The promoting effects of digital targeted cognitive training in medication treatment for children with ADHD: A randomized controlled trial

SUPPLEMENTARY MATERIALS

1. Supplementary Table 1

Demographic data and baseline clinical symptoms for participants in both the combined treatment group and the group receiving medication alone, regardless of the medication type, are detailed in Supplementary Table 1. Between these two groups, there were no significant differences observed in sex, age, intelligence quotient (IQ), or symptom severity at baseline and follow-up (all $p > 0.1$).

Supplementary Table 1. Demographics and baseline clinical characteristics of participants (mean ± standard deviation).

	Cognitive Training + Medication treatment	Medication treatment	t / χ^2	p
N	60	64		
Male/female	53/7	58/6	0.173	0.677
Age (years)	9.4 ± 1.4	9.5 ± 1.7	-0.412	0.681
IQ	108.1 ± 13.9	110.2 ± 14.4	-0.827	0.410
ADHD-RS total (baseline)	34.2 ± 8.9	32.4 ± 6.7	1.259	0.211
Inattention (baseline)	19.5 ± 4.5	18.8 ± 3.3	1.005	0.317
Hyperactivity/impulsivity (baseline)	14.7 ± 5.7	13.6 ± 5.0	1.122	0.264
ADHD-RS total (follow-up)	23.1 ± 7.1	23.6 ± 6.9	-0.406	0.686
Inattention (follow-up)	13.0 ± 3.3	13.5 ± 3.8	-0.746	0.457
Hyperactivity/impulsivity (follow-up)	10.0 ± 4.9	10.1 ± 4.5	-0.035	0.972

2. Supplementary Table 2. Cognitive task performance at baseline and follow-up in four groups (mean ± standard deviation).

	CT + MPH		MPH		CT + ATX		ATX	
	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up	Baseline	Follow-up
Vigilance task (RT, ms)	604 ± 191	525 ± 136	591 ± 196	558 ± 138	737 ± 244	630 ± 292	906 ± 400	629 ± 122
Flanker task (RT, ms)	1255 ± 464	1019 ± 223	1243 ± 283	1058 ± 223	1377 ± 402	1113 ± 332	1546 ± 378	1307 ± 329
Random dot motion task (Acc.)	0.69 ± 0.21	0.79 ± 0.11	0.69 ± 0.19	0.79 ± 0.12	0.6 ± 0.27	0.76 ± 0.21	0.69 ± 0.19	0.73 ± 0.16
Go/No-Go task (Acc.)	0.84 ± 0.12	0.81 ± 0.2	0.78 ± 0.15	0.81 ± 0.11	0.73 ± 0.17	0.78 ± 0.2	0.74 ± 0.12	0.74 ± 0.14
Digit span forward task (Ca.)	6.93 ± 1.66	7.44 ± 1.72	7.82 ± 1.78	8.27 ± 2.24	6.19 ± 2.14	6.69 ± 1.91	7.18 ± 1.47	7 ± 1.48
Digit span backward task (Ca.)	4.41 ± 1.42	4.74 ± 1.68	5.64 ± 2.69	6.64 ± 2.11	4.58 ± 1.75	4.35 ± 2.15	4.55 ± 1.51	5 ± 2.1
K-score task (Ca.)	1.38 ± 0.53	1.45 ± 0.58	1.24 ± 0.67	1.3 ± 0.62	1.06 ± 0.54	1.2 ± 0.54	1.13 ± 0.45	1.14 ± 0.5
Emotion recognition task (Acc.)	0.81 ± 0.07	0.84 ± 0.08	0.9 ± 0.09	0.89 ± 0.1	0.78 ± 0.19	0.79 ± 0.13	0.8 ± 0.14	0.83 ± 0.1
Paper folding task (Acc.)	0.75 ± 0.13	0.78 ± 0.15	0.85 ± 0.18	0.9 ± 0.13	0.7 ± 0.19	0.78 ± 0.13	0.66 ± 0.19	0.68 ± 0.19
Time selection task (Acc.)	0.66 ± 0.25	0.64 ± 0.18	0.66 ± 0.34	0.81 ± 0.23	0.63 ± 0.19	0.65 ± 0.21	0.62 ± 0.24	0.6 ± 0.28

Abbreviations: CT, cognitive training; MPH, methylphenidate; ATX, atomoxetine; RT, reaction time; Acc., accuracy; Ca., capacity.

3. Detailed description of cognitive tasks

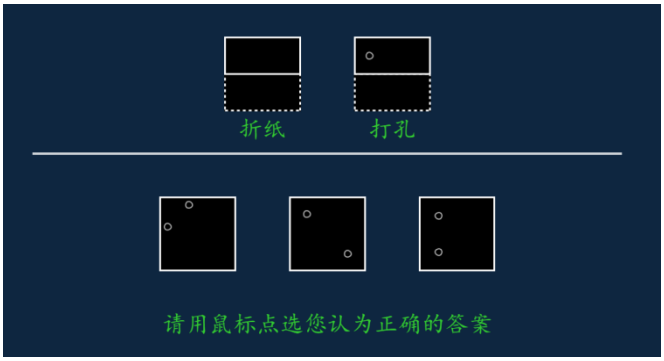
This study used 10 cognitive tasks to evaluate five domains (perception, attention, memory, executive function, and social cognition).

3.1. There were three perception tasks: a time selection task for assessing temporal perception, a paper folding task to evaluate spatial perception, and a random dot motion task aimed at examining motor perception.

In the time selection task, participants observed a red time bomb displayed on the screen and were tasked with estimating the duration until detonation, choosing the correct duration from five available options. Higher accuracy and shorter reaction time indicated better temporal perception.

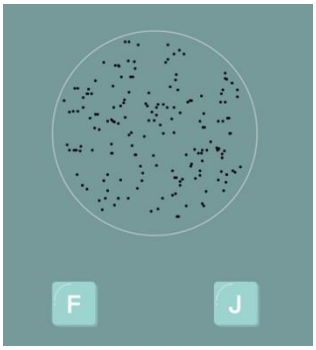
(Time selection task)

In the paper folding task, participants were guided to imagine folding a square piece of paper along specified lines, followed by using a pencil to penetrate all the layers of the folded paper before fully unfolding it. Subsequently, they were required to choose from three options, presented below a horizontal separation line, the accurate number and locations of the holes in the paper after it was completely unfolded. Higher accuracy and shorter reaction time indicated better spatial perception.



(Paper folding task)

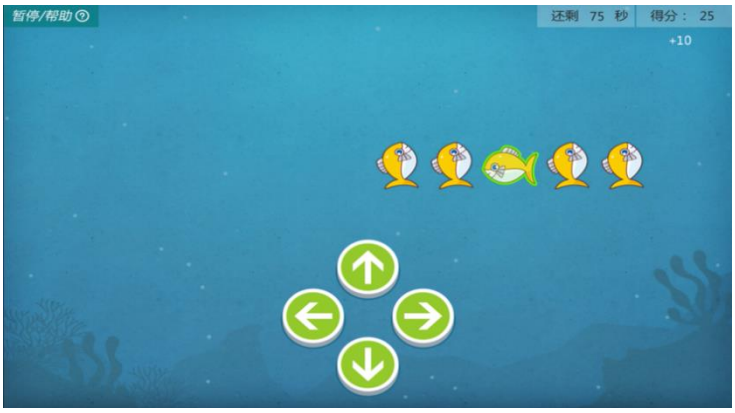
In the random dot motion task, the subject was asked to carefully observe a moving dot on the screen. Among dots with seemingly random motion, some dots regularly moved to the left or right. The subject had to determine the movement direction of these dots and press the corresponding button (F for left and J for right). Higher accuracy indicated better motor perception.



(Random dot motion task)

3.2. The attention domain included two tasks: the flanker task was used to evaluate selective attention, and the vigilance test was used to evaluate continuous attention.

In the flanker task, five fish would appear on the screen, and the subject was asked to determine the direction of the central fish while ignoring the other four fish and pressing the corresponding button. Higher accuracy and shorter reaction time indicated better selective attention.



(Flanker task)

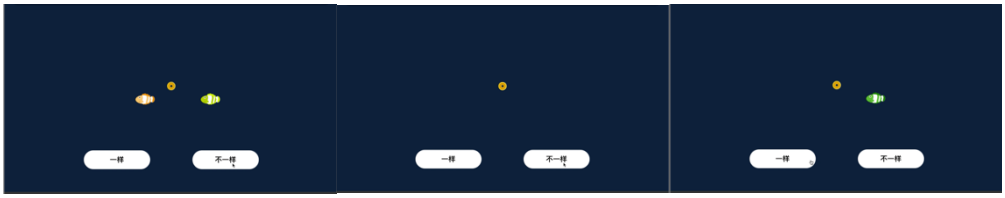
The psychomotor vigilance task (PVT) (Dinges et al., 1997) was used for the vigilant attention task. This task primarily tests whether the subject can promptly respond to sudden signals to evaluate vigilant attention. In the vigilance task, the screen would display “00000,” and the subject was instructed to focus their attention on this string of numbers and immediately press the “stop” button when the numbers in the “00000” string increased. Higher accuracy and shorter reaction time indicated better continuous attention.

(Vigilant attention task)

3.3. The memory domain included two tasks: A K-score task to evaluate spatial memory and a digit span forward task to evaluate memory span.

The K-score task used a delay-probe paradigm, divided into three stages: onset, delay, and

probe. First, two or four fish of different colors and sizes would appear on the screen, and the fish would quickly disappear. After a blank screen, a small fish would appear in a position where it used to be. The subject was instructed to determine if the fish was the same as the previous stimuli in the same position and press the button to respond. The first half of the task used two fish, and the second half used four fish. Higher accuracy indicated better spatial memory.



(K-score task)

In the digit span forward task, a string of numbers would appear on the screen one at a time. The participant had to remember these numbers and the sequence of appearance. After the string of numbers had been displayed, the patient was instructed to press the buttons 0–9 on the screen based on the sequence of appearance. The length of the number string started at two and became longer. Two attempts are given per length, increasing by one when the correct answer was chosen. The test ended when the participant provided the incorrect answer for both attempts for the same length. The longer the digit span at the end of the task, the better the memory span.

3.4. The executive function domain included two tasks. The go/no-go task was used to evaluate inhibition, while the digit span backward task was used to evaluate working memory.

The go/no-go measured reaction inhibition. This task randomly presented two alternating letters or pictures, and the subject was instructed to respond to one of them (go response) and not to the other stimuli (no-go response). This task had two rounds in total. In the first round, the screen would display the image of a “tiger” or “giraffe”. When the “tiger” image appeared, the subject was instructed to tap anywhere on the screen. The subject was instructed to not tap the screen when the “giraffe” image was displayed. In the second round, the screen would display the image of a

“lion” or “elephant”. When the “lion” image appeared, the subject was instructed to tap anywhere on the screen. The subject was instructed to not tap the screen when the “elephant” image appeared. Higher accuracy and shorter reaction time for the final task indicated better reaction inhibition.

(go/no-go task)

In the digit span backward task, a string of numbers would appear on the screen one at a time. The participant had to remember these numbers and the sequence of appearance. After the string of numbers had been displayed, the participant was instructed to press the buttons 0–9 on the screen in reverse based on the sequence of appearance. The length of the number string started at two and became longer. Two attempts were given per length, and the length was increased by one when the correct answer was provided. The test ended when the participant provided the incorrect answer for both attempts of the same length. The longer the digit span at the end of the task, the better the working memory.

3.5 The social cognition domain included one task: the emotion recognition task was used to evaluate emotion perception.

Emotional faces were used for the emotion recognition task to evaluate the accuracy of emotion recognition in the subject. During the test, the center of the screen would display images of an emoting face, and the participant was instructed to identify the expression on each face by

selecting from four options (happy, sad, angry, or scared). Higher accuracy indicated better emotional perception.

(Emotion recognition task)

4. Steps to develop personalized training plans

A personalized cognitive training plan was constructed based on the results of cognitive function assessments for each patient with ADHD. Specifically, the individual's cognitive ability scores were computed and utilized to determine the training value of each training task. Out of the five daily tasks sent, three were selected as training tasks with the highest training value for the individual, while the system randomly chose two additional training tasks to ensure the training regimen covered multiple cognitive domains. After the completion of daily training tasks, the system would reassess each cognitive ability based on the initial assessment results and the training outcomes at that day. These assessments were then used to determine the subsequent training regimen. In summary, personalized training was ensured by incorporating all cognitive paradigms and their variations, specifically targeting those with lower baseline scores.

5. Elaboration on the assessments used at baseline and follow-up

5.1 ADHD symptoms.

The ADHD Rating Scale (ADHD-RS) was used to assess changes in ADHD symptoms. The scale comprised 18 items, consisting of attention-deficit and hyperactivity/impulsivity subscales. A 0–3 Likert scale was employed for scoring, with "Never" assigned 0 points, "Occasionally" assigned 1 point, "Often" assigned 2 points, and "Always" assigned 3 points. The difference in ADHD-RS scores between baseline and follow-up (baseline minus follow-up) and the reduction rate [(baseline - follow-up)/baseline] were calculated. A higher reduction rate indicates more significant symptom improvement.

5.2 Ecological executive function.

The Behavior Rating Inventory of Executive Function (BRIEF) parental version comprises 86 items, from which two indices and eight factors can be derived. The indices include the Metacognition Index (MI) and Behavioral Regulation Index (BRI). The factors include Inhibit, Shift, Emotional Control, Initiate, Working Memory, Plan/Organize, Organization of Materials, and Monitor. Higher index and factor scores indicate that the corresponding impairment is more severe. Differences in various BRIEF indices and factor scores between the baseline and follow-up were calculated. A greater difference indicates more significant symptom improvement.

5.3 Laboratory cognitive function.

Ten tasks were employed to evaluate cognitive function, including perception, attention, memory, executive function, and social cognition, with the same tasks used for both baseline and follow-up assessments. Differences in task accuracy and reaction time between baseline and follow-up were calculated. A decrease in accuracy and an increase in reaction time differences indicated more significant symptom improvements.

5.4 Social functioning.

The Chinese Version of the WEISS Functional Impairment Scale - Parent (WFIRS-P) was used. This scale mainly includes six functional domains: family, learning and school, life skills, self-concept, social activities, and risky activities. A 4-point Likert scoring system is used for each item, with scores ranging from 0 to 3. The sum of the scores of valid items in each domain is divided by the number of valid items to obtain the corresponding domain score. A higher score indicates a more severe impairment of the corresponding social function.