as weighting of single and double updating, accounted best for the current behavioral data (see Fig. 3a).

3.4. Computational modeling of behavior: Model parameters

ADHD patients showed an overall lower reinforcement sensitivity (group effect: $t = 4.30$ ($df = 6$), $p < 0.001$), especially for positive feedback (group*feedback effect: $t = 4.00$ ($df = 6$), $p < 0.001$), see Fig. 3b). The learning rate of ADHD patients was increased compared to healthy controls (group effect: $t = 2.50$ ($df = 6$), $p = 0.016$), but this effect was mainly driven by the higher learning rate for negative feedback (group*feedback effect: $t = 3.20$ ($df = 6$), $p = 0.003$, see Fig. 3d). The parameter kappa, which defines the strength of the update weighting between chosen and unchosen option, showed a marginal difference between groups. ADHD patients updated the selected option slightly stronger than the unselected option compared to healthy controls ($p = 0.09$, Cohen's $d = 0.59$, see Fig. 3c).

3.5. Correlations

In an exploratory analysis, which was not corrected for multiple comparisons, we correlated all five modeling parameters and the three core symptoms in ADHD patients (inattention, hyperactivity and impulsivity). Stronger hyperactivity symptoms ($r = -0.50$, $p = 0.04$) and marginally stronger impulsivity symptoms ($r = -0.42$, $p = 0.09$) were associated with a weaker updating of the unchosen option. Stronger impulsivity symptoms in ADHD patients were associated with lower learning rate for positive feedback ($r = -0.51$, $p = 0.03$). Stronger hyperactivity symptoms were marginally associated with a lower learning rate for negative feedback ($r = -0.5$, $p = 0.07$). Scatter plots of the significant correlations are shown in Supplementary Fig. S5a and b. Exploratory analysis showed no further associations between modeling parameters and clinical symptoms ($r < 0.27$, $p > 0.3$).

3.6. fMRI analysis

Across both groups, single update prediction errors were significantly correlated with activity in the left and right ventral striatum (xyz: −13/8/−15, $t = 3.51$, $p_{FWE} = 0.003$, cluster size (k): 31 voxel; xyz: 12/

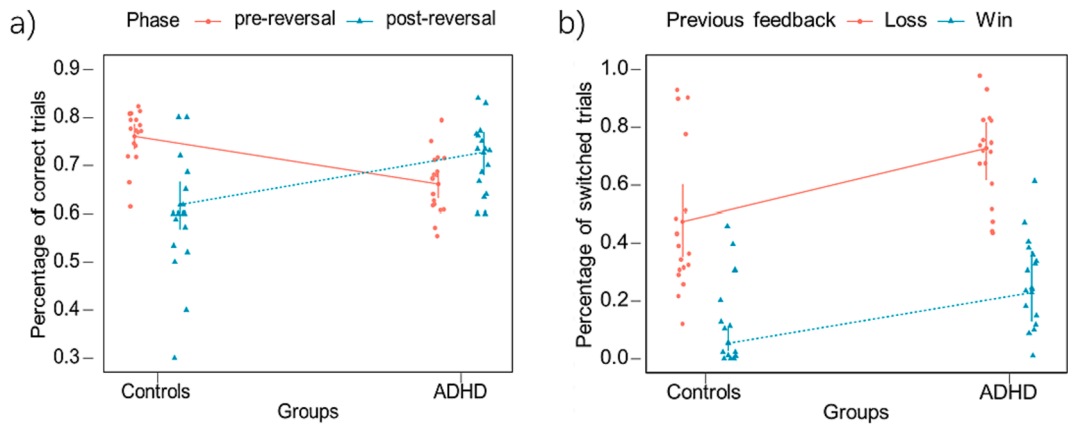

Fig. 2. **A)** While controls chose the correct option more often during the stable pre-reversal phases, ADHD patients chose the correct answer more frequently during the post-reversal phase. Error bars depict standard errors of the mean. **b)** Subjects chose an option more often if it had been rewarded previously. ADHD patients were less likely to choose the same option twice in a row, regardless of the previous feedback. **c)** Both groups were faster to respond if they had lost in the previous trial. ADHD patients responded slower overall than controls.

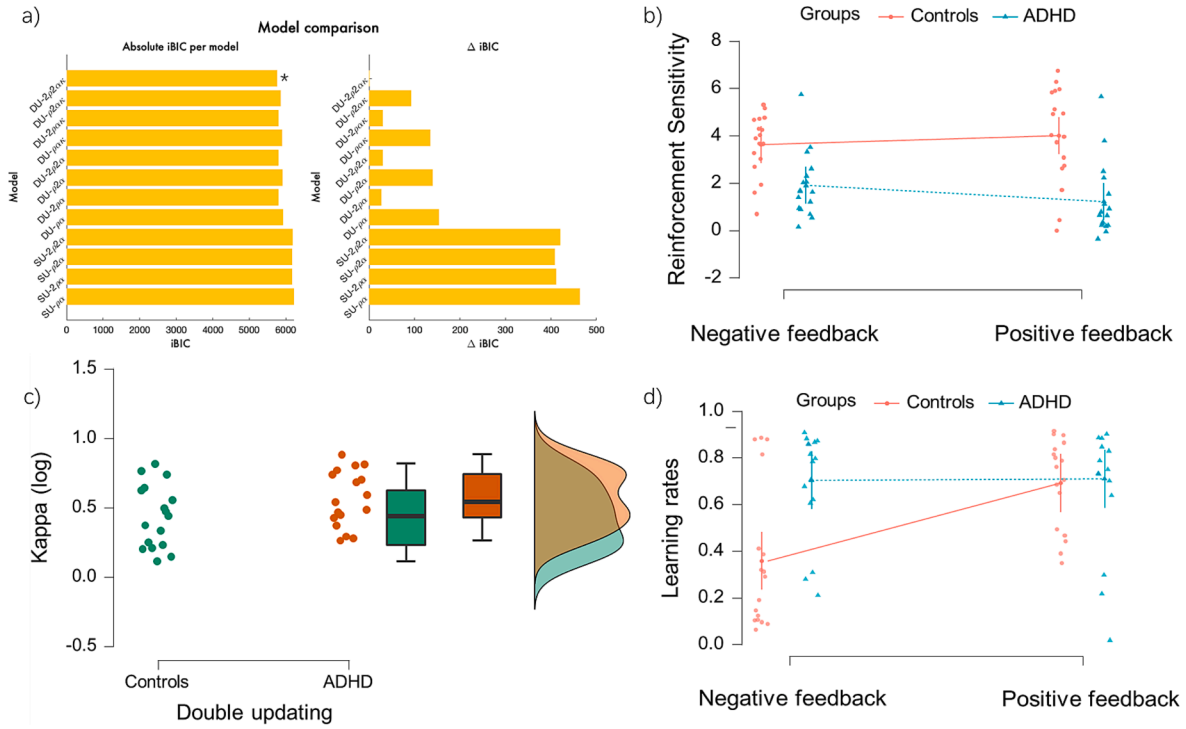

Fig. 3. **A)** Model comparison showing the integrated Bayesian information criterion (iBIC) and delta iBIC (distance from the model with the best evidence) of the respective models. The most complex Q-learning model with separate learning rates and reinforcement sensitivities for wins and losses and a weighting parameter between single and double updating (marked with an asterisk) showed the best fit to the behavioral data. **b)** ADHD patients showed an overall lower reinforcement sensitivity. Reinforcement sensitivity for positive feedback was especially lower in ADHD patients. **c)** ADHD patients and healthy controls differed marginally in weighting the value of the chosen and non-chosen option to update, with ADHD patients updating the unchosen option slightly weaker. **d)** ADHD patients showed a higher learning rate than controls. This effect was driven by the higher learning rate for negative feedback in ADHD patients.

10/−15, t = 2.55, p_{FWE} = 0.045, k: 46 voxel, see Fig. 4a), but not in the ventromedial prefrontal cortex (xyz: 2/33/−2, t = 2.30, p_{FWE} = 0.520). There were no significant differences between groups for the ROIs (VS L: xyz: −18/6/−15, t = 0.84, p_{FWE} = 0.423; VS R: xyz: 14/3/−15, t = 1.23, p_{FWE} = 0.342; VMPFC: VS L: xyz: −10/50/0, t = 2.74, p_{FWE} = 0.277) nor on the whole brain level.

Across both groups, double update prediction errors were significant in the ventromedial prefrontal cortex cortex (xyz: 0/53/8, t = 3.62, p_{FWE} = 0.043, k: 32 voxel) and only marginally significant in left (xyz: −13/10/−15, t = 2.31, p_{FWE} = 0.060, k: 31 voxel) and right ventral striatum (xyz: 7/6/−8, t = 2.37, p_{FWE} = 0.066, k: 46 voxel). We found a marginally higher double update prediction error signal in the right ventral striatum of healthy controls (type of PE x group, xyz: 7/6/−8, t = 2.12, p_{FWE} = 0.097, k: 46 voxel, see Fig. 4b) but not in the vmPFC (xyz: 0/6/−8, t = 2.95, p_{FWE} = 0.191). No other group effects emerged on the whole brain level.

Across both groups, trial-by-trial choice probability was only marginally significantly related to activity in the vmPFC (xyz: 7/50/−8, t = 3.66, p_{FWE} = 0.083, k: 3 voxel). There was no group difference in the vmPFC (xyz: 4/16/−8, t = 2.16, p_{FWE} = 0.690). However, at the whole brain level, there was a significantly weaker neural representation of choice probability in ADHD as compared to controls in the posterior

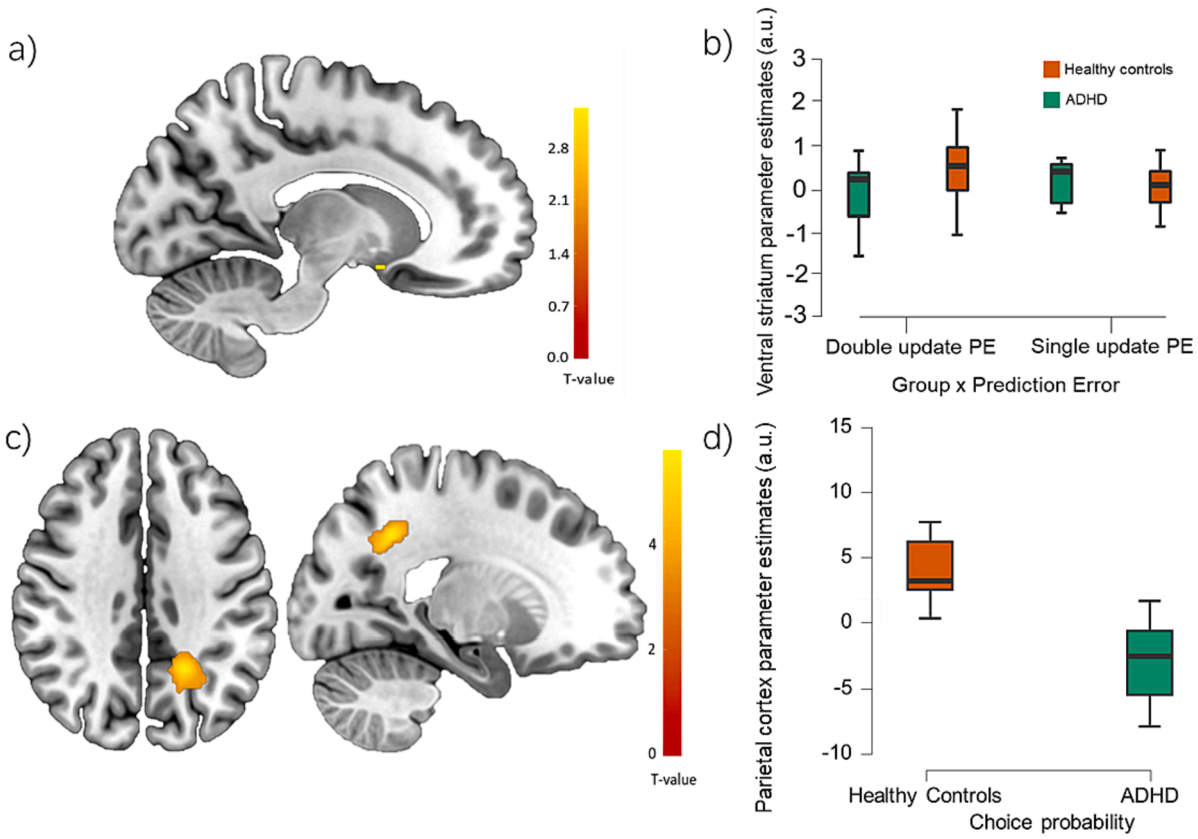

Fig. 4. A) Single update prediction errors of both groups were represented in the left nucleus accumbens (xyz: −13, 8, −15). **b)** There was a marginally significant prediction error x group interaction. ADHD patients that had a weaker double update prediction error (DU PEs) signal in the right ventral striatum drove this effect. PE: Prediction Error, a. u.: Arbitrary units. **c)** Choice probability representation in the left posterior parietal cortex was weaker in ADHD patients. Slices MNI coordinates in 4c: −18, 5, 38. The color bars represent the t-values. The images are radiologically oriented.

parietal cortex (xyz: −20/−54/38, t = 5.82, pFWE = 0.04, k: 167 voxel, see Fig. 4c and d).

Across both groups, we found neural representations of noisy/exploratory trials in the left and right insular cortex in both groups at the whole brain level (left: xyz: −38/16/−12, t = 7.52, pFWE = 0.001, k: 527 voxel; right: xyz: 40/28/−8, t = 6.08, pFWE = 0.024, k: 622 voxel, see Supplementary Fig. S6). Using a ROI of activation in these trials covering the same regions from an independent study (Waltmann et al., 2023), there was no group difference in these regions (xyz 40/30/−8, t = 3.34, pFWE = 0.734).

4. Discussion

In this study, adult ADHD patients showed impaired performance specifically when the learning environment was stable while performance was slightly improved after a reversal had occurred. Both effects (pre- and post-reversal) can be understood as results of an overall enhanced choice switching, which is maladaptive when the environment is stable but beneficial when environmental changes occur. Our RL modelling explains this choice switching most clearly by a blunted sensitivity to positive and negative reinforcement. This blunted sensitivity results in in less distinguishable values for each of the two choice options. Additionally, an enhanced learning rate after negative feedback as well as a subtle tendency for reduced double-updating also contribute to elevated levels of choice switching. On the neural level, this was mirrored by a weaker representation of choice probability (which is scaled by reinforcement sensitivity) in the parietal cortex and weak indications for reduced double-update PEs in the right ventral striatum of ADHD patients. These results should be treated with caution, in particular with regard to double updating, due to the limited sample size of the current study.

A similar study in adolescent ADHD patients and healthy controls with a probabilistic reversal learning task in fMRI showed only partially overlapping results (Hauser et al., 2014). This is probably partly due to different analysis methods such as the underlying models or the inclusion of pre- and post-reversal phases. The study also found no group difference in terms of overall accuracy, but did not test for possible phase effects, which we found to be significant. While modeling implementation was slightly different (in our modeling, we used a fixed softmax function and variable reinforcement sensitivities instead of variable temperatures of the softmax function) we find comparative results indicating enhanced decision noise leading to increased exploratory behavior.

As described above, individuals with ADHD had a significantly weaker neural representation of choice probability in the parietal cortex, compared to the control group. The parietal cortex is a crucial part of the attention network (Rushworth et al., 2001, Ptak, 2012). A weaker attentional system might disrupt the processing of reinforcement information, making it difficult for an individual to accurately perceive and control the positive and negative consequences of their actions. This in turn might result in a reduced sensitivity to reinforcement, as seen in our modeling data, suggesting that ADHD patients probably need stronger reinforcements to update their choice values and to maintain certainty in decision making. A lower reinforcement sensitivity and a weaker processing of choice probability could lead to noisy/exploratory choice switching behavior independent of prior feedback, which was clearly evident in our behavioral data.

We did not replicate the weaker prediction error signals found in ADHD patients in the ventromedial prefrontal cortex in the previous study (Hauser et al., 2014). Instead, we found a trend of weaker learning

signals of the double update prediction error in the nucleus accumbens in ADHD patients. The decreased reinforcement sensitivity could make it more difficult for ADHD patients to build an internal model of contingencies. Therefore, they are more likely to respond to acute changes, which is beneficial post-reversal but detrimental pre-reversal. This is in line with solid evidence that ADHD patients prefer smaller immediate rewards for easier tasks as opposed to larger delayed rewards for more difficult tasks (Tripp and Alsop, 2001, De Meyer et al., 2019). For ADHD patients, this could also lead to poorer retrieval of internal choice values, resulting in more variability in reaction times (Kofler et al., 2013, Véronneau-Veilleux et al., 2022). We speculate that this decreased reinforcement sensitivity could be linked to our observation of weaker learning signals of the ventral striatum that incorporate chosen and unchosen action values. This finding could result from a reduced integration of environmental information (external sensory information or internal states) to learning signals due to aberrant dopaminergic neuromodulation. Research with animal subjects has already shown that dopaminergic neurons have a modulatory effect on neuronal and circuit flexibility, which ultimately leads to changes in behavior (Siju et al., 2021). With further necessary empirical evidence, this could be regarded as an extension of existing ADHD dopamine theories (Tripp and Wickens, 2008, Ziegler et al., 2016). In this regard, computational modelling is a helpful tool to further elucidate dopamine-based learning mechanisms in ADHD.

The literature regarding the learning rate of ADHD patients is not yet congruent. One theory proposes that the performance difficulties of ADHD patients in reward learning tasks may not be associated with deficits in learning, but with the sensitivity to reinforcements and the storing of cue-outcome contingences (Luman et al., 2009). This is supported by studies with tasks that explicitly test learning from losses and wins (Agay et al., 2010). However, in our data and model analyses, ADHD patients showed an increased learning rate for negative feedback, whereas the learning rate for positive feedback did not differ between groups. A higher learning rate for negative feedback would ensure that choice values are extinguished more quickly after a reversal and thus support enhanced switching (Ziegler et al., 2016) as seen in our data (but not in feedback-specific manner). Valence-dependent learning deficits have also been observed in a disease in which decreased cerebral dopamine concentrations play a role: Parkinson's disease (PD). In one study, unmedicated PD patients learned less well from positive feedback compared to healthy controls, but this effect was reversed for negative feedback (Frank et al., 2004). The authors attribute this to the different direct (D1 receptor) and indirect (D2 receptor) pathways of the basal ganglia: While reduced phasic dopamine bursts would decrease sensitivity to positive feedback (via D1 receptors), reduced tonic dopamine could provide increased D2 receptor activity supporting learning from negative feedback. According to this theory, influences of the tonic dopamine concentration and thus on the indirect pathway would influence the learning behavior of ADHD patients. Lower tonic dopamine levels and thus higher D2 receptor activity could thus enhance learning from negative feedback and in addition to augmented decision noise lead to increased switching behavior. While this would explain our data, our study cannot prove this theory. Further animal studies, for example by influencing tonic and phasic dopamine bursts by genetic manipulation (Beeler et al., 2010) are necessary to draw clearer conclusions.

Our observation that the anterior insula plays a role in exploratory decisions aligns with findings in the general population reported by other studies (Reiter et al., 2017; Li et al., 2021; Zhen et al., 2022; Waltmann et al., 2023). Furthermore, in adolescents, the distinction between 'explorers' and 'non-explorers' during a temporal decision-making task is marked by greater resting-state connectivity between the rostrolateral PFC and the insula in the 'explorers' (Kayser et al., 2016). Notably, the administration of L-dopa seems to mitigate decision uncertainty associated with the anterior insula (Chakroun et al., 2020). This suggests a potential mechanism where the insula, as part of the salience network, could drive exploration under conditions of heightened overall uncertainty. This may involve facilitating a switch from the presently exploited option to more uncertain yet discernible alternative choices (Chakroun et al., 2020; Li et al., 2021).

It is interesting to speculate on daily life implications of our findings: a generally lower reinforcement sensitivity means that negative and positive events have a smaller influence on choice values and, thus, subsequent actions. This could result in a situation where the internally assigned values of these actions do not significantly differ when deciding between action options, which leads to greater uncertainty about which decision to take next. In real-life situations where certain conditions provide stability and the consequences of actions remain relatively constant, ADHD patients might explore different action options more due to this higher level of uncertainty, and are less likely to exploit the beneficial options. For example, this observation aligns with behaviour often seen in children with ADHD, who tend to frequently switch between play activities. Such behavior presents a potential disadvantage, as it contrasts with the inclination of other same-age children to engage in a single activity for a more extended period. In adolescents and adults with ADHD this may result in switching conversation topics quickly in a manner that annoys members of a peer group. This behavior may stem from the difficulty in discerning preferences. However, in situations where the consequences of actions are unpredictable, quick shifts in preferred actions could be advantageous.

4.1. Limitations

The sample size of this study, n = 17 per group is too small to draw more than preliminary conclusions. Caution is necessary also because small sample sizes can inflate effect sizes (Button et al., 2013). While we matched the two groups for age, handedness, and gender, group differences emerged with respect to nicotine use, intelligence, and working memory. While it is known that ADHD patients are more likely to smoke (Ilbegi et al., 2018) and perform worse on tests of intelligence (Bridgett and Walker, 2006) and working memory (Kofler et al., 2020), these differences between groups may explain some of the behavioral differences. However, we believe that our binary choice task places relatively low demands on working memory. Nevertheless, one should be cautious in interpreting the results as generalizing to all ADHD patients. Finally, the question arises to what extent the task structure can really represent exploratory behavior. While there is some uncertainty in the currently used task, the strictly anti-correlated structure scarcely represents the real complex exploration behavior of ADHD patients in their everyday life. The task structure with two response options has another limitation: It makes it difficult to pinpoint why ADHD patients perform better in the post-reversal phase of the task. One possibility is that ADHD patients artificially benefit from this simple task structure. The frequent choice switching observed in ADHD patients could align with the task's inherent environmental changes due to frequent reversals. Alternatively, ADHD patients might be able to adjust their internal parameters more quickly after the reversal. This could be due to their higher learning rate from negative feedback and lower sensitivity to reinforcement, allowing them to purposely choose a new option more quickly. To clarify this, future studies could use a task design with three response options in combination with modelling equipped to tackle state space learning. However, a more complex task structure could also be associated with drawbacks (e.g., greater dependence on the individual working memory of the test subjects). In terms of model fitting, it is relevant to note that the same (empirical) priors were used in fitting the model to both groups. Thus, we adopted a conservative modelling approach in which we assume that parameters of each group are drawn from the same distribution. This introduces a conservative bias (increasing the risk of type 2 error). However, modelling the data separately for each group has the opposite effect – introducing an anti-conservative bias – which increases the risk of overestimating group differences (type 1 errors).

5. Conclusion

Using computational reinforcement learning models, this study provides insight into the neurocognitive processes that facilitate behavioral differences in motivational learning and decision making in ADHD patients. Noisier behavior of ADHD patients was associated with decreased reinforcement sensitivity in our study. This behavior could be due to a reduced neural representation of dopaminergic prediction error signals in the nucleus accumbens and a reduced representation of choice probability in the posterior parietal cortex in ADHD patients. We speculate that lower tonic dopamine levels might lead to faster relearning after negative feedback via D2 receptor activation in ADHD, which may prove beneficial in rapidly changing environments.

CRediT authorship contribution statement

Hans-Christoph Aster: Formal analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. **Maria Waltmann:** Formal analysis, Methodology, Software, Validation, Writing – review & editing. **Anika Busch:** Formal analysis, Methodology, Writing – review & editing. **Marcel Romanos:** Conceptualization, Funding acquisition, Project administration, Writing – review & editing. **Matthias Gamer:** Formal analysis, Methodology, Supervision, Validation, Writing – review & editing. **Betteke Maria van Noort:** Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing. **Anne Beck:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. **Viola Kappel:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. **Lorenz Deserno:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Funding

H-CA was supported by a Clinician Scientist Program at the Interdisciplinary Centre of Clinical Research at the Medical Faculty of the University of Würzburg. LD was supported by the IFB Adiposity Diseases, Federal Ministry of Education and Research (BMBF), Germany, GN: 01EO150, and a grant on reinforcement learning in ADHD by the German Research Foundation (DFG, 533682086). LD and AB are supported by the DFG as part of the Collaborative Research Centre 265 Losing and Regaining Control over drug intake (402170461, Project A02 and C02).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2024.103588.

References

Agay, N., Yechiam, E., Carmel, Z., Levkovitz, Y., 2010. Non-specific effects of methylphenidate (ritalin) on cognitive ability and decision-making of ADHD and healthy adults. Psychopharmacology 210 (4), 511–519.
Ashburner, J., Friston, K.J., 2005. Unified segmentation. Neuroimage 26 (3), 839–851.
Barr, D.J., Levy, R., Scheepers, C., Tily, H.J., 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. J Mem Lang 68 (3).
Beeler, J.A., Daw, N., Frazier, C.R., Zhuang, X., 2010. Tonic dopamine modulates exploitation of reward learning. Front Behav Neurosci 4, 170.
Bridgett, D.J., Walker, M.E., 2006. Intellectual functioning in adults with ADHD: a meta-analytic examination of full scale IQ differences between adults with and without ADHD. Psychol Assess 18 (1), 1–14.
Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nature Reviews Neuroscience 14 (5), 365–376.
Calabro, F.J., Montez, D.F., Larsen, B., Laymon, C.M., Foran, W., Hallquist, M.N., Price, J. C., Luna, B., 2023. Striatal dopamine supports reward expectation and learning: a simultaneous PET/FMRI study. Neuroimage 267, 119831.
Cazé, R.D., van der Meer, M.A.A., 2013. Adaptive properties of differential learning rates for positive and negative outcomes. Biological Cybernetics 107 (6), 711–719.
Chakroun, K., Mathar, D., Wiehler, A., Ganzer, F., Peters, J., 2020. Dopaminergic Modulation of the Exploration/exploitation Trade-off in Human Decision-Making. eLife 9, e51260.
Chantiluke, K., Barrett, N., Giampietro, V., Brammer, M., Simmons, A., Murphy, D.G., Rubia, K., 2015. Inverse effect of fluoxetine on medial prefrontal cortex activation during reward reversal in ADHD and autism. Cereb Cortex 25 (7), 1757–1770.
Christiansen, H., Hirsch, O., Philipsen, A., Oades, R.D., Matthies, S., Hebebrand, J., Ueckermann, J., Abdel-Hamid, M., Kraemer, M., Wiltfang, J., Graf, E., Colla, M., Sobanski, E., Alm, B., Rösler, M., Jacob, C., Jans, T., Huss, M., Schimmelmann, B.G., Kis, B., 2013. German validation of the conners adult ADHD rating scale-self-report: confirmation of factor structure in a large sample of participants with ADHD. J Atten Disord 17 (8), 690–698.
Coren, S., 1993. Measurement of handedness via self-report: the relationship between brief and extended inventories. Perceptual and Motor Skills 76 (3), 1035–1042.
D'Ardenne, K., McClure, S.M., Nystrom, L.E., Cohen, J.D., 2008. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. Science 319 (5867), 1264–1267.
Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., Dolan, R.J., 2011. Model-based influences on humans' choices and striatal prediction errors. Neuron 69 (6), 1204–1215.
De Meyer, H., Beckers, T., Tripp, G., van der Oord, S., 2019. Reinforcement contingency learning in children with ADHD: Back to the basics of behavior therapy. J Abnorm Child Psychol 47 (12), 1889–1902.
Derogatis, L. R. and K. L. Savitz (1999). The SCL-90-R, Brief Symptom Inventory, and Matching Clinical Rating Scales. The use of psychological testing for treatment planning and outcomes assessment, 2nd ed. Mahwah, NJ, US, Lawrence Erlbaum Associates Publishers: 679-724.
Deserno, L., Beck, A., Huys, Q.J., Lorenz, R.C., Buchert, R., Buchholz, H.G., Plotkin, M., Kumakara, Y., Cumming, P., Heinze, H.J., Grace, A.A., Rapp, M.A., Schlagenhauf, F., Heinz, A., 2015a. Chronic alcohol intake abolishes the relationship between dopamine synthesis capacity and learning signals in the ventral striatum. Eur J Neurosci 41 (4), 477–486.
Deserno, L., Huys, Q.J., Boehme, R., Buchert, R., Heinze, H.J., Grace, A.A., Dolan, R.J., Heinz, A., Schlagenhauf, F., 2015b. Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. Proc Natl Acad Sci U S A 112 (5), 1595–1600.
Deserno, L., Moran, R., Michely, J., Lee, Y., Dayan, P., Dolan, R.J., 2021. Dopamine enhances model-free credit assignment through boosting of retrospective model-based inference. Neuroscience.
Dreher, J.C., Kohn, P., Kolachana, B., Weinberger, D.R., Berman, K.F., 2009. Variation in dopamine genes influences responsivity of the human reward system. Proc Natl Acad Sci U S A 106 (2), 617–622.
Eppinger, B., Kray, J., 2011. To choose or to avoid: age differences in learning from positive and negative feedback. Journal of Cognitive Neuroscience 23 (1), 41–52.
Faraone, S.V., Asherson, P., Banaschewski, T., Biederman, J., Buitelaar, J.K., Ramos-Quiroga, J.A., Rohde, L.A., Sonuga-Barke, E.J., Tannock, R., Franke, B., 2015. Attention-deficit/hyperactivity disorder. Nat Rev Dis Primers 1, 15020.
Frank, M.J., Seeberger, L.C., O'Reilly, C.R., 2004. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306 (5703), 1940–1943.
Frank, M.J., Moustafa, A.A., Haughey, H.M., Curran, T., Hutchison, K.E., 2007. Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc Natl Acad Sci U S A 104 (41), 16311–16316.
Fusar-Poli, P., Rubia, K., Rossi, G., Sartori, G., Balottin, U., 2012. Striatal dopamine transporter alterations in ADHD: pathophysiology or adaptation to psychostimulants? a meta-analysis. Am J Psychiatry 169 (3), 264–272.
Geisler, D., Ritschel, F., King, J.A., Bernardoni, F., Seidel, M., Boehm, I., Runge, F., Goschke, T., Roessner, V., Smolka, M.N., Ehrlich, S., 2017. Increased anterior cingulate cortex response precedes behavioural adaptation in anorexia nervosa. Sci Rep 7, 42066.
Hauser, T.U., Iannaccone, R., Ball, J., Mathys, C., Brandeis, D., Walitza, S., Brem, S., 2014. Role of the medial prefrontal cortex in impaired decision making in juvenile attention-deficit/hyperactivity disorder. JAMA Psychiatry 71 (10), 1165–1173.
Hollerman, J.R., Schultz, W., 1998. Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci 1 (4), 304–309.

Humphreys, K.L., Tottenham, N., Lee, S.S., 2018. Risky decision-making in children with and without ADHD: a prospective study. Child Neuropsychol 24 (2), 261–276.
Huys, Q.J., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., Roiser, J.P., 2012. Bonsai trees in your head: how the pavlovian system sculpts goal-directed choices by pruning decision trees. PLoS Comput Biol 8 (3), e1002410.
Huys, Q.J., Pizzagalli, D.A., Bogdan, R., Dayan, P., 2013. Mapping anhedonia onto reinforcement learning: a behavioural meta-analysis. Biol Mood Anxiety Disord 3 (1), 12.
Huys, Q. J. M. (2017). Bayesian Approaches to Learning and Decision Making. Computational Psychiatry - Mathematical Modeling of Mental Illness. J. M. Alan Anticevic, Elsevier Wordmark.
Ilbegi, S., Groenman, A.P., Schellekens, A., Hartman, C.A., Hoekstra, P.J., Franke, B., Faraone, S.V., Rommelse, N.N.J., Buitelaar, J.K., 2018. Substance use and nicotine dependence in persistent, remittent, and late-onset ADHD: a 10-year longitudinal study from childhood to young adulthood. Journal of Neurodevelopmental Disorders 10 (1), 42.
Katahira, K., 2015. The relation between reinforcement learning parameters and the influence of reinforcement history on choice behavior. Journal of Mathematical Psychology 66, 59–69.
Kayser, A.S., Op de Macks, Z., Dahl, R.E., Frank, M.J., 2016. A neural correlate of strategic exploration at the onset of adolescence. Journal of Cognitive Neuroscience 28 (2), 199–209.
Kofler, M.J., Rapport, M.D., Sarver, D.E., Raiker, J.S., Orban, S.A., Friedman, L.M., Kolomeyer, E.G., 2013. Reaction time variability in ADHD: a meta-analytic review of 319 studies. Clin Psychol Rev 33 (6), 795–811.
Kofler, M.J., Singh, L.J., Soto, E.F., Chan, E.S.M., Miller, C.E., Harmon, S.L., Spiegel, J.A., 2020. Working memory and short-term memory deficits in ADHD: a bifactor modeling approach. Neuropsychology 34 (6), 686–698.
Li, C.-W., Lin, C.-Y.-Y., Chang, T.-T., Yen, N.-S., Tan, D., 2021. Motivational system modulates brain responses during exploratory decision-making. Scientific Reports 11 (1), 15810.
Luman, M., Van Meel, C.S., Oosterlaan, J., Sergeant, J.A., Geurts, H.M., 2009. Does reward frequency or magnitude drive reinforcement-learning in attention-deficit/hyperactivity disorder? Psychiatry Res 168 (3), 222–229.
Marx, I., Hacker, T., Yu, X., Cortese, S., Sonuga-Barke, E., 2021. ADHD and the choice of small immediate over larger delayed rewards: a comparative meta-analysis of performance on simple choice-delay and temporal discounting paradigms. Journal of Attention Disorders 25 (2), 171–187.
Mazaika, P., Whitfield, S., Cooper, J.C., 2005. Detection and repair of transient artifacts in fMRI data. Neuroimage 26, S36.
Mowinckel, A.M., Pedersen, M.L., Eilertsen, E., Biele, G., 2015. A meta-analysis of decision-making and attention in adults with ADHD. Journal of Attention Disorders 19 (5), 355–367.
Nussenbaum, K., Hartley, C.A., 2019. Reinforcement learning across development: what insights can we draw from a decade of research? Dev Cogn Neurosci 40, 100733.
O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., Dolan, R.J., 2004. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304 (5669), 452–454.
Pessiglione, M., Seymour, B., Flandin, G., Dolan, R.J., Frith, C.D., 2006. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442 (7106), 1042–1045.
Pine, A., Sadeh, N., Ben-Yakov, A., Dudai, Y., Mendelsohn, A., 2018. Knowledge acquisition is governed by striatal prediction errors. Nature Communications 9 (1), 1673.
Plichta, M.M., Scheres, A., 2014. Ventral–striatal responsiveness during reward anticipation in ADHD and its relation to trait impulsivity in the healthy population: a meta-analytic review of the fMRI literature. Neurosci Biobehav Rev 38, 125–134.
Ptak, R., 2012. The frontoparietal attention network of the human brain: action, saliency, and a priority map of the environment. Neuroscientist 18 (5), 502–515.
Reitan, R.M., 1958. Validity of the trail making test as an indicator of organic brain damage. Perceptual and Motor Skills 8 (3), 271–276.
Reiter, A.M.F., Deserno, L., Kallert, T., Heinze, H.-J., Heinz, A., Schlagenhauf, F., 2016. Behavioral and neural signatures of reduced updating of alternative options in alcohol-dependent patients during flexible decision-making. The Journal of Neuroscience 36 (43), 10935–10948.
Reiter, A.M., Heinze, H.J., Schlagenhauf, F., Deserno, L., 2017. Impaired flexible reward-based decision-making in binge eating disorder: evidence from computational modeling and functional neuroimaging. Neuropsychopharmacology 42 (3), 628–637.
Retz-Junginger, P., Retz, W., Blocher, D., Weijers, H.G., Trott, G.E., Wender, P.H., Rössler, M., 2002. Wender Utah rating scale. the short-version for the assessment of the attention-deficit hyperactivity disorder in adults. Nervenarzt 73 (9), 830–838.
Rösler, M., Retz, W., Retz-Junginger, P., Thome, J., Supprian, T., Nissen, T., Stieglitz, R. D., Blocher, D., Hengesch, G., Trott, G.E., 2005. Instrumente zur diagnostik der aufmerksamkeitsdefizit-/hyperaktivitätsstörung (ADHS) im erwachsenenalter. Der Nervenarzt 76 (1), 129–130.
Rostami Kandroodi, M., Cook, J.L., Swart, J.C., Froböse, M.I., Geurts, D.E.M., Vahabie, A. H., Nili Ahmadabadi, M., Cools, R., den Ouden, H.E.M., 2021. Effects of methylphenidate on reinforcement learning depend on working memory capacity. Psychopharmacology (berl) 238 (12), 3569–3584.
Rushworth, M.F., Krams, M., Passingham, R.E., 2001. The attentional role of the left parietal cortex: the distinct lateralization and localization of motor attention in the human brain. J Cogn Neurosci 13 (5), 698–710.
Schlagenhauf, F., Rapp, M.A., Huys, Q.J., Beck, A., Wustenberg, T., Deserno, L., Buchholz, H.G., Kalbitzer, J., Buchert, R., Bauer, M., Kienast, T., Cumming, P., Plotkin, M., Kumakura, Y., Grace, A.A., Dolan, R.J., Heinz, A., 2013. Ventral striatal prediction error signaling is associated with dopamine synthesis capacity and fluid intelligence. Hum Brain Mapp 34 (6), 1490–1499.
Schlagenhauf, F., Huys, Q.J., Deserno, L., Rapp, M.A., Beck, A., Heinze, H.J., Dolan, R., Heinz, A., 2014. Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. Neuroimage 89, 171–180.
Scholz, V., Waltmann, M., Herzog, N., Reiter, A., Horstmann, A., Deserno, L., 2023. Cortical Grey matter mediates increases in model-based control and learning from positive feedback from adolescence to adulthood. The Journal of Neuroscience 43 (12), 2178–2189.
Schultz, W., 2013. Updating dopamine reward signals. Curr Opin Neurobiol 23 (2), 229–238.
Siju, K.P., De Backer, J.F., Grunwald Kadow, I.C., 2021. Dopamine modulation of sensory processing and adaptive behavior in flies. Cell Tissue Res 383 (1), 207–225.
Sutton, R.S., Barto, A.G., 1998. Reinforcement learning: an introduction. MA, MIT Press, Cambridge.
Tripp, G., Alsop, B., 2001. Sensitivity to reward delay in children with attention deficit hyperactivity disorder (ADHD). Journal of Child Psychology and Psychiatry 42 (5), 691–698.
Tripp, G., Wickens, J.R., 2008. Research review: dopamine transfer deficit: a neurobiological theory of altered reinforcement mechanisms in ADHD. Journal of Child Psychology and Psychiatry 49 (7), 691–704.
Véronneau-Veilleux, F., P. Robaey, M. Ursino and F. Nekka (2022). "A mechanistic model of ADHD as resulting from dopamine phasic/tonic imbalance during reinforcement learning." Frontiers in Computational Neuroscience **16**.
Von Aster, M., A. Neubauer and R. Horn (2006). Hamburg-Wechsler-Intelligenz-Test für Erwachsene III, Harcourt, Frankfurt.
von Rhein, D., Cools, R., Zwiers, M.P., van der Schaaf, M., Franke, B., Luman, M., Oosterlaan, J., Heslenfeld, D.J., Hoekstra, P.J., Hartman, C.A., Faraone, S.V., van Rooij, D., van Dongen, E.V., Lojowska, M., Mennes, M., Buitelaar, J., 2015. Increased neural responses to reward in adolescents and young adults with attention-deficit/hyperactivity disorder and their unaffected siblings. Journal of the American Academy of Child & Adolescent Psychiatry 54 (5), 394–402.
Waltmann, M., Schlagenhauf, F., Deserno, L., 2022. Sufficient reliability of the behavioral and computational readouts of a probabilistic reversal learning task. Behav Res Methods.
Waltmann, M., Herzog, N., Reiter, A.M.F., Villringer, A., Horstmann, A., Deserno, L., 2023. 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 60, 101226.
Watkins, C.J.C.H., Dayan, P., 1992. Q-learning. Machine Learning 8 (3), 279–292.
Weiss, E. O., J. A. Kruppa, G. R. Fink, B. Herpertz-Dahlmann, K. Konrad and M. Schulte-Rüther (2021). "Developmental Differences in Probabilistic Reversal Learning: A Computational Modeling Approach." Frontiers in Neuroscience **14**.
Weiß, R.H., Weiß, B., 2008. CFT 20-R mit WS/ZF-R: grundintelligenztest skala 2-revision (CFT 20-R) mit wortschatztest und zahlenfolgentest-revision (WS/ZF-R). Hogrefe.
Westbrook, A., van den Bosch, R., Maatta, J.I., Hofmans, L., Papadopetraki, D., Cools, R., Frank, M.J., 2020. Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work. Science 367 (6484), 1362–1366.
Wittchen, H.-U. and H. Pfister (1997). "DIA-X-interviews: manual für screening-Verfahren und interview; Interviewheft.".
Zhen, S., Yaple, Z.A., Eickhoff, S.B., Yu, R., 2022. To learn or to gain: neural signatures of exploration in human decision-making. Brain Structure and Function 227 (1), 63–76.
Ziegler, S., Pedersen, M.L., Mowinckel, A.M., Biele, G., 2016. Modelling ADHD: a review of ADHD theories through their predictions for computational models of decision-making and reinforcement learning. Neurosci Biobehav Rev 71, 633–656.