

## ARTICLE OPEN



# Exploring the steps of learning: computational modeling of initiatory-actions among individuals with attention-deficit/hyperactivity disorder

Gili Katabi <sup>1</sup>✉ and Nitzan Shahar <sup>1,2</sup>

© The Author(s) 2024

Attention-deficit/hyperactivity disorder (ADHD) is characterized by difficulty in acting in a goal-directed manner. While most environments require a sequence of actions for goal attainment, ADHD was never studied in the context of value-based sequence learning. Here, we made use of current advancements in hierarchical reinforcement-learning algorithms to track the internal value and choice policy of individuals with ADHD performing a three-stage sequence learning task. Specifically, 54 participants (28 ADHD, 26 controls) completed a value-based reinforcement-learning task that allowed us to estimate internal action values for each trial and stage using computational modeling. We found attenuated sensitivity to action values in ADHD compared to controls, both in choice and reaction-time variability estimates. Remarkably, this was found only for first-stage actions (i.e., initiatory actions), while for actions performed just before outcome delivery the two groups were strikingly indistinguishable. These results suggest a difficulty in following value estimation for initiatory actions in ADHD.

*Translational Psychiatry* (2024)14:10; <https://doi.org/10.1038/s41398-023-02717-7>

## INTRODUCTION

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder that is characterized by inattention, impulsivity, and hyperactivity [1, 2]. Approximately 10% of school-aged children are affected by ADHD [3], with the majority experiencing significant symptoms throughout adolescence and into adulthood [4]. Individuals with ADHD frequently struggle with the completion of tasks that require the execution of multiple actions (e.g., doing homework) [5–7]. Despite the known difficulties experienced by those with ADHD in completing multi-action tasks in natural environments, there has been no prior investigation of ADHD in the context of value-based sequence learning. In this study, we employed reinforcement learning modeling to examine the extent to which ADHD is associated with reduced value-based learning for initial actions (i.e., actions at the start of the sequence) using a multi-stage reinforcement learning task.

Reinforcement learning studies have demonstrated that human individuals estimate the value of potential actions based on their outcome history [8–10]. This estimation process enables individuals to act persistently, even when the overall goal or outcome is distant [11, 12]. Specifically, reinforcement learning algorithms suggest that when making a sequence of actions that lead to a certain outcome, the value of the observed outcome is mentally “backpropagated” to preceding actions [13]. For example, when making a cup of coffee, multiple actions must be completed before achieving the end goal (i.e., consumption of the beverage). These actions should be associated with the value of the reached outcome, thereby increasing the likelihood of repeating/avoiding the sequence based on the value of the outcome. The

backpropagation of values to actions is considered a “glue” that binds the actions into a sequence [14]. However, during the backpropagation process, value is also discounted across the sequence, meaning that initiatory actions (i.e., actions performed at the start of a sequence) receive smaller value updates than terminal actions (i.e., actions performed immediately before outcome delivery) [10, 15]. The challenge is that without making correct value-based choices early in the sequence, the end goal cannot be reached, resulting in less consistent and more arbitrary value-based choices overall. Nevertheless, since initiatory actions are distant in time from the outcome, they are also assigned smaller value updates [10]. Many studies have already suggested that the processes described by reinforcement learning equations do indeed predict human behavior [15–17] and can explain neural activity [9, 18–22]. However, no study that we know of to date has used this approach to investigate whether the value-based process can explain the difficulty of participants with ADHD to engage and maintain a multi-stage task in order to reach a desired goal.

Here, we hypothesize that difficulties acting according to obtained rewards in ADHD will be more pronounced in initiatory actions (i.e., actions that are performed at the start of a sequence) compared with terminal actions (i.e., actions that are performed at the end of a sequence). Specifically, we suggest that the process of updating the value of initiatory actions is less efficient in ADHD, leading to noisier choices at the start of a sequence. Difficulty holding and following distinct action values at the beginning of a sequence can lead to less consistent choices and difficulty achieving the desired goal, which is highly

<sup>1</sup>School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel. <sup>2</sup>Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel. ✉email: [gilikatabi@mail.tau.ac.il](mailto:gilikatabi@mail.tau.ac.il)

Received: 2 July 2023 Revised: 13 December 2023 Accepted: 14 December 2023

Published online: 08 January 2024

relevant to ADHD symptomatology. For example, if the initiatory action of a sequence for making a beverage is not clearly defined, other competing actions that are closer to the outcome and assigned with similar values might take precedence. Alternatively, if the value of the initiatory action is internally well defined, it will lead more frequently to the selection of the best action, generating a more consistent and uninterrupted sequence of actions.

Although no studies have examined action–outcome associations in ADHD for initiatory actions, several findings in the literature support this hypothesis. First, ADHD individuals exhibit marked differences in reward-related neural processing [23–30]. Second, ADHD individuals are known to show steeper temporal discounting of future values [31–33]. Furthermore, ADHD shows inconsistent behavior when performing sustained attention/executive function tasks without immediate reward or feedback [34–40]. Finally, when sustained attention/executive function tasks are performed under conditions of immediate reward (where every action can be considered a terminal action, immediately leading to reward), ADHD shows marked improvements [41]. Therefore, there is reason to believe that individuals with ADHD have a reduced ability to assign and follow internal action values, mostly for actions that are at the beginning of a sequence and distance in time from outcome delivery. However, no empirical investigation has directly addressed this question.

The current study, therefore, examined to what extent individuals diagnosed with ADHD were able to update and act upon action–outcome associations across different stages of a sequential decision-making task. Specifically, we hypothesized that the value-based behavior of ADHD participants would be more inconsistent (noisier choices and higher RT variability) compared with their healthy control (HC) peers, mostly for initiatory but not terminal actions. We further hypothesize that the reason for such group differences will be due to reduced action value updating for outcome-distant stages in ADHD. To examine this question empirically, ADHD/HC participants performed a sequential reinforcement learning task. In each trial, they were asked to make three choices to gain rewards. We then used computational modeling to assess the internal value each individual might be assigning for each action in each stage. We replicated previous findings showing that for all participants and across both groups, choice

accuracy, and RT variability estimates improved when the latent action values were more differentiated (larger internal value difference between the two available actions). We then examined the ability of ADHD individuals to act according to the reward history across stages of the task.

## METHODS

### Participants

Fifty-four undergraduate Tel-Aviv University students completed the study in return for course credits or monetary compensation. (For demographics including gender, age, and IQ, see Table 1; all participants were white caucasian; 53 came from the Jewish sector and 1 from the Arab–Christian sector). All participants reported normal or corrected vision and signed informed consent before participating in the study. Participants using ADHD medications were asked to abstain for at least 24 hours before the lab session. The study was approved by the Tel-Aviv University IRB ethics committee. Exclusion criteria included any neurological or psychiatric history past or current, history of significant head injury, and current drug or alcohol abuse. Specifically, nine participants were excluded from the study due to a concurrent diagnosis of additional disorders, including anxiety, depression, and OCD, while an additional five participants were excluded because they reported ongoing psychiatric treatment and two additional participants were excluded due to drug abuse.

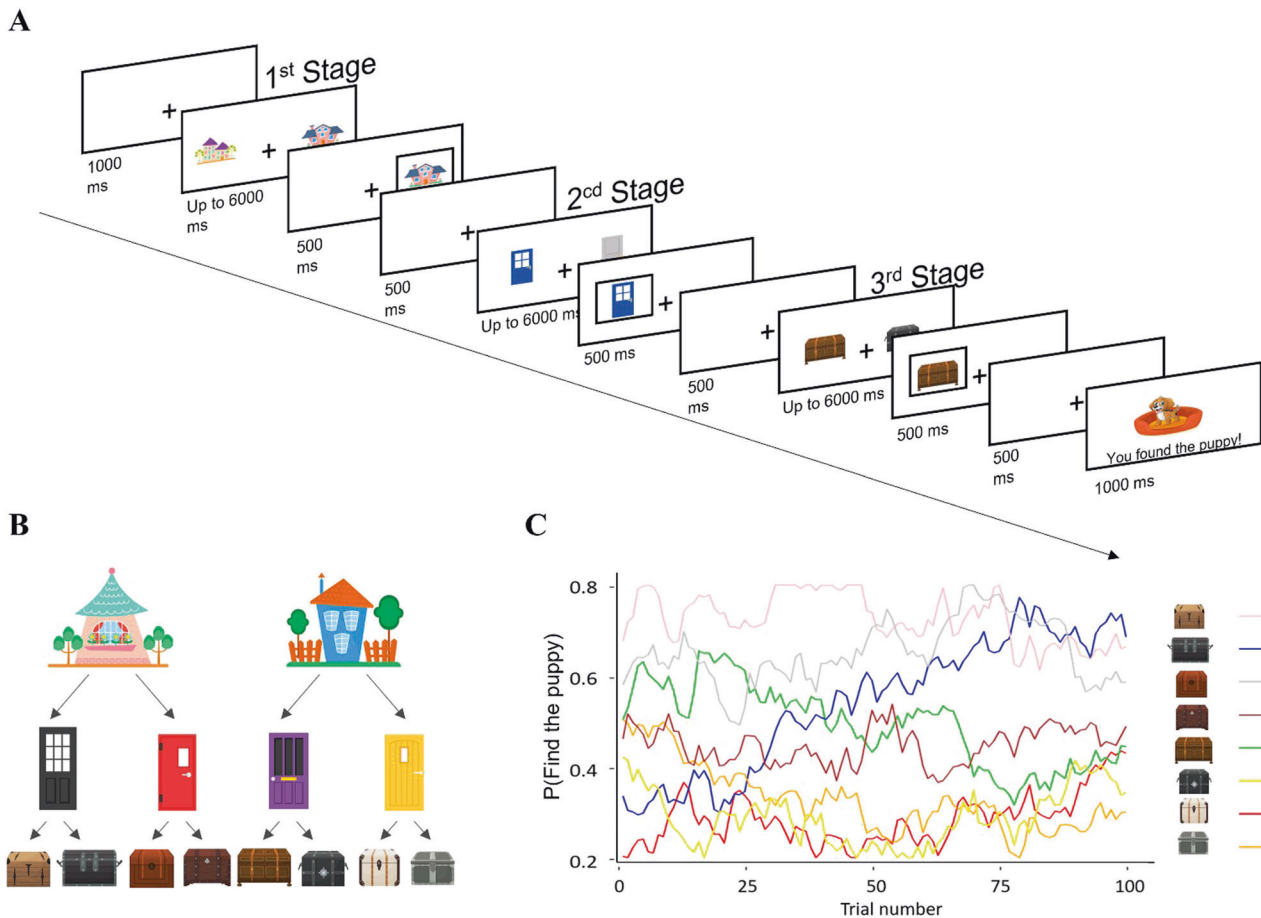
### ADHD diagnosis

To ensure ADHD diagnosis (present for the ADHD group, absent for the control group), we performed the following steps: (1) we made sure that all ADHD participants received an ADHD diagnosis in the community by a licensed mental health professional (approved by the Israeli Government Health Department in making ADHD diagnosis). (2) We made sure that no participants in the control group received an ADHD diagnosis in the community. (3) We questioned the participants in both groups to ensure they were not diagnosed in the community with any additional mental health disorder (current or past). (4) All participants completed a well-validated ADHD assessment in our lab that included a semi-structured interview using the Diagnostic Interview for ADHD in adults (DIVA [42]) to reconfirm the presence/absence of an ADHD diagnosis for the clinical/control group. DIVA assessments were conducted by well-trained psychology students under the direct supervision of a licensed clinical psychologist expert in ADHD. (5) All participants completed self-report questionnaires for ADHD symptoms in adulthood (ASRS [43]) and childhood (WURS [44]) to provide an estimate for the difference of ADHD symptoms between the groups in our cohort (see Table 1 and Fig. S1). (6)

**Table 1.** Characteristics of ADHD and Healthy control groups.

	Healthy control	ADHD	p-Value
N	26 (female: 20)	28 (female: 15)	
Age	19–34 (SD = 4.30)	20–32 (SD = 2.68)	0.38
ASRS score	38.2 (SD = 10.5)	61.7 (SD = 6.6)	<0.001
WURS score	20.0 (SD = 18.5)	44.4 (SD = 13.1)	<0.001
OCI score	40.6 (SD = 28.1)	32.3 (SD = 23.6)	0.25
STAI score	41.6 (SD = 9.6)	45.4 (SD = 9.9)	0.16
BDI score	9.1 (SD = 9.7)	11.3 (SD = 7.2)	0.35
AQ score	15.9 (SD = 5.6)	16.9 (SD = 5.8)	0.44
Raven score	12.2 (SD = 3.4)	12.5 (SD = 3.7)	0.76
AUDIT	2.5 (SD = 2.95)	2.88 (SD = 1.36)	0.62
CUDIT	4.61 (SD = 6.07)	3.23 (SD = 3.32)	0.41
Stimulants Medication	–	10—Methylphenidate based 10—Amphetamine based	

Note. ASRS adult ADHD self-report scale, WURS Wender Utah rating scale for attention deficit hyperactivity, OCI obsessive-compulsive inventory, STAI state-trait anxiety inventory, BDI Beck depression inventory, AQ autism spectrum quotient, Raven Raven progressive matrices, AUDIT alcohol use disorders identification test, CUDIT Cannabis use disorders identification test. Note that CUDIT and AUDIT scores were obtained after the study's completion (as a response to a reviewer's comment) and thus included 18/17 participants from the HC/ADHD group, respectively. All other measurements were completed by all participants included in study.



**Fig. 1 Trial sequences for the multiple-stage task.** Participants were told that a puppy was hiding behind a chest and that their task was to locate the puppy. **A** Each trial individuals made choices across three stages (houses, doors, and chests) to try and locate the puppy. **B** State-action transition structure shows a deterministic transition from stage one (houses), stage two (doors), and stage three (chests). **C** We used a typical reinforcement learning design where the true expected value (probability of finding the puppy for each chest) for each chest drifted slowly across the block, thus requiring participants to keep learning across the task, similar to conventional reinforcement-learning paradigms [10, 94–97].

To further describe our cohort and examine group differences in other relevant estimates, participants completed an IQ test (Ravens [45]), self-report estimates for symptoms of depression (beck depression inventory; BDI [46]), anxiety (state trait anxiety inventory; STAI [47]), obsessive-compulsive (obsessive-compulsive inventory; OCI [48]), autism (autism spectrum quotient; AQ [49]), alcohol use disorder (Alcohol Use Disorders Identification Test, AUDIT [50]) and cannabis use disorder (Cannabis Use Disorders Identification Test, CUDIT [51]). Overall, we found no statistically significant group difference in these scores, while we did find a substantial group difference in ADHD self-report estimates (ASRS and WURS), as should be expected.

### Reinforcement learning task

Participants performed a multiple-stage reinforcement learning task (Fig. 1) that included three sequential stages. The participants were told that their task was to find a puppy that was hiding behind different chests and that they needed to make three choices each trial to try and locate the puppy. They were told that the task included two houses, each house containing two rooms and each room containing two chests. Accordingly, at the first stage of each trial, participants chose one of two “houses” (denoted by cartoon houses stimuli, see Fig. 1B), each leading them deterministically to a second-stage state where a pair of “doors” (cartoon stimuli, see Fig. 1B) was offered. Choosing one of the doors led participants deterministically to a third-stage state where a pair of “chests” (cartoon stimuli; see Fig. 1B) was offered. After making all three choices, the participants were presented with the outcome, which was a cartoon image of a happy puppy or an

empty pet pillow with no puppy (see Fig. 1A). Specifically, each chest delivered a reward (finding the puppy) on a drifting probability (see Fig. 1C), and participants were asked to learn which house, door and chest are most likely to lead to reward throughout the task. (For more information, see SI).

### Choice accuracy and difficulty

For the purposes of analysis, we wanted to code choice accuracy and choice difficulty for each trial and stage. In each stage, participants were asked to make a choice between two actions. A choice was coded as “correct” if the participant selected the action that took part in the sequence with the highest true expected value at a certain stage. Choice difficulty was defined in a similar way, according to the true expected values of the sequence associated with each offered option. (For more information, see SI).

### Using ex-Gaussian distribution to estimate RT variability

An ex-Gaussian distribution is a non-theoretical statistical model that was extensively used in cognitive [52] and ADHD research [53–55] to describe the distribution of reaction times. It combines two different distributions: the Gaussian (normal) distribution, which represents the more frequent and typically quicker reaction time, and an exponential distribution, which represents the tail of slower reaction times. Specifically,  $\tau$  is the parameter that describes the exponential portion of the distribution. By using the ex-Gaussian distribution,

researchers can more accurately model the full range of reaction times, including the long tail of slower responses. (For more information, see SI).

### Data treatment

The first trial in each block, trials with implausibly quick reaction times (<200 ms) or exceptionally slow reaction times (>4000 ms) were omitted (5.15% of all trials). Participants with more than 25% excluded trials (one participant, HC group) or higher than 5% no response rate (none), were further excluded from the analysis. Due to technical issues, we found after data collection that some participants did not receive choice feedback on a minority of trials. Those trials were omitted (69 trials overall), and participants with more than 5% of omitted trials were excluded (4 participants, 3 participants from the HC group, and 1 participant from the ADHD group).

### Computational modeling

To estimate the internal values participants might assign to each available action, we fitted a well-established hierarchical reinforcement learning model to participants' choice behavior [16, 56, 57]. First, we defined a prediction error (PE) for each state and action according to its expected value and the observed rewards as follows (see Eqs. 1–3):

$$PE_1 = V_{\text{choice2}} - V_{\text{choice1}} \quad (1)$$

$$PE_2 = V_{\text{choice3}} - V_{\text{choice2}} \quad (2)$$

$$PE_3 = R - V_{\text{choice3}} \quad (3)$$

Where  $V_{\text{choice1}}$ ,  $V_{\text{choice2}}$ , and  $V_{\text{choice3}}$  are the latent state-action values for the chosen actions in each state (in the reinforcement learning literature, also known as  $Q$  values).  $R$  refers to the outcome observed following the terminal state so that  $R \in \{0,1\}$ . We followed the convention in model-free reinforcement learning modeling [10], where each PE is backpropagated and assigned to all past actions under a discounting factor (i.e., free parameter). Therefore, the first action is updated using  $PE_1$ ,  $PE_2$ , and  $PE_3$ . The second action is updated using  $PE_2$  and  $PE_3$ . The third action is updated using  $PE_3$ . Thus, the value of every action was updated as follows (see Eqs. 4–6):

$$V_{\text{choice1}} = V_{\text{choice1}} + \alpha_1[V_{\text{choice2}} - V_{\text{choice1}}] + \alpha_1\lambda[V_{\text{choice3}} - V_{\text{choice2}}] + \alpha_1\lambda^2[R - V_{\text{choice3}}] \quad (4)$$

$$V_{\text{choice2}} = V_{\text{choice2}} + \alpha_2[V_{\text{choice3}} - V_{\text{choice2}}] + \alpha_2\lambda[R - V_{\text{choice3}}] \quad (5)$$

$$V_{\text{choice3}} = V_{\text{choice3}} + \alpha_3[R - V_{\text{choice3}}] \quad (6)$$

Where  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$  are the learning rates for Stages 1, 2, and 3, respectively  $\lambda$  is an eligibility trace factor referring to the extent of prediction error backpropagation. Specifically, for  $\lambda$  of 1, the agent will learn and update whole sequences, while for  $\lambda = 0$  the agent will learn individual state action values for each stage. Finally, choices were predicted using a softmax decision policy (see Eqs. 7–9):

$$p(\text{choice1}) = \frac{\exp(\beta_1 V_{\text{choice1}})}{\sum \exp(\beta_1 V_i)} \quad (7)$$

$$p(\text{choice2}) = \frac{\exp(\beta_2 V_{\text{choice2}})}{\sum \exp(\beta_2 V_i)} \quad (8)$$

$$p(\text{choice3}) = \frac{\exp(\beta_3 V_{\text{choice3}})}{\sum \exp(\beta_3 V_i)} \quad (9)$$

where  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  are the decision temperature for Stages 1, 2, and 3, respectively. Moreover, to ensure that the model can be used adequately to estimate latent parameters, we performed a parameter recovery analysis, which indicated excellent parameter recovery with no indication of bias (see Fig. S3; for more information, see SI).

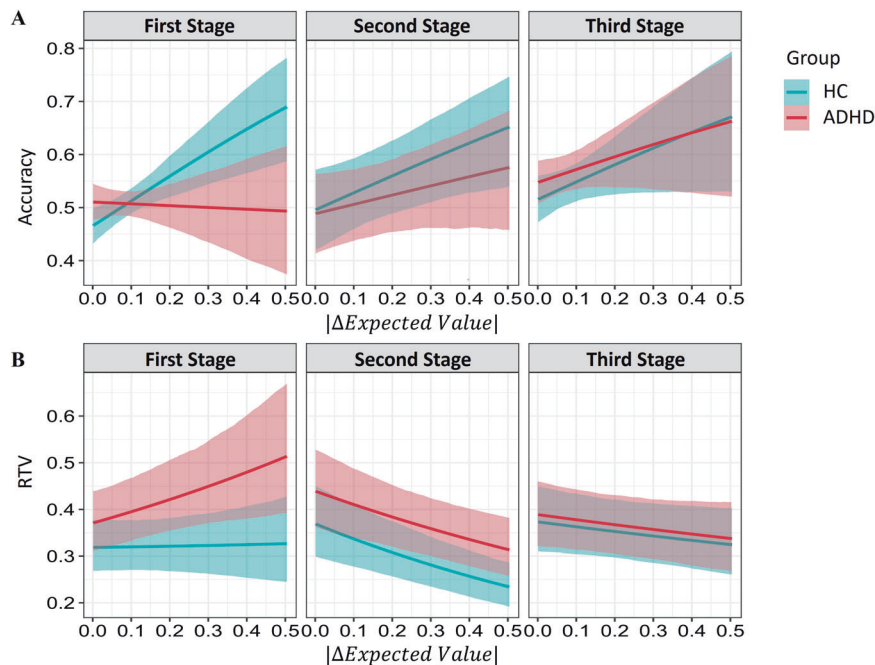
## RESULTS

### Theory-independent analysis

**Accuracy rates—group difference analysis.** To examine the difference in performance between ADHD and HC participants on accuracy rates, we performed a hierarchical Bayesian regression analysis (for more details, see SI) where we predicted choice accuracy (0/1 for choosing the offer with the lower/higher true expected value, respectively), using choice difficulty (absolute difference between the true expected values of the two offers, see Methods), stage (1, 2, or 3) and group (HC vs. ADHD). When examining the interaction between group and choice difficulty, we found that the effect of choice difficulty on accuracy rates was smaller for ADHD compared to HC only for first-stage choices. Specifically, a change of 0.5 in the difference between the expected value of the two offered actions in the first stage was associated with an increase of 22.6% in accuracy rates for the HC group and only 1.8% for the ADHD group [posterior median of the groupXchoice-difficulty effect in the first stage =  $-2.00$   $CI_{89\%}$  between  $-3.35$  and  $-0.71$ ,  $CI_{95\%}$  between  $-3.58$  and  $-0.48$ , probability of direction (pd) = 99.30; see full posterior in Fig. S5A]. Strikingly, group differences vanished completely and even changed direction when examining second and third-stage choices. Specifically, we found a tendency toward a group X choice difficulty interaction in the opposite direction to that of the first stage when examining the second stage [groupXchoice-difficulty effect posterior median = 1.45  $CI_{89\%}$  between  $-0.38$  and 3.22,  $CI_{95\%}$  between  $-0.80$  and 3.69, pd = 89.45%; See full posterior in Fig. S5B] and the third stage [groupXchoice-difficulty median = 1.67,  $CI_{89\%}$  between  $-0.19$  to 3.58,  $CI_{95\%}$  between  $-0.65$  and 4.00, pd = 92.15%; see full posterior in Fig. S5C]. Therefore, we conclude that ADHD participants were not as sensitive to differences in choice difficulty as the HC group when making first-stage choices. However, in the second and third choices, we found no evidence for group differences in accuracy rates (see Fig. 2A). (For additional analysis regarding accuracy rates across groups and the effect of IQ on accuracy rates, see SI).

**Reaction-time variability—group difference analysis.** We next examined the effect of choice difficulty (absolute difference between the true expected values of the two offers, see Methods), group (HC vs. ADHD), and stage (1, 2, or 3) on RT variability using a hierarchical Bayesian regression analysis (for more details see SI). We found that the effect of choice difficulty on RT variability was smaller for HC compared to ADHD in first-stage choices (see Fig. 2B). Specifically, a change of 0.5 in the difference between the expected value of the two offered actions was associated with an increase of 8 ms in the tail of the RT distribution for the HC group. However, the ADHD showed 136 ms change for the same amount of difference [posterior median for group X choice-difficulty effect on  $\tau$  estimates on the first stage = +0.60  $CI_{89\%}$  between 0.03 to 1.16,  $CI_{95\%}$  between  $-0.08$  and 1.28, pd = 95.25%; see full posterior in Fig. S6A]. Strikingly, for the second and third stages, this effect reversed with a slightly larger effect of choice difficulty on RT variability estimates for ADHD compared to HC. Specifically, the group effect on  $\tau$  estimates was negative for both the second stage [posterior median =  $-0.37$   $CI_{89\%}$  between  $-1.05$  and 0.29,  $CI_{95\%}$  between  $-1.19$  and 0.43, pd = 80.00%; see full posterior in Fig. S6B] and third stage [posterior median =  $-0.60$   $CI_{89\%}$  between  $-1.28$  and 0.05,  $CI_{95\%}$  between  $-1.41$  and 0.24, pd = 92.35%; see full posterior in Fig. S6C; for coefficients report see Table S2]. (For additional analysis regarding reaction time variability rates across groups, see SI).





**Fig. 2** Effect of choice difficulty on accuracy rates and reaction-time variability. **A** The influence of choice difficulty, group (HC vs. ADHD), and stage (1, 2, or 3) on participants' accuracy rates. First, results demonstrate that across stages and groups, accuracy rates improve as a function of choice difficulty. Importantly, ADHD individuals showed an attenuated sensitivity to choice difficulty changes only in the first stage but not in the second and third stages. **B** The influence of choice difficulty, group, and stage on participants' RT variability. Results demonstrate that across stages and groups, RT variability improved when choices were easier. Importantly, ADHD individuals showed an attenuated sensitivity to choice difficulty changes only in the first stage (y-axis represents  $\log(\tau)$  estimates for an ex-gaussian distribution fitted to empirical data).

### Computational reinforcement learning modeling

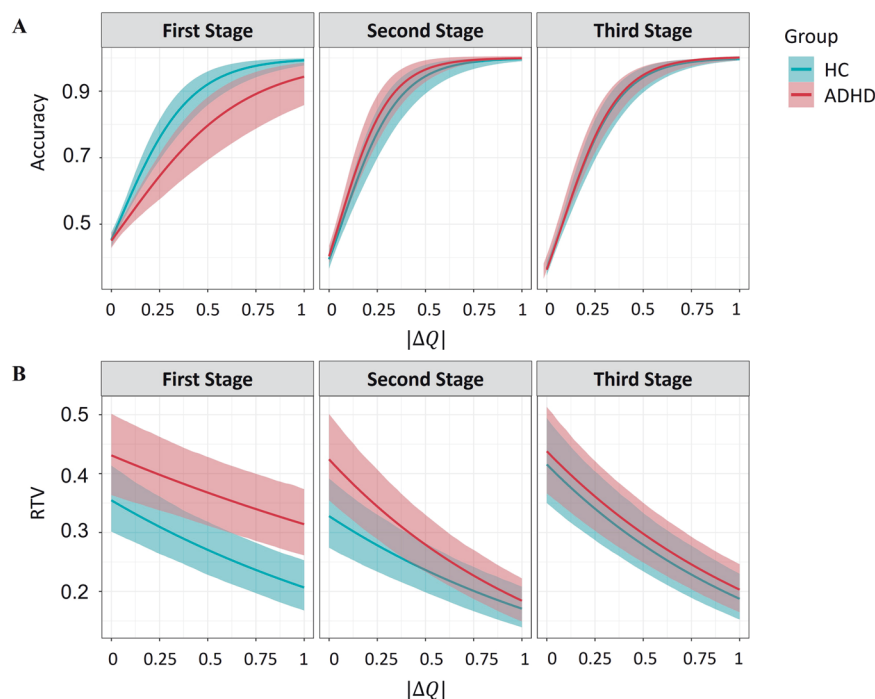
The previous analysis showed that ADHD individuals were less sensitive to the difference in the true expected values of the different choices in the first stage using both choice-accuracy and RT variability estimates. However, true expected values are quantified according to the pre-defined experimental design. Since the current task is a reinforcement learning task, the reward history is not only a function of the experimental design but also a function of the participant's choices. To track and gain a better estimate of the individual internal action value for each trial and stage (also referred to as  $Q$ -values in reinforcement learning), we used a well-established hierarchical reinforcement learning modeling (for details, see SI). We fitted the model to each group separately using hierarchical Bayesian parameter estimation (for details, see SI). We further performed a nested model comparison and found that a saturated seven-parameter model had the best fit to the empirical data (for details, see SI).

**Accuracy rates estimated using internal action values.** We used the expected value of the  $Q$ -value posterior distribution obtained for each action in our model Bayesian sampling as an estimate for the internal action values for each action (see "Methods"). To estimate choice accuracy, we first defined an accurate choice as one that reflects a selection in the higher internal value action in each stage (defined as 1 if the action with the higher internal value was taken and 0 otherwise). We repeated the same regression analyses reported in the previous section, with absolute differences in internal value (i.e.,  $|\Delta Q|$ ), this notion is used due to the common notion of " $Q$ -values" for internal values in the reinforcement learning literature) group (ADHD vs. HC), stage (1, 2, and 3) and their interactions as predictors for choice accuracy.

First, we found that higher  $Q$ -value differences were coupled with higher choice accuracy across stages and groups [posterior median = 5.39  $CI_{89\%}$  between 4.19 and 6.50,  $CI_{95\%}$  between 3.90 and 6.75,  $pd \sim 100\%$ ; see full posterior in Fig. S7A]. Specifically, a change of 0.5 in the difference between the two offered action values was associated

with an increase of 51% in accuracy rates. This finding shows that the model was able to capture some estimates for values that indeed predict participants' behavior well. Second, we further found a substantial group  $\times |\Delta Q|$  interaction for Stage 1, so that ADHD exhibited lower sensitivity to  $Q$ -values difference in terms of choice accuracy compared with HC [posterior median of group  $\times |\Delta Q|$  effect in the first stage =  $-2.19$   $CI_{89\%}$  between  $-3.75$  and  $-0.63$ ,  $CI_{95\%}$  between  $-4.06$  and  $-0.25$ ,  $pd = 98.65\%$ ; see full posterior in Fig. S7B; see Fig. 3A]. Strikingly, similarly to former results, group differences vanished completely and even changed direction when examining second and third stage choice. Specifically, we found a tendency toward a group  $\times |\Delta Q|$  interaction in the opposite direction to that of the first stage when examining the second stage [posterior median = 3.11  $CI_{89\%}$  between 1.21 to 5.13,  $CI_{95\%}$  between 0.67 and 5.54,  $pd = 99.30\%$ ; see full posterior in Fig. S7C] and the third stage [posterior median = 2.40  $CI_{89\%}$  between 0.40 and 4.34,  $CI_{95\%}$  between  $-0.12$  and 4.82,  $pd = 97\%$ ; see full posterior in Fig. S7D]. Overall, these findings resonated with the finding that ADHD had lower decision noise parameters at Stage 1 compared to their peers.

**Reaction-time variability as a function of internal action values.** We sampled a hierarchical Bayesian regression where the absolute difference in  $Q$ -values (i.e.,  $|\Delta Q|$ ) RT variability group (ADHD vs. HC), stage (1, 2, or 3), and their paired interaction as predictors for the  $\tau$  parameter in an ex-Gaussian distribution fitted to trial-by-trial reaction-time estimates (see "Methods"). First, we found that higher  $Q$ -value differences were coupled with lower  $\tau$  values across all stages [posterior median =  $-0.54$   $CI_{89\%}$  between  $-0.67$  and  $-0.41$ ,  $CI_{95\%}$  between  $-0.70$  and  $-0.38$ ,  $pd \sim 100\%$ ; see full posterior in Fig. S8A]. This finding supports the validity of our  $|\Delta Q|$  since longer RT tails are expected for decisions with similar values [58–60]. Furthermore, our model estimates of internal action values were based only on choices, without introducing RTs to the model during parameter estimation. Thus, finding this expected association  $|\Delta Q|$  serves as an important sanity check for our modeling approach.



**Fig. 3 Effect of latent internal values difference for each action on accuracy rates and reaction-time variability.** **A** Choice accuracy estimates as a function of group, stage, and absolute difference between the internal values estimated using a reinforcement learning model (i.e.,  $|\Delta Q|$ ). **B** RT variability as a function of group, stage and absolute difference between internal action values (i.e.,  $|\Delta Q|$ ). (y-axis represents  $\log(\tau)$  estimates for an ex-gaussian distribution fitted to empirical data).

Second, we found a substantial  $\text{Group} \times |\Delta Q|$  interaction for Stage 1, so that ADHD exhibited lower sensitivity to Q-values difference in terms of  $\tau$  estimates [posterior median = 0.23  $\text{Cl}_{99\%}$  between 0.07 and 0.38,  $\text{Cl}_{95\%}$  between 0.03 and 0.41,  $\text{pd} = 99\%$ ; see full posterior in Fig. S8B; see Fig. 3B]. However, for the second stage, this effect reversed with a significant effect of Q-value difference on RT variability estimates for ADHD compared to HC [posterior median =  $-0.41$   $\text{Cl}_{99\%}$  between  $-0.63$  and  $-0.18$ ,  $\text{Cl}_{95\%}$  between  $-0.68$  and  $-0.14$ ,  $\text{pd} = 99.9\%$ ; See full posterior in Fig. S8C]. For the third stage, this effect remained negative with a slightly smaller effect of Q-value difference on RT variability estimates for ADHD compared to HC [posterior median =  $-0.20$   $\text{Cl}_{99\%}$  between  $-0.42$  and 0.03,  $\text{Cl}_{95\%}$  between  $-0.46$  and 0.08,  $\text{pd} = 91.2\%$ ; see full posterior in Fig. S8D; for coefficients report see Table S4].

## DISCUSSION

The current study examined the ability of ADHD individuals to assign and act upon internal action values across stages of a sequential decision-making task. Participants with/without ADHD made decisions across a three-step deterministic tree task. First, we used reinforcement learning modeling to gain estimates of the latent and internal action values for each individual, trial, and step. We replicated previous findings showing that across both groups, estimates of choice accuracy and RT variability improved when the two available actions were well differentiated in terms of their latent values [61, 62]. Second, we found that ADHD individuals were less sensitive to initiatory action value differences across both choice accuracy and RT variability estimates compared with their HC. Specifically, a larger difference between the values of the two available first-stage actions had smaller benefits on ADHD choices and RTs compared to HC. Remarkably, these group differences diminished and disappeared when making terminal third-stage actions to a point where the two groups were remarkably indistinguishable. Our results, therefore, suggested a

striking attenuated sensitivity to value actions in ADHD, only in initiatory but not terminal task action.

The current finding extends previous literature in several ways. First, previous studies have already demonstrated that individuals with ADHD are uniquely characterized by inconsistent decision-making, as apparent in both choice and reaction time estimates [61, 62]. These studies further suggested that these behavioral inconsistencies in ADHD tend to wash out under conditions of immediate reward. For example, ADHD has been found to exhibit marked behavioral inconsistency in reaction times (i.e., increased RT variability) when performing sustained attention/executive function tasks (for review, see [59, 62, 63]). In a meta-analysis report, RT variability was further found to show the highest effect size for group differences. Thus, the mechanisms that contribute to this ADHD phenotype can be of much importance to how we understand ADHD. The sustained attention/executive functions tasks that reported increased RT variability in ADHD were mostly performed in the absence of any immediate reward or feedback. Remarkably, when immediate feedback or reward was delivered, ADHD demonstrated a significant improvement in RT variability estimates. Improved RT variability with the delivery of immediate reward was found in a variety of paradigms, including go/no-go [63], choice reaction time task [64, 65], simple reaction time [66], and time discrimination task [67, 68]. However, these studies are characterized almost exclusively by choice-feedback coupling, such that the outcome is given after a single choice and uses between-block manipulations (blocks with/without feedback). Here, we show both increased RT variability in ADHD compared to HC and the absence of group difference in RT variability estimates *within a single trial*. This is important since it refutes claims that a general motivation induced across a block is the sole reason for improved performance in ADHD under conditions of immediate reward/feedback. Instead, our findings point toward a more refined value-action mechanism that leads to increased inconsistency in ADHD.

Second, we find that only for first-stage choices ADHD show attenuated sensitivity to action values in terms of RT variability. Therefore, one explanation for both increased inconsistency in RTs

in ADHD [59, 62, 63] and previous findings regarding the improvement in this estimate under immediate reward [63–66, 68] might be an attenuated ability to assign and follow internal action values. Specifically, when performing a task in the absence of immediate reward/feedback, the individual is required to hold in mind the overall value of task completion. This might lead to noisier action values in ADHD compared with HC, which in turn is reflected in noisier reaction times. Therefore, individuals need to hold an overall value of task completion to keep interest and guide the selection of accurate actions. Since individuals with ADHD have known deficits in reward neural processing and tend to discount future outcomes [68], they may encounter particular difficulty in assigning value to current actions based on a distant goal. Indeed, reinforcement learning studies have confirmed both in simulations and empirical studies that a reduced ability to assign value to actions leads to prolonged and more inconsistent RTs, a finding that we also replicated in the current study across groups. Therefore, we speculate that the process of internally assigning value to actions can be a major contributing mechanism to increased RT variability in ADHD.

Why do ADHD individuals exhibit noisier choices during first-stage choices? The current study cannot offer a clear explanation, yet we will discuss a few explanations. First, it might be that ADHD individuals are more explorative in their behavior [6, 25, 61, 69, 70]. Specifically, a few reinforcement learning studies suggested that less value-directed behavior in ADHD might be the result of the directed decision to explore rather than a reduced ability to follow action values. For example, Dubois et al. [69] examined explore/exploit decision-making among children in early and late adolescence. They found that participants with high scores in self-report symptoms of ADHD made more exploratory decisions. Next, Dubois & Hauser [71] further linked exploration and the impulsivity component within ADHD in an explore/exploit decision-making task. Addicott et al. [72] also examined explore/exploit decision-making using the 6-armed bandit task where, on each trial, six bandit options were depicted on a computer screen, and participants selected one and received a reward in various probabilities. They compared ADHD and HC groups of adults on explore/exploit decisions that were modeled using reinforcement learning algorithms. They found that ADHD participants made more exploratory decisions (i.e., chose options without the highest expected reward value) and earned fewer points than HC in all three study days. Frank et al. [25] examined go/no-go decision-making between ADHD and HC groups of young adults. They additionally used computational models in order to predict participants' behavior. They found that ADHD participants were more likely to display inconsistent choices from trial to trial and that increased inconsistent behavior was correlated with increased variability in reaction times.

Another explanation of the reduced tendency to act according to value actions in first stage choice in ADHD might be the temporal distance from the outcome. That is, in the current task, we cannot rule out the fact that it is not the actual sequence or the fact that first-stage choices were initiatory actions in a sequence of three decisions. Rather, it might be that the temporal delay per se is the reason for the reduced ability to follow action values. We note that a previous study suggested that, counter-intuitively, ADHD individuals improve their ability to learn from feedback when the action outcome epoch is prolonged [73]. However, dedicated studies should carefully disentangle temporal delay from task stages to account for this explanation.

Moreover, it is interesting to speculate the role of another cognitive process that was identified as associated with ADHD and the current finding. For example, the ability to assign and follow action values for initiatory actions might require executive functions such as working memory updating and sustained attention. Working memory plays a fundamental role in cognition, allowing one to hold information "in mind" and is considered a

well-known attentional control system [40]. Since sequential learning consists of multiple stages that need to be learned, working memory capacity might be an important trait that affects this ability. Previous studies have highlighted individual differences in working memory capacity, demonstrating that lower capacity is associated with higher attentional deficits [36, 40, 74–78]. Furthermore, our findings may be explained by considering sustained attention, which pertains to an individual's capacity to sustain focus on a task or stimulus over an extended period. Given the relatively prolonged duration of the multiple-stage task (approximately 50 min) and the demanding nature of the cognitive processes required to navigate its stages and outcomes, sustained attention emerges as a potentially influential factor affecting task performance. Difficulties in sustained attention are one of the core deficits in ADHD [79] and are reflected by increased omission errors [80, 81], lack of inhibitory control by an increased number of commission errors [80, 81], and unstable cognitive processing by increased reaction time variability [82] as shown using a continuous performance task (CPT). Hence, future studies can examine with a larger sample size the ability of working memory and/or sustain attention to moderate the group differences.

In terms of generalizing our results to the overall ADHD population, we note that our study was performed using young adults with/without ADHD. The sample was a clinical sample that included participants who sought treatment in the community, and as such, can be seen as part of the much-needed growing literature regarding ADHD in this age group [83]. However, it would be important to examine whether the observed attenuated sensitivity to initiatory action-value is also reflected in young children. Specifically, we speculate that since executive functions are only developing during early ages [84], and since these executive functions might take a role in action-value assignment, we should expect similar and even larger effects in children with ADHD compared to what was observed in young adults. Therefore, there is a need for further studies in different age groups that would also control for the moderation effect of different executive functions (i.e., working memory/response inhibition). Furthermore, our sample included only college students and, therefore, might show higher general cognitive abilities than the overall young adult population of the same age. Therefore, this might somewhat limit our ability to generalize our findings to the overall young adult population. However, IQ testing suggested that general cognitive abilities for both ADHD and HC groups are within the expected age norms and widely spread across the IQ range (see SI). Furthermore, we found no differences in IQ between the ADHD and HC groups. Therefore, we argue that despite being a college population, the current sample is not dramatically different from the overall young adult population.

Another point that should be considered in future studies is that deficits in value-based actions are not exclusive to ADHD and were also observed in certain impulsive disorders, including substance use disorders (SUD) and alcohol use disorders (AUD). Given the well-documented comorbidities between ADHD and SUD/AUD, perhaps attributable to their shared core symptoms of impulsivity [85, 86], we should consider the potential influence of these impulsive disorders on action-value processing effects [87, 88]. E.g., Sebold et al. [89] demonstrated that AUD is associated with alterations in value-based decision-making. Furthermore, in the context of SUD, Groman et al. [90] conducted a review and showed evidence suggesting that value updating following positive outcomes, but not negative outcomes, predicts escalation in SUD. However, given none of the participants were diagnosed with impulse disorders, and the self-report questionnaires assessing AUD and SUD yielded similar results between the groups, we can conclude that our results are not mainly driven by increased addiction/impulsive disorders in the experimental group.

Finally, a few limitations should be mentioned. First, a major limitation of this study pertains to the limited sample size. Specifically, the sample for this study comprised 54 participants, with 26 HC and 28 ADHD participants. Due to the possible inflated influence of small sample sizes on effect sizes that tend to characterize ADHD research [91], future studies with larger numbers of patients are required to confirm our results and conclusions. Notwithstanding, we employed Bayesian statistical methodologies in our analysis, a choice informed by the recognition that Bayesian approaches often exhibit reduced susceptibility to sample size constraints compared to classical frequentist statistical methods (For a review, see [92]) and allow to more delicately describe the uncertainty in the estimates. Second, our design does not allow us to differentiate whether ADHD individuals adopted different speed-accuracy response criteria in first-stage choices. In future work, integration of evidence accumulation model (e.g., DDM) parameters may also provide more clarity, as these models can provide estimates of choice sensitivity that are informed by both RT and accuracy [60, 93]. Third, our study cannot disentangle whether temporal distance from the outcome, unique properties of first-stage actions, or both are behind the effect, which should be examined using dedicated experimental manipulation in future research.

In conclusion, individuals with ADHD exhibit reduced choice accuracy and higher RT variability compared to HC individuals for first-stage, initiatory action but not for the third-stage, terminal action. Our computational modeling uncovered that individuals with and without ADHD assign and act upon internal action values in sequential decision-making tasks, and as shown here, there are differences between the groups in the ability to do so. This highlights the need to address and deeper understand the internal action values as an important mechanism in ADHD.

#### DATA AVAILABILITY

Data, regression analysis, and computational modeling analysis codes are publicly available at <https://osf.io/ws47k/>.

#### REFERENCES

- American Psychiatric Association (ed.). *Diagnostic and statistical manual of mental disorders: DSM-5*. 5th ed. American Psychiatric Association: Washington, D.C; 2013.
- Faraone SV, Banaschewski T, Coghill D, Zheng Y, Biederman J, Bellgrove MA, et al. The World Federation of ADHD International Consensus Statement: 208 Evidence-based conclusions about the disorder. *Neurosci Biobehav Rev*. 2021;128:789–818.
- Thomas R, Sanders S, Doust J, Beller E, Glasziou P. Prevalence of attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. *Pediatrics*. 2015;135:e994–e1001.
- Lara C, Fayyad J, de Graaf R, Kessler RC, Aguilar-Gaxiola S, Angermeyer M, et al. Childhood predictors of adult attention-deficit/hyperactivity disorder: results from the World Health Organization World Mental Health Survey Initiative. *Biol Psychiatry*. 2009;65:46–54.
- Langberg JM, Smith ZR, Dvorsky MR, Molitor SJ, Bourchtein E, Eddy LD, et al. Factor structure and predictive validity of a homework motivation measure for use with middle school students with attention-deficit/hyperactivity disorder (ADHD). *Sch Psychol Q*. 2018;33:390–8.
- Seesjärvi E, Puhakka J, Aronen ET, Lipsanen J, Mannerkoski M, Hering A, et al. Quantifying ADHD symptoms in open-ended everyday life contexts with a new virtual reality task. *J Atten Disord*. 2022;26:1394–411.
- Sibley MH, Graziano PA, Ortiz M, Rodriguez L, Coxe S. Academic impairment among high school students with ADHD: the role of motivation and goal-directed executive functions. *J Sch Psychol*. 2019;77:67–76.
- Dayan P, Niv Y. Reinforcement learning: the good, the bad and the ugly. *Curr Opin Neurobiol*. 2008;18:185–96.
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science*. 1997;275:1593–9.
- Sutton RS, Barto A. Reinforcement learning: an introduction. Second edition. The MIT Press: Cambridge, MA London; 2018.
- Lehmann MP, Xu HA, Liakoni V, Herzog MH, Gerstner W, Preusschoff K. One-shot learning and behavioral eligibility traces in sequential decision making. *eLife*. 2019;8:e47463.
- Singh SP, Sutton RS. Reinforcement learning with replacing eligibility traces. *Mach Learn*. 1996;22:123–58.
- Botvinick MM, Niv Y, Barto AC. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. *Cognition*. 2009;113:262–80.
- Schultz W, Dickinson A. Neuronal coding of prediction errors. *Annu Rev Neurosci*. 2000;23:473–500.
- Dezfouli A, Balleine BW. Actions, action sequences and habits: evidence that goal-directed and habitual action control are hierarchically organized. *PLoS Comput Biol*. 2013;9:e1003364.
- Botvinick MM. Hierarchical reinforcement learning and decision making. *Curr Opin Neurobiol*. 2012;22:956–62.
- Dezfouli A, Balleine BW. Habits, action sequences and reinforcement learning. *Eur J Neurosci*. 2012;35:1036–51.
- Fiorillo CD, Tobler PN, Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*. 2003;299:1898–902.
- Houk JC, Adams JL, Barto AG. A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: *Models of information processing in the basal ganglia*. The MIT Press: Cambridge, MA, US; 1995. p. 249–70.
- Montague PR, Hyman SE, Cohen JD. Computational roles for dopamine in behavioural control. *Nature*. 2004;431:760–7.
- Schultz W. Dopamine reward prediction-error signalling: a two-component response. *Nat Rev Neurosci*. 2016;17:183–95.
- Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. *Science*. 2005;307:1642–5.
- Biederman J, Faraone SV. Current concepts on the neurobiology of attention-deficit/hyperactivity disorder. *J Atten Disord*. 2002;6:7–16.
- Del Campo N, Fryer TD, Hong YT, Smith R, Brichard L, Acosta-Cabronero J, et al. A positron emission tomography study of nigro-striatal dopaminergic mechanisms underlying attention: implications for ADHD and its treatment. *Brain*. 2013;136:3252–70.
- Frank MJ, Santamaria A, O'Reilly RC, Willcutt E. Testing computational models of dopamine and noradrenaline dysfunction in attention deficit/hyperactivity disorder. *Neuropsychopharmacology*. 2007;32:1583–99.
- Luman M, Tripp G, Scheres A. Identifying the neurobiology of altered reinforcement sensitivity in ADHD: a review and research agenda. *Neurosci Biobehav Rev*. 2010;34:744–54.
- Plichta MM, Scheres A. 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*. 2014;38:125–34.
- Sagvolden T, Johansen EB, Aase H, Russell VA. A dynamic developmental theory of attention-deficit/hyperactivity disorder (ADHD) predominantly hyperactive/impulsive and combined subtypes. *Behav Brain Sci*. 2005;28:397–419.
- Solanto MV, Arnsten AFT, Castellanos FX. Stimulant drugs and ADHD: basic and clinical neuroscience. Oxford University Press; 2001.
- Volkow ND, Wang G-J, Newcorn JH, Kollins SH, Wigal TL, Telang F, et al. Motivation deficit in ADHD is associated with dysfunction of the dopamine reward pathway. *Mol Psychiatry*. 2011;16:1147–54.
- Demurie E, Roeyers H, Baeyens D, Sonuga-Barke E. Common alterations in sensitivity to type but not amount of reward in ADHD and autism spectrum disorders. *J Child Psychol Psychiatry*. 2011;52:1164–73.
- Luman M, Oosterlaan J, Sergeant JA. The impact of reinforcement contingencies on AD/HD: a review and theoretical appraisal. *Clin Psychol Rev*. 2005;25:183–213.
- Sonuga-Barke EJS, Taylor E, Sembi S, Smith J. Hyperactivity and delay aversion—I. The effect of delay on choice. *J Child Psychol Psychiatry*. 1992;33:387–98.
- Hart H, Radua J, Nakao T, Mataix-Cols D, Rubia K. Meta-analysis of functional magnetic resonance imaging studies of inhibition and attention in attention-deficit/hyperactivity disorder: exploring task-specific, stimulant medication, and age effects. *JAMA Psychiatry*. 2013;70:185–98.
- Krieger V, Amador-Campos JA, Guàrdia-Olmos J. Executive functions, personality traits and ADHD symptoms in adolescents: a mediation analysis. *PLoS ONE*. 2020;15:e0232470.
- Martinussen R, Hayden J, Hogg-Johnson S, Tannock R. A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2005;44:377–84.
- Mullane JC, Corkum PV. The Relationship Between Working Memory, Inhibition, and Performance on the Wisconsin Card Sorting Test in Children With and Without ADHD. 2007. <https://journals.sagepub.com/doi/10.1177/0734282906297627>. Accessed 6 Oct 2022.
- Pedersen A, Ohmann P. Impaired behavioral inhibition in implicit sequence learning in adult ADHD. *J Atten Disord*. 2018;22:250–60.



39. Pievsky MA, McGrath RE. The neurocognitive profile of attention-deficit/hyperactivity disorder: a review of meta-analyses. *Arch Clin Neuropsychol*. 2018;33:143–57.
40. Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF. Validity of the Executive Function Theory of Attention-deficit/hyperactivity Disorder: A Meta-analytic Review. *Biol Psychiatry*. 2005;57:1336–46.
41. Hulsbosch A-K, De Meyer H, Beckers T, Danckaerts M, Van Liefferinge D, Tripp G, et al. Systematic review: attention-deficit/hyperactivity disorder and instrumental learning. *J Am Acad Child Adolesc Psychiatry*. 2021;60:1367–81.
42. Kooij JJS, Francken MH, Bron TI. Diagnostic interview for ADHD in Adults 2.0 (DIVA 2.0). 2010.
43. Zohar AH, Konfortes H. Diagnosing ADHD in Israeli adults: the psychometric properties of the adult ADHD self report scale (ASRS) in Hebrew. *Isr J Psychiatry Relat Sci*. 2010;47:308–15.
44. Ward MF, Wender PH, Reimherr FW. The Wender Utah Rating Scale: an aid in the retrospective diagnosis of childhood attention deficit hyperactivity disorder. *Am J Psychiatry*. 1993;150:885–90.
45. Raven J, Raven JC, Court JH. Manual for Raven's progressive matrices and vocabulary scales. Pearson: San Antonio, TX; 1998.
46. Beck AT, Steer RA, Brown G. Beck Depression Inventory–II. 1996. <https://doi.org/10.1037/t00742-000>.
47. Spielberger CD. State-Trait Anxiety Inventory for Adults. 1983. <https://doi.org/10.1037/t06496-000>.
48. Foa EB, Kozak MJ, Salkovskis PM, Coles ME, Amir N. The validation of a new obsessive–compulsive disorder scale: the obsessive–compulsive inventory. *Psychol Assess*. 1998;10:206–14.
49. Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E. The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J Autism Dev Disord*. 2001;31:5–17.
50. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. AUDIT: The alcohol use disorders identification test: Guidelines for use in primary health care. 2nd ed. Geneva: World Health Organization; 2001.
51. Adamson SJ, Kay-Lambkin FJ, Baker AL, Lewin TJ, Thornton L, Kelly BJ, et al. An improved brief measure of cannabis misuse: the Cannabis Use Disorders Identification Test-Revised (CUDIT-R). *Drug Alcohol Depend*. 2010;110:137–43.
52. Matzke D, Wagenmakers E-J. Psychological interpretation of the ex-Gaussian and shifted Wald parameters: a diffusion model analysis. *Psychon Bull Rev*. 2009;16:798–817.
53. Alderson RM, Rapport MD, Kofler MJ. Attention-deficit/hyperactivity disorder and behavioral inhibition: a meta-analytic review of the stop-signal paradigm. *J Abnorm Child Psychol*. 2007;35:745–58.
54. Castellanos FX, Sonuga-Barke EJS, Scheres A, Di Martino A, Hyde C, Walters JR. Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biol Psychiatry*. 2005;57:1416–23.
55. Epstein JN, Langberg JM, Rosen PJ, Graham A, Narad ME, Antonini TN, et al. Evidence for higher reaction time variability for children with ADHD on a range of cognitive tasks including reward and event rate manipulations. *Neuropsychology*. 2011;25:427.
56. Dezfouli A, Lingawi NW, Balleine BW. Habits as action sequences: hierarchical action control and changes in outcome value. *Philos Trans R Soc B Biol Sci*. 2014;369:20130482.
57. Solway A, Diuk C, Córdova N, Yee D, Barto AG, Niv Y, et al. Optimal behavioral hierarchy. *PLoS Comput Biol*. 2014;10:e1003779.
58. Peters J, D'Esposito M. The drift diffusion model as the choice rule in inter-temporal and risky choice: a case study in medial orbitofrontal cortex lesion patients and controls. *PLoS Comput Biol*. 2020;16:e1007615.
59. Shahar N, Teodorescu AR, Karmon-Presser A, Anholt GE, Meiran N. Memory for action rules and reaction time variability in attention-deficit/hyperactivity disorder. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2016;1:132–40.
60. Shahar N, Hauser TU, Moutoussis M, Moran R, Keramati M, NSPN consortium. et al. Improving the reliability of model-based decision-making estimates in the two-stage decision task with reaction-times and drift-diffusion modeling. *PLoS Comput Biol*. 2019;15:e1006803.
61. 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;71:1165–73.
62. Kofler MJ, Rapport MD, Sarver DE, Raiker JS, Orban SA, Friedman LM, et al. Reaction time variability in ADHD: A meta-analytic review of 319 studies. *Clin Psychol Rev*. 2013;33:795–811.
63. Adamo N, Hodson J, Asherson P, Buitelaar JK, Kuntsi J. Ex-Gaussian, frequency and reward analyses reveal specificity of reaction time fluctuations to ADHD and not autism traits. *J Abnorm Child Psychol*. 2019;47:557–67.
64. Tye C, Johnson KA, Kelly SP, Asherson P, Kuntsi J, Ashwood KL, et al. Response time variability under slow and fast-incentive conditions in children with ASD, ADHD and ASD+ADHD. *J Child Psychol Psychiatry*. 2016;57:1414–23.
65. Vainieri I, Adamo N, Michelini G, Kitsune V, Asherson P, Kuntsi J. Attention regulation in women with ADHD and women with bipolar disorder: an ex-Gaussian approach. *Psychiatry Res*. 2020;285:112729.
66. Gopin CB, Berwid O, Marks DJ, Mlodnicka A, Halperin JM. ADHD preschoolers with and without ODD: do they act differently depending on degree of task engagement/reward? *J Atten Disord*. 2013;17:608–19.
67. Marx I, Höpcke C, Berger C, Wandschneider R, Herpertz SC. The impact of financial reward contingencies on cognitive function profiles in adult ADHD. *PLoS ONE*. 2013;8:e67002.
68. Marx I, Hacker T, Yu X, Cortese S, Sonuga-Barke E. 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. *J Atten Disord*. 2021;25:171–87.
69. 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;22:969–83.
70. 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.
71. Dubois M, Hauser TU. Value-free random exploration is linked to impulsivity. *Nat Commun*. 2022. <https://doi.org/10.1038/s41467-022-31918-9>.
72. Addicott MA, Pearson JM, Schechter JC, Sapyta JJ, Weiss MD, Kollins SH. Attention-deficit/hyperactivity disorder and the explore/exploit trade-off. *Neuropsychopharmacology*. 2021;46:614–21.
73. Gabay Y, Shahbari-Khateb E, Mendelsohn A. Feedback timing modulates probabilistic learning in adults with ADHD. *Soc Rep*. 2018;8:15524.
74. Dovis S, Van der Oord S, Wiers RW, Prins PJM. What part of working memory is not working in ADHD? Short-term memory, the central executive and effects of reinforcement. *J Abnorm Child Psychol*. 2013;41:901–17.
75. Ehliis A-C, Bähne CG, Jacob CP, Herrmann MJ, Fallgatter AJ. Reduced lateral prefrontal activation in adult patients with attention-deficit/hyperactivity disorder (ADHD) during a working memory task: a functional near-infrared spectroscopy (fNIRS) study. *J Psychiatr Res*. 2008;42:1060–7.
76. Kofler MJ, Rapport MD, Bolden J, Sarver DE, Raiker JS, Alderson RM. Working memory deficits and social problems in children with ADHD. *J Abnorm Child Psychol*. 2011;39:805–17.
77. Massat I, Slama H, Kavac M, Linotte S, Mary A, Baleriaux D, et al. Working memory-related functional brain patterns in never medicated children with ADHD. *PLoS ONE*. 2012;7:e49392.
78. Rapport MD, Alderson RM, Kofler MJ, Sarver DE, Bolden J, Sims V. Working memory deficits in boys with attention-deficit/hyperactivity disorder (ADHD): the contribution of central executive and subsystem processes. *J Abnorm Child Psychol*. 2008;36:825–37.
79. Tsal Y, Shalev L, Mevorach C. The diversity of attention deficits in ADHD: the prevalence of four cognitive factors in ADHD versus controls. *J Learn Disabil*. 2005;38:142–57.
80. Losier BJ, McGrath PJ, Klein RM. Error patterns on the continuous performance test in non-medicated and medicated samples of children with and without ADHD: a meta-analytic review. *J Child Psychol Psychiatry*. 1996;37:971–87.
81. Metin B, Roeyers H, Wiersma JR, van der Meere J, Sonuga-Barke E. A meta-analytic study of event rate effects on Go/No-Go performance in attention-deficit/hyperactivity disorder. *Biol Psychiatry*. 2012;72:990–6.
82. Karalunas SL, Geurts HM, Konrad K, Bender S, Nigg JT. Annual research review: Reaction time variability in ADHD and autism spectrum disorders: measurement and mechanisms of a proposed trans-diagnostic phenotype. *J Child Psychol Psychiatry*. 2014;55:685–710.
83. Song P, Zha M, Yang Q, Zhang Y, Li X, Rudan I. The prevalence of adult attention-deficit hyperactivity disorder: a global systematic review and meta-analysis. *J Glob Health*. 2021;11:04009.
84. Rapport MD, Orban SA, Kofler MJ, Friedman LM. Do programs designed to train working memory, other executive functions, and attention benefit children with ADHD? A meta-analytic review of cognitive, academic, and behavioral outcomes. *Clin Psychol Rev*. 2013;33:1237–52.
85. Robbins TW, Gillan CM, Smith DG, de Wit S, Ersche KD. Neurocognitive endophenotypes of impulsivity and compulsivity: towards dimensional psychiatry. *Trends Cogn Sci*. 2012;16:81–91.
86. Winstanley CA, Eagle DM, Robbins TW. Behavioral models of impulsivity in relation to ADHD: translation between clinical and preclinical studies. *Clin Psychol Rev*. 2006;26:379–95.
87. Gueguen MCM, Schweitzer EM, Konova AB. Computational theory-driven studies of reinforcement learning and decision-making in addiction: what have we learned? *Curr Opin Behav Sci*. 2021;38:40–8.
88. Sweis BM, Thomas MJ, Redish AD. Beyond simple tests of value: measuring addiction as a heterogeneous disease of computation-specific valuation processes. *Learn Mem Cold Spring Harb N*. 2018;25:501–12.

89. Sebold M, Kiebel SJ, Smolka MN, Heinz A, Deserno L. Computational theories of alcohol use disorder: mapping learning and choice mechanisms on symptoms. *Neuropsychobiology*. 2022;81:339–56.
90. Groman SM, Thompson SL, Lee D, Taylor JR. Reinforcement learning detuned in addiction: integrative and translational approaches. *Trends Neurosci*. 2022;45:96–105.
91. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci*. 2013;14:365–76.
92. McNeish D. On using bayesian methods to address small sample problems. *Struct Equ Model Multidiscip J*. 2016;23:750–73.
93. Pedersen ML, Frank MJ, Biele G. The drift diffusion model as the choice rule in reinforcement learning. *Psychon Bull Rev*. 2017;24:1234–51.
94. Akam T, Costa R, Dayan P. Simple plans or sophisticated habits? State, transition and learning interactions in the two-step task. *PLOS Comput Biol*. 2015;11:e1004648.
95. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. *Nature*. 2006;441:876–9.
96. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-based influences on humans' choices and striatal prediction errors. *Neuron*. 2011;69:1204–15.
97. Moran R, Dayan P, Dolan RJ. Human subjects exploit a cognitive map for credit assignment. *Proc Natl Acad Sci USA*. 2021;118:e2016884118.

### ACKNOWLEDGEMENTS

We would like to acknowledge the undergraduate students Eli Greisman and Tali Bogachek for their assistance in data collection. The study was supported by the Minducate Center for the Science of Learning (Sagol School of Neuroscience, Tel Aviv University), the Israel Ministry of Science and Technology, and the Israeli Science Foundation (Grant 2536/20 awarded to NS). Both authors contributed equally to all stages of the study.

### AUTHOR CONTRIBUTIONS

All aspects of the research, including conceptualization, methodology, formal analysis, investigation, resources management, data curation, and writing (original draft, review, and editing), were equally undertaken by GK and NS.

### COMPETING INTERESTS

The authors declare no competing interests.

### ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41398-023-02717-7>.

**Correspondence** and requests for materials should be addressed to Gili Katabi.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024