Supplementary Material

Results and Discussion S1. Results and discussion of cancellation tests with ET in ADHD children.

Figure S1. Study flowchart of cancellation tests with ET in children with or without ADHD.

Figure S2. Results of three cancellation tests in male and female ADHD children. **A1.** Comparisons of the total scores between genders in ADHD children. **A2.** Comparisons of the correct scores between genders in ADHD children. **A3.** Comparisons of the wrong scores between genders in ADHD children. **A4.** Comparisons of the total scores between genders in ADHD children. **B1.** Comparisons of the total scores of the colored-fruit second and third subtests between genders in ADHD children. **B2.** Comparisons of the correct scores of the number fourth subtest, and colored-fruit second and third subtests between genders in ADHD children. **B3.** Comparisons of the wrong scores of the number fourth subtest, and colored-fruit second and third subtests between genders in ADHD children. **B4.** Comparisons of the wrong ratios of the number fourth subtest, and colored-fruit second and third subtests between genders in ADHD children. Data are presented as mean ± SEM. * $p < 0.05$.

Figure S3. Correlations between the correct/wrong scores and clinical scale evaluation indicators in male ADHD children.

Figure S4. Correlations between the correct/wrong scores and clinical scale evaluation indicators in female ADHD children.

Table S1. Demographic and clinical characteristics.

Table S2. ET indicators.

Table S3. Multiple linear regression analysis of correct and wrong scores of three-cancellation tests using ET and clinical scale indicators in male ADHD children.

Table S4. Multiple linear regression analysis of correct and wrong scores of three-cancellation tests using ET and clinical scale indicators in female ADHD children.

Table S5. Results of identification of patients with ADHD and male/female ADHD using ET indicators.

1. Results and Discussion S1

Results

1.1 Attentive bias during three cancellation testing in ADHD children

We compared the total scores (**Fig. S5. A1**), correct scores (**Fig. S5. A2**), wrong scores (**Fig. S5. A3**) and wrong ratios (**Fig. S5. A4**) between the ADHD and the HC group. We found that the correct scores of all three tests in ADHD group were lower than those in HC group, and the correct scores of colored fruit test were significantly decreased in ADHD group (**Fig. S5. A2**), also the total scores and wrong scores of colored fruit test were significantly decreased and increased respectively in ADHD group (**Fig. S5. A1A3**). Then we compared each of the five subtests of colored fruit test between two groups, and found that the wrong ratios of the third to fifth subtests in the ADHD group was significantly higher than that of the HC group (**Fig. S5. B1**), also the wrong scores of the third subtests in the ADHD group was significantly higher than that of the HC group (**Fig. S5. B2**).

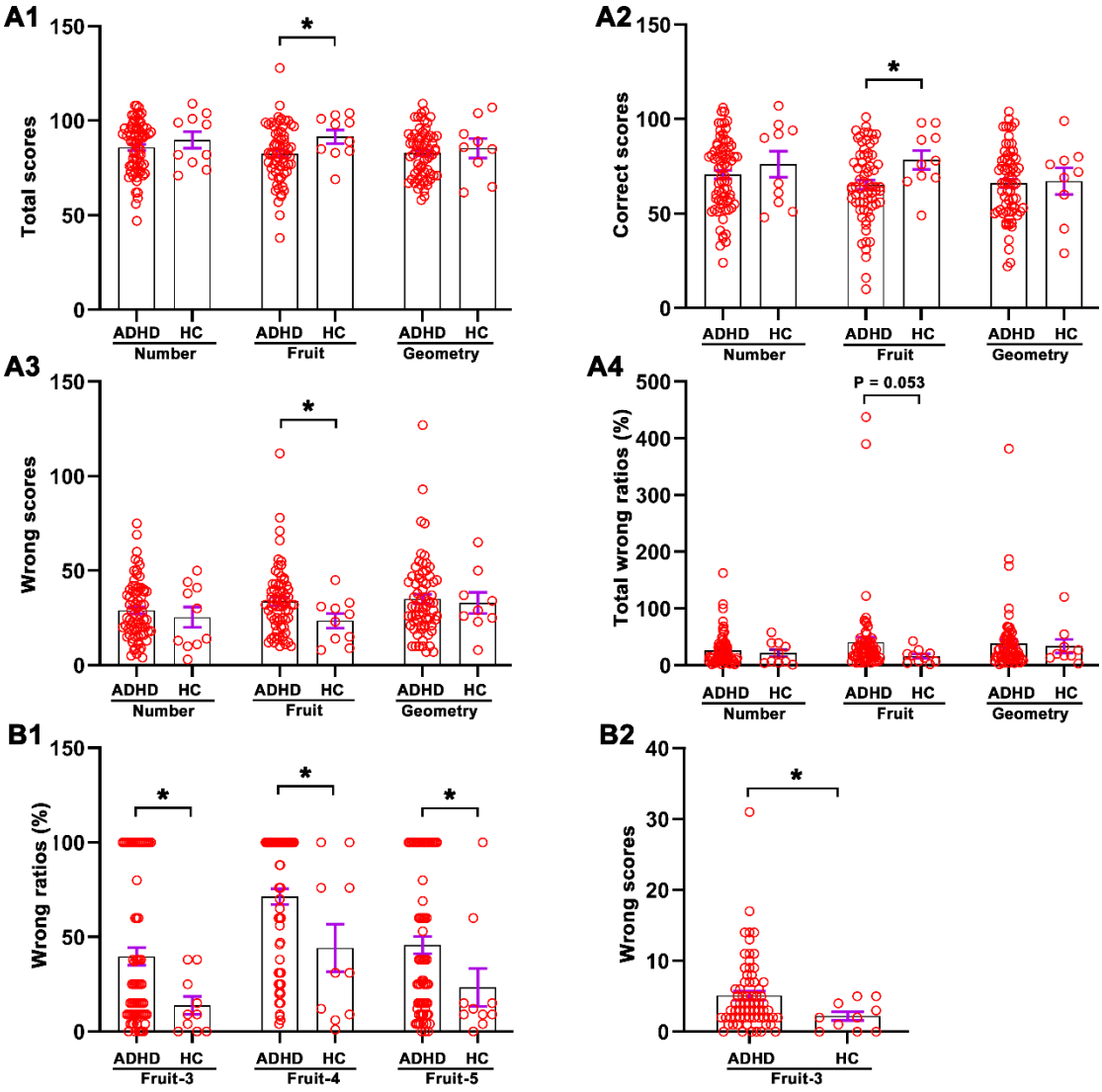

Figure S5. Results of three cancellation tests in ADHD children. **A1.** Comparisons of the total scores between the ADHD and the HC group. **A2.** Comparisons of the correct scores between the ADHD and the HC group. **A3.** Comparisons of the wrong scores between the ADHD and the HC group. **A4.** Comparisons of the wrong ratios between the ADHD and the HC group. **B1.** Comparisons of the wrong ratios for the third to fifth subtests of the colored fruit test between two groups. **B2.** Comparisons of the wrong scores for the third subtests of colored fruit test between two groups. Data are presented as mean ± SEM. * $p < 0.05$.

Next, we compared a total of 37 ET indicators (**Table S2**) of each test between two groups, and found that Targets without eye movements (TWOEM) was significantly decreased for all three tests in the ADHD group (**Fig. S6. A**), the Eye movements without mouse clicks (EMWOMC) was significantly increased for the number test in the ADHD group (**Fig. S6. B**), and the Minimum duration of fixations (MIDOF) was significantly decreased for the number test in the ADHD group (**Fig. S6. C**).

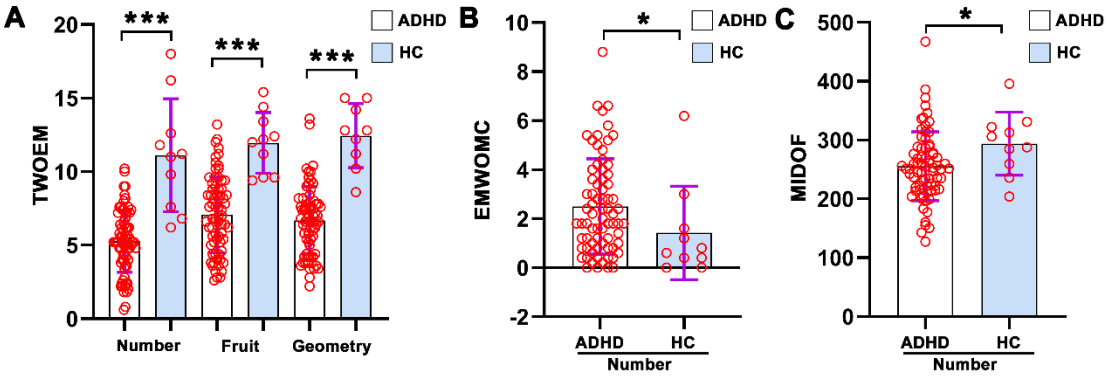

Figure S6. Attentive bias during three-cancellation testing in ADHD children. **A.** TWOEM was significantly decreased for all three tests in the ADHD group when compared to that of HC group. **B.** EMWOMC was significantly increased for the number test in the ADHD group when compared to that of HC group. **C.** MIDOF was significantly decreased for the number test in the ADHD group when compared to that of HC group. Data are presented as mean ± SEM. * $p < 0.05$, *** $p < 0.0001$.

We further compared the ET indicators of each of the five subtests of three tests between the two groups. The results showed that the third and fifth subtests in the ADHD group showed a significant decrease in TWOEM compared to the HC group; Also, the ADHD group showed a significant decrease in TWOEM in the fourth subtest of the colored fruit test compared to the HC group. In addition, there were significant differences in the fourth subtest of EMWOMC and Time from first fixation to click (TFFFTC), as well as the fifth subtest of Mouse clicks (MC), EMWOMC, Average fixation duration (AFD) and MIDOF between the two groups; The ADHD group had significant differences in AFD, Targets with eye movements (TWEM), and TFFFTC

of the third, fourth, and fifth subtests of the colored fruit test compared to the HC group; The ADHD group geometry test showed significant differences compared to the HC group in the first subtest and the second subtest of MC and MIDOF, and the fourth subtest of AFD, Average visit duration (AVD), Maximum duration of fixations (MADOF), Maximum duration of visit (MADOV) and Number of saccades in AOI (NOSIA) (see Additional file 3: **Table S6**).

Table S6. Attentive bias of five-subtests during three-cancellation testing in ADHD children

Cancellation test	Subtest-Variables	ADHD (n = 78)	HC (n = 11)	t[a] /Mann-Whitney U	*p*
Number	3-TWOEM	73(4.055±0.242)	10(18.000±0.919)	0.000	0.000
	4- EMWOMC	73(5.603±0.439)	10(3.200±0.929)	224.000	0.048
	4-TFFFTC	73(5.003±0.663)	10(2.035±0.240)	186.000	0.012
	5-TWOEM	73(3.863±0.278)	10(17.400±0.600)	0.000	0.000
	5-MC	73(8.685±0.365)	10(10.700±0.667)	1.977[a]	0.051
	5-EMWOMC	73(1.849±0.207)	10(0.400±0.267)	172.000	0.005
	5-AFD	73(0.377±0.015)	10(0.476±0.045)	221.000	0.044
	5-MIDOF	73(240.214±10.337)	10(300.922±34.096)	206.000	0.026
Colored fruit	3-TWOEM	69(4.362±0.294)	10(17.500±0.687)	0.000	0.000
	3-AFD	69(0.333±0.010)	10(0.378±0.014)	1.697 [a]	0.094
	4-TWEM	69(18.145±0.525)	10(21.400±1.002)	2.268 [a]	0.026
	4-TWOEM	69(9.855±0.525)	10(6.600±1.002)	-2.268 [a]	0.026
	5-TWOEM	69(4.609±0.334)	10(18.300±1.096)	0.000	0.000
	5-TFFFTC	67(7.256±1.833)	10(2.255±0.295)	150.000	0.005

Geometry	1-MC	69(21.406±0.535)	9(18.333±0.500)	179.500	0.040
	1-CC	69(23.174±0.627)	9(19.444±0.603)	-4.287 [a]	0.000
	2-MC	69(21.232±0.749)	9(18.889±1.230)	175.000	0.033
	2-CC	69(22.275±0.822)	9(19.667±1.374)	184.000	0.047
	2-MIDOF	69(259.471±9.002)	9(328.013±34.864)	2.465 [a]	0.016
	3-TWOEM	69(4.565±0.292)	9(17.778±0.596)	0.000	0.000
	4-AFD	69(0.311±0.008)	9(0.357±0.018)	2.104 [a]	0.039
	4-AVD	69(0.401±0.011)	9(0.487±0.027)	2.712 [a]	0.008
	4-MADOF	69(469.341±14.415)	9(576.689±49.385)	179.000	0.040
	4-MADOV	69(602.418±20.167)	9(753.119±53.955)	147.000	0.011
	4-ADOG	69(427.692±10.891)	9(511.851±26.629)	2.656 [a]	0.010
	4-MADOG	69(628.878±20.251)	9(769.607±56.006)	166.000	0.024
	4-NOSIA	69(8.101±0.474)	9(14.000±0.928)	4.434 [a]	0.000
	5-TWOEM	69(4.232±0.354)	9(19.222±0.813)	0.000	0.000

Based on the above results, a heatmap analysis of the effective target fixation duration was conducted by selecting the third and fifth subtests with significant differences in ET indicators. The results showed that in the third subtest, the ADHD group had shorter effective fixation time in the target area, was less limited to the target area, and was not widely distributed throughout the entire test area compared to HC group; The geometry test was more pronounced, indicating that ADHD's attention selectivity was limited, the time was short, and the distribution was uneven; The fifth subtest also found a similar situation, with the difference between the ADHD and HC groups being similar among the three tests, indicating that ADHD children have deficiencies in attention span and allocation ability. See (**Fig. S5**).

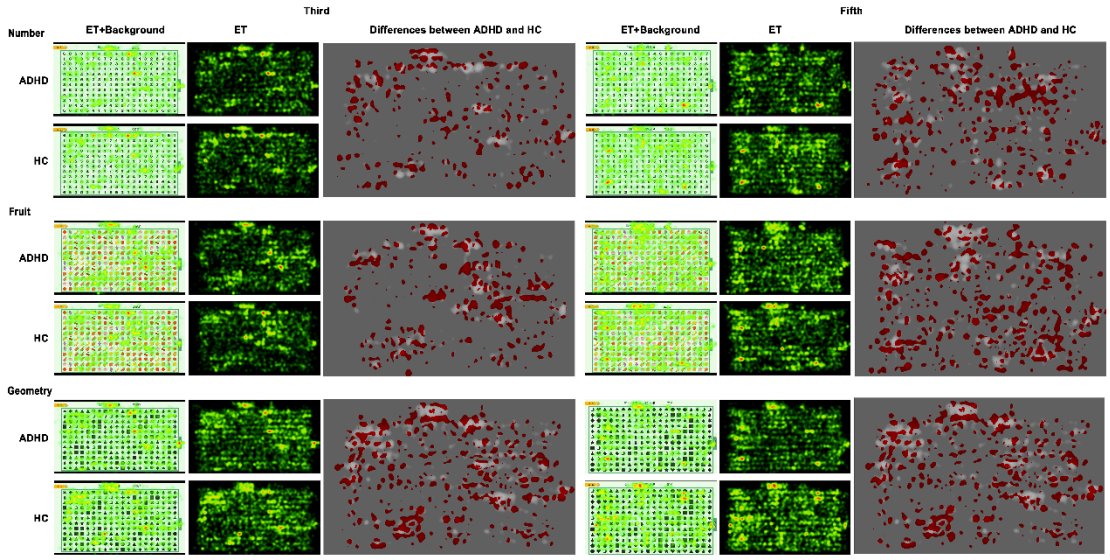

Figure S5. Heatmap analysis of the effective target fixation duration for the third and fifth subtests of the three cancellation tests in ADHD children. ET=eye tracking.

1.2 Correlation analyses of gender differences of attentive bias during three-cancellation testing in ADHD children

Firstly, a correlation analysis was conducted between the correct/wrong scores and ET indicators of three tests in ADHD children. It was found that the correlation was basically similar among three tests. The correct score was positively correlated with TWEM, MC, Mouse clicks without eye movements (MCWOEM), CC, and negatively correlated with TWOEM, EMWOMC, TFF, TTFC, TFFFTC, MADOF, MADOV, TTFV, Maximum duration of glance (MADOG), TTFG, TTENS, and TTEXS, while the wrong score was mostly opposite to the correct score. We also analyzed the consistency of the correlation among three tests, i.e., the correlation between the correct/wrong scores and ET indicators of one test and the other two tests. It was found that the consistency was high with the correlation coefficient of the main test was higher than that of the other two tests. See Additional file 5: **Fig. S6**.

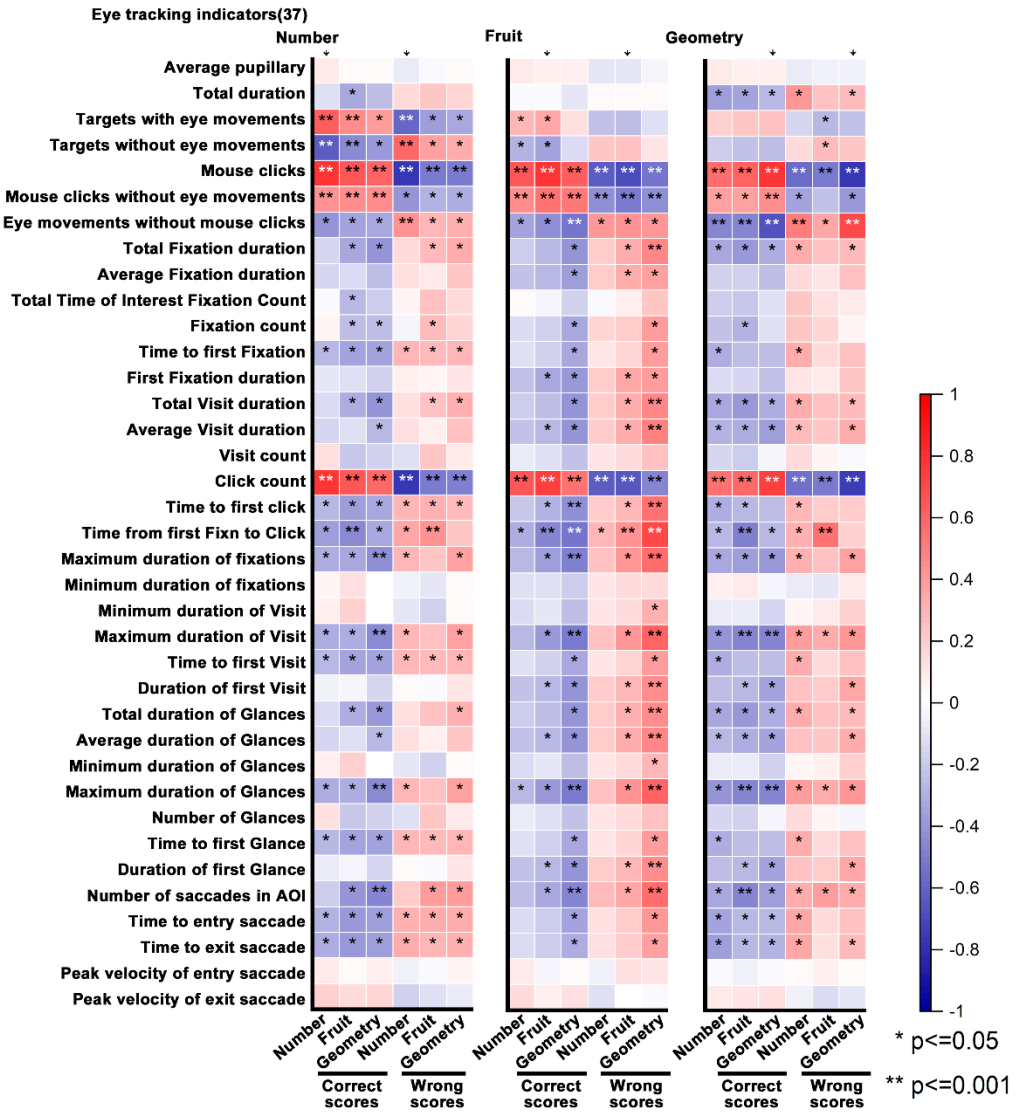

Figure S6. Correlations between the correct/wrong scores and ET indicators of three tests in ADHD children.

Secondly, we analyzed the correlation between the correct and wrong scores and their ET indicators of each subtest of three tests. It was found that correlation among the three tests were roughly the same: the correct score was positively correlated with TWEM, MC and CC, and negatively correlated with TWOEM, TFFFTC, etc., while the wrong score was the opposite.

There were many ET indicators negatively correlated with correct scores in the second subtest of the number test, including EMWOMC, TFF, AVD, TTFC, MADOF, MADOV, TTFV, Average duration of glances (ADOG), MADOG, TTFG, TTENS and TTEXS.

There were fewer ET indicators related to the first and second subtest scores of the colored fruit test, while there were more ET indicators related to the third, fourth, and fifth subtest scores; Surprisingly, the correct and wrong scores in the fourth subtest were significantly negatively correlated with AFD and First fixation duration (FFD);

The correct and wrong scores of the fifth subtest showed a significant negative correlation with FFD, while the correlation with AFD was not significant.

There were many ET indicators related to the first subtest of the geometry test. The correct score was positively correlated with MC, MCWOEM, CC and MIDOF, and negatively correlated with TD, EMWOMC, TTIFC, Fixation count (FC), TFF, VC, Time to first click (TTFC), TFFFTC, TTFV, NOG, TTFG, NOSIA, TTENS, TTEXS, etc.; The ET indicators related to the second, third, fourth and fifth subtests decreased, similar to the subtests of other tests. See Additional file 6: **Fig. S7**.

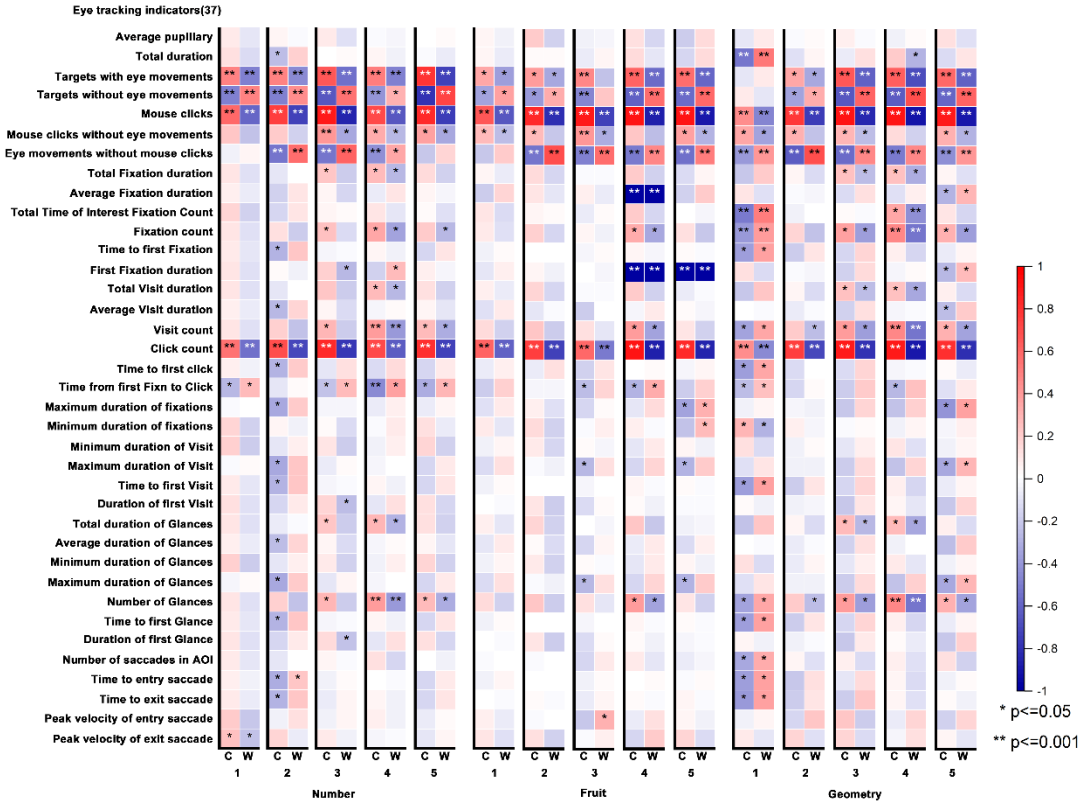

Figure S7. Correlations between the correct/wrong scores and ET indicators of each subtest of three tests in ADHD children.

We also analyzed the correlation between correct/wrong scores and ET indicators across of each subtest in male and female ADHD children:

The first subtest: The correct scores of the three tests for ADHD boys were significantly positively correlated with MC and CC, and there was no common negative correlation with ET indicators among the three tests; The number and fruit tests also showed a significantly positive correlation with TWEM, and a significantly negative correlation with TWOEM; The geometry test did not show a significant correlation with TWEM and TWOEM, but showed a significantly positive correlation with MCWOEM and MIDOF, and a significantly negative correlation with TD, EMWOMC, TTIFC, FC, TFF, VC, TTFC, TTFV, NOG, TTFG, NOSIA, TTENS and TTEXS. There was no common significantly related indicator among the three tests for ADHD girls; The number (TFFFTC negative correlation) and geometry (AP and Peak velocity of entry saccade (PVOENS) negative correlation) tests had few significant correlation indicators, while the colored fruit test had many correlation indicators, which were similar to that of ADHD boys. The correct scores in ADHD girls were positively

correlated with MC and CC, and also with MCWOEM; while negatively correlated with TFD, AFD, FFD, TVD, AVD, MADOF, MIDOF, MIDOV, MADOV, DOFV, TDOG, ADOG, MIDOG, MADOG, DOFG and NOSIA. The correlation of the wrong scores in ADHD girls was approximately opposite to the correct scores.

The second subtest: The correct scores of the three tests for ADHD boys were positively correlated with TWEM, MC and CC, and negatively correlated with EMWOEM; There were many indicators with significantly negative correlation in number test, such as TFF, AVD, TTFC, MADOF, MADOV, ADOG, MADOG, TTFG, TTENS and TTEXS. There was no common significant correlation index among the three tests for ADHD girls; Number tests were similar to that of the boys; Except for MC and CC; there were more significant negative correlation indicators (a total of 22 items) in the geometry test. The correlation of the wrong scores was approximately opposite to the correct scores.

The third subtest: The correct scores of the three tests for ADHD boys were positively correlated with TWEM, MC and CC, and negatively correlated with TWOEM and EMWOEM; The number test also showed a significantly positive correlation with MCWOEM, TFD, FC, TVD, VC, TDOG and NOG, and a significantly negative correlation with TFFFTC; The geometry test also showed a significantly positive correlation with TFD, FC, TVD, VC, TDOG, NOG, PVOENS and Peak velocity of exit saccade (PVOEXS). There was an imbalance among the three tests for ADHD girls, with an increase in significant correlation indicators for colored fruits and geometry tests; The number test for ADHD girls was similar to that of ADHD boys; In addition to TWOEM being the same as that of ADHD boys and EMWOMC disappeared, 9 other negative correlation indicators such as AVD, MADOF, MADOV, DOFV, ADOG, MADOG, DOFG, NOSIA and PVOENS were added to the colored fruit test; The geometry test had added positive correlation indicators (TTEXS) and 9 negative correlation indicators such as AFD, AVD, MADOF, MIDOV, MADOV, ADOG, MIDOG, MADOG and PVOEXS. The correlations of the wrong scores were approximately opposite to the correct scores.

The fourth subtest: In addition to a significantly positive correlation with TWEM, MC and CC, the correct scores of the three tests for ADHD boys also added significantly positive correlation indicators such as TFD, TTIFC, FC, TVD, VC, TDOG and NOG, while showing a significantly negative correlation with TWOEM, TFFFTC and EMWOEM; Colored fruits also showed a significantly negative correlation with AFD and FFD. There was an imbalance among the three tests in ADHD girls, with number and geometry similar to those in ADHD boys, while colored fruits show an increased number of negative correlation indicators; The positive correlation indicators of number test in ADHD girls decreased compared to ADHD boys(only MC and CC were the same as ADHD boys, adding MCWOEM), while the negative correlation indicators disappearing (TWOEEM), adding PVOENS and PVOEXS; In addition to AFD and FFD being the same as ADHD boys, TWOEM disappeared, and five other negative correlation indicators such as AVD, MADOF, MADOV, ADOG and

MADOG were added for the colored fruit test; The geometry test was similar to that of ADHD boys, with 6 positive correlation indicators disappeared (including TFD, FC, TVD, VC, TDOG and NOG), and a negative correlation indicator (MCWOEM) added. Most correlations of the wrong scores were opposite to correct scores; However, there was a significantly negative correlation between the correct/wrong scores and AFD/FFD of both genders in the colored fruit test.

The fifth subtest: The correct scores of the three tests for ADHD boys were positively correlated with TWEM, MC and CC, and negatively correlated with TWOEM; Colored fruits also showed a significant negative correlation with FFD. There was an imbalance among the three tests in ADHD girls, with number and colored fruits similar to those in boys, while geometry showed an increased numbers of negative correlation indicators; The positive correlation indicators of ADHD girls of number tests decreased compared to ADHD boys (only TWEM, MC and CC were the same as ADHD boys), while the negative correlation indicators decreased TFFFTC; The significantly negative correlation indicators in the colored fruit test disappeared except for EMWOMC and FFD, which were the same as those in ADHD boys; In addition to the significantly positive correlation between TWEM, MC and CC, the geometry test decreased by 3 indicators and added 12 negative correlation indicators (AFD, FFD, AVD, TFFFTC, MADOF, MIDOV, MADOV, DOFV, ADOG, MIDOG, MADOG, MADOG and DOFG). Most correlations of the wrong scores were opposite to correct scores; However, there was a significantly negative correlation between the correct/wrong scores and FFD of both genders in the colored fruit test. See **Fig. S8**.

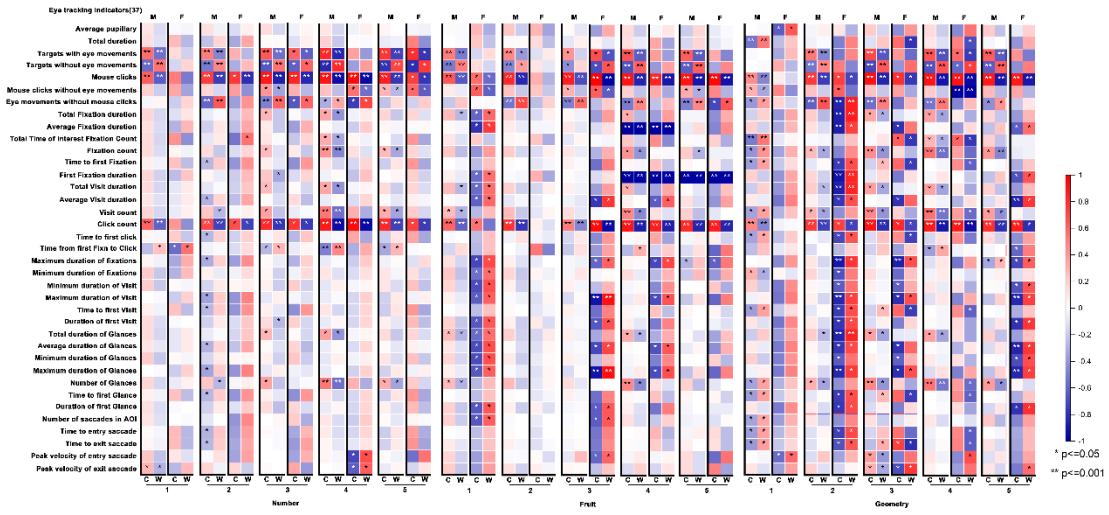

Figure S8. Correlations between the correct/wrong scores and ET indicators of five subtest of three cancellation tests in male/female ADHD children.

1.3 Relevant factors affecting gender differences of attentive bias during three-cancellation testing in ADHD children

A correlation analysis was conducted between the correct/wrong scores of three tests and clinical scale evaluation indicators in ADHD children. It was found that only in number test did PSQ-behavior show a significantly positive correlation with the correct scores. We also analyzed the correlation between correct/wrong scores across of each subtest and clinical scale evaluation indicators in ADHD children. The correct scores were found to be positively correlated with SNAP-IV-AD (number-fifth subtest), but negatively correlated with PSQ-anxiety (number-third subtest), ADHD rating score (fruit-first subtest), EF-organization of materials (fruit-second subtest), PSQ-behavior (geometry-first subtest), EF-inhibition (geometry -fifth subtest), EF-emotional control (geometry -fifth subtest), and EF-BRI (geometry-fifth subtest); Wrong scores were the opposite of correct scores; But it increased the positive correlation with PSQ behavior (number-first subtest), decreased the negative correlation with SNAP-IV-AD (number-first subtest) and the positive correlation with EF-organization of materials (fruit-second subtest), and increased the negative correlation with SNAP-IV-AD (fruit-third subtest). See Additional file 7: **Fig. S9**.

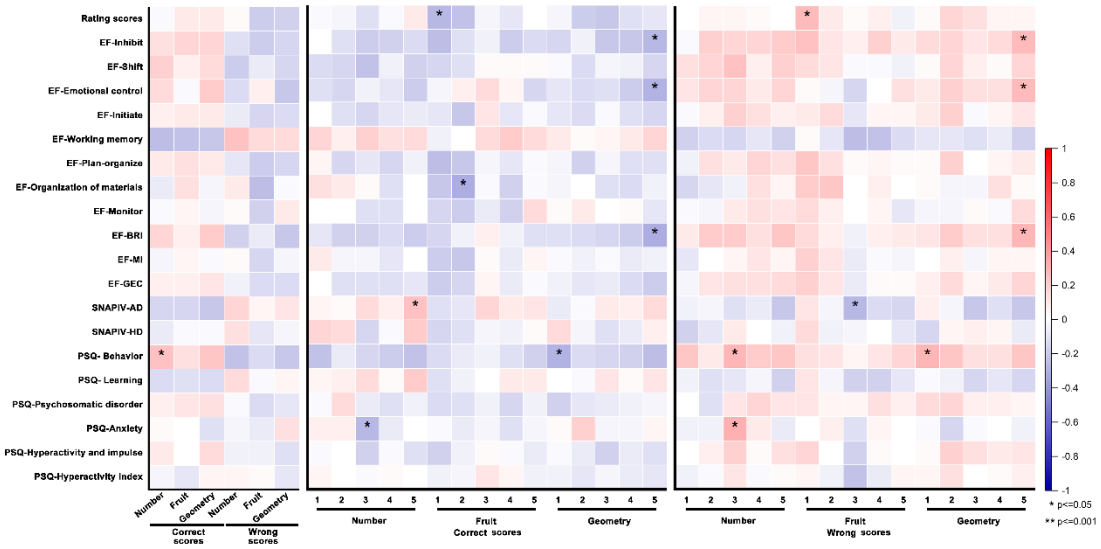

Figure S9. Correlations between the correct/wrong scores and clinical scale evaluation indicators in ADHD children.

Then, we conducted a correlation analysis between the correct/wrong scores of three tests and clinical scale evaluation indicators in male and female ADHD children. It was found that there were differences between the males and the females. There was a significant positive correlation between correct scores and PSQ-behavior (number) and PSQ-behavior (geometry) in male patients; There was a significant negative correlation between wrong scores and PSQ-behavior (number), EF-monitor (fruit) (Additional file 8: **Fig. S10**). However, no significantly correlated indicators were found in female patients (Fig.13). We also conducted a correlation analysis between the correct/wrong scores of each subtest and clinical scale evaluation indicators. It was found that the correct scores of ADHD boys were negatively correlated with multiple clinical scale evaluation indicators, such as EF-inhibit (number-third subtest), EF-shift (number-third subtest), EF-BRI (number-third subtest), PSQ-behavior (number-fifth subtest), ADHD-

rating score (fruit-first subtest), EF-inhibit (fruit-first subtest), EF-plan-organize (fruit-first subtest), ADHD-rating score (fruit-second subtest), EF-organization of materials (fruit-second subtest), PSQ-behavior (geometry-first subtest), EF-emotional control (geometry-fifth subtest), and EF-BRI (geometry-fifth subtest); while no positive correlation indicators were found in ADHD boys. Wrong score was positively correlated with many indicators, in addition to indicators that were opposite to correct score, there were still differences (mainly in the Number-third and fifth subtests), such as the addition of EF-initiate (number-third subtest), PSQ-behavior (number-third subtest), PSQ-anxiety (number-third subtest), EF-initiate (number-fifth subtest), and EF-BRI (number-fifth subtest); No significantly negative correlation indicators were found in ADHD boys (see Additional file 8: **Fig. S10**). Compared to males, females had both positive and negative correlations with correct/wrong scores; The correct score was negatively correlated with PSQ-psychosomatic disorder (number-fourth subtest), positively correlated with EF-initiate (fruit-second subtest), positively correlated with PSQ-behavior (fruit-second subtest), and positively correlated with PSQ-hyperactivity (fruit-second subtest); In addition to being opposite to correct scores, wrong scores were also negatively correlated with SNAP-IV-AD (geometry-fourth subtest) and positively correlated with PSQ psychosomatic disorders (geometry-fifth subtest) (see Additional file 9: **Fig. S11**).

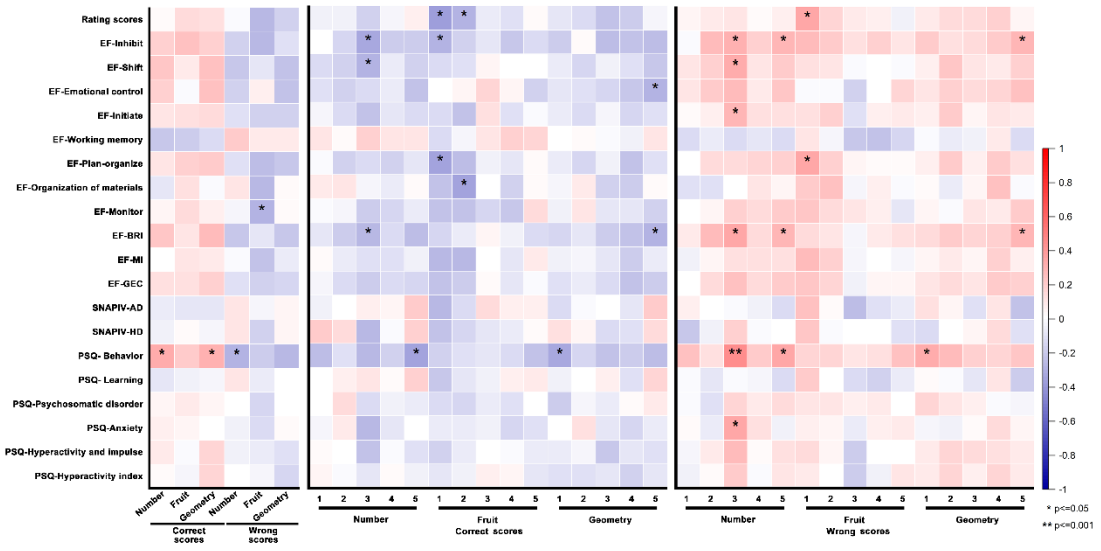

Figure S10. Correlations between the correct/wrong scores and clinical scale evaluation indicators in male ADHD children.

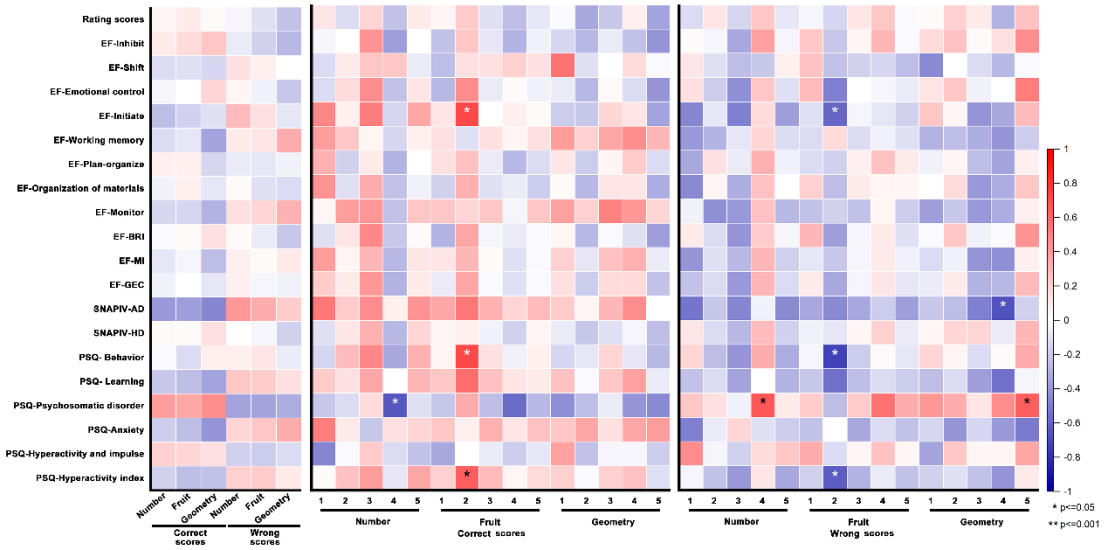

Figure S11. Correlations between the correct/wrong scores and clinical scale evaluation indicators in female ADHD children.

1.4 Relevant factors affecting gender differences of attentive bias during three-cancellation testing in ADHD children by stepwise linear regressions

We conducted a stepwise linear regression analysis on the correct and wrong scores of three tests for male ADHD children using ET and clinical scale indicators. Number Test: for the correct score, CC (β= 0.536) and TWEM (β= 0.211 entered the regression model, while colored fruit MC also entered the model; for wrong score, CC (β= -0.617), TTIFC (β= -0.252), TWOEM (β= -0.683) and VC (β= 0.216) entered the model, while the geometry CC and TTENS, as well as EF-initiate and EF-organization of materials, also entered the model. Colored fruit test: for the correct score, MC (β= 0.709) and TTFG (β= 0.185) entered the regression model, while number MC, TTIFC and Peak velocity of entry saccade (PVOENS), and PSQ-learning also entered the model (consistent with the overall sample); for the wrong score, MC (β= -0.452) entered the model, while the number CC and geometry NOSIA also entered the model (consistent with the overall sample). Geometry test: for the correct score, MC (β= 0.550) entered the model, while number CC, VC and PVOENS, as well as fruit PVOENS, also entered the model; for the wrong score, MC (β= -0.327) and EMWOMC (β= 0.412) entered the model, while number CC and NOSIA also entered the model. See **Table S7**.

Table S7

Multiple linear regression analysis of correct and wrong scores of three-cancellation tests using ET and clinical scale indicators in children with ADHD

		Variable	B	Standard error	β	t	P
Number	Correct scores	Constant	9.703	6.676		1.453	0.153
		N-MC	2.684	0.236	0.675	11.351	0.000
		F-MC	2.041	0.299	0.427	6.831	0.000
		N-VC	0.287	0.109	0.156	2.625	0.012
		G-TTIFC	-0.018	0.006	-0.191	-3.014	0.004
		N-PVOENS	0.043	0.021	0.127	2.042	0.047
	Wrong scores	Constant	112.875	7.306		15.450	0.000
		N-TWOEM	-3.048	0.217	-0.683	-14.042	0.000
		G-CC	-1.758	0.270	-0.362	-6.510	0.000
		SNAPIV-HD	4.933	1.157	0.209	4.265	0.000
		EF-initiate	-8.982	2.033	-0.243	-4.418	0.000
		N-VC	-0.595	0.116	-0.288	-5.144	0.000
		G-TTIFC	0.013	0.006	0.122	2.340	0.024
		F-TFFFTC	0.787	0.372	0.120	2.115	0.041
Colored fruit	Correct scores	Constant	25.740	8.066		3.191	0.003
		F-MC	3.204	0.322	0.668	9.934	0.000
		N-MC	1.575	0.262	0.395	6.018	0.000
		PSQ-learning	-4.527	1.338	-0.223	-3.384	0.002
		N-TTIFC	-.032	0.008	-0.322	-4.150	0.000

		F-TTFG	0.000	0.000	0.164	2.298	0.027
		N-PVOENS	0.049	0.024	0.145	2.032	0.049
	Wrong	Constant	87.189	7.784		11.201	0.000
	scores	**F-MC**	-2.440	0.368	-0.509	-6.633	0.000
		N-CC	-1.752	0.295	-0.457	-5.939	0.000
		G-NOSIA	1.203	0.393	0.237	3.061	0.004
Geometry	Correct	Constant	10.156	8.722		1.164	0.251
	scores	**G-MC**	3.272	0.417	0.813	7.843	0.000
		N-CC	1.854	0.234	0.547	7.934	0.000
		N-FC	-0.367	0.096	-0.284	-3.809	0.000
		N-PVOENS	0.059	0.023	0.195	2.516	0.016
		G-EMWOMC	1.739	0.696	0.257	2.499	0.016
	Wrong	Constant	96.832	9.385		10.317	0.000
	scores	**G-MC**	-3.289	0.424	-0.601	-7.748	0.000
		N-MC	-2.127	0.219	-0.444	-9.711	0.000
		N-NOSIA	1.734	0.290	0.341	5.981	0.000
		G-EMWOMC	2.146	0.735	0.234	2.922	0.006
		G-MIDOF	0.142	0.034	0.521	4.155	0.000
		G-MCWOEM	1.454	0.637	0.129	2.283	0.028

G-MIDOV	-0.087	0.027	-0.417	-3.263	0.002
N-TTIFC	-0.020	0.007	-0.164	-2.899	0.006

Nest, we conducted a stepwise linear regression analysis on the correct and wrong scores of three tests for female ADHD children using ET and clinical scale indicators. Number test: for the correct score, fruit MCWOEM and TTFC, SNAPIV-HD and EF-emotional control entered the regression model; for the wrong score, fruit MCWOEM, EF-BRI, and SNAPIV-HD entered the model. Colored fruit test: for the correct score, FFD (β= 0.222) and TTIFC (β= 0.053) entered the model, while the number TTIFC, DOFV, and PVOENS, as well as the geometry MADOG and TFD PVOENS, also entered the model; for the wrong score, AFD (β= -0.215) entered the model, while number TTIFC and DOFG, geometry MADOG, PVOENS, and AP also entered the model. Geometry test: for the correct score, AFD (β= -0.831), TTENS (β= -0.323) and TFD (β= 0.065) entered the model, while number AP, fruit Minimum duration of fixations (MIDOF) and CC, and EF-GEC also entered the model; for the wrong score, AVD (β= -0.737), CC (β= -0.281), MIDOF (β= 0.154), and TTFC (β= 0.035) entered the model, while number TWOEM, fruit ADOG, and EF-GEC also entered the model. See **Table S8**.

Table S8

Multiple linear regression analysis of correct and wrong scores of three-cancellation tests using ET and clinical scale indicators in male ADHD children

		Variable	B	Standard error	β	t	*P*
Number	Correct	Constant	-0.224	6.659		-0.034	0.973
	scores	**N-CC**	2.216	0.356	0.536	6.221	0.000
		F-MC	2.092	0.345	0.424	6.058	0.000
		N-TWEM	1.202	0.520	0.211	2.309	0.027
	Wrong	Constant	104.748	13.115		7.987	0.000
	scores	**N-CC**	-2.887	0.362	-0.617	-7.977	0.000

		G-CC	-1.908	0.367	-0.391	-5.202	0.000
		N-TTFC	-0.366	0.098	-0.252	-3.755	0.001
		EF-initiate	-10.464	2.259	-0.301	-4.631	0.000
		EF-organization of materials	5.267	2.035	0.152	2.588	0.014
		N-TWOEM	1.389	0.518	0.216	2.681	0.012
		G-TTENS	0.000	0.000	0.143	2.139	0.040
Color	Correct	Constant	28.344	8.331		3.402	0.002
Fruit	scores	**F-MC**	3.551	0.359	0.709	9.880	0.000
		N-MC	1.527	0.296	0.358	5.155	0.000
		PSQ- learning	-5.402	1.476	-0.258	-3.660	0.001
		N-TTIFC	-0.042	0.008	-0.393	-4.994	0.000
		F-TTFG	0.000	0.000	0.185	2.512	0.017
		N-PVOENS	0.051	0.024	0.159	2.139	0.040
	Wrong	Constant	87.186	8.761		9.952	0.000
	scores	N-CC	-2.028	0.364	-0.498	-5.578	0.000
		F-MC	-2.194	0.439	-0.452	-4.993	0.000
		G-NOSIA	1.392	0.459	0.269	3.035	0.004
Geometry	Correct	Constant	23.986	6.182		3.880	0.000
	scores	**G-MC**	2.206	0.305	0.550	7.245	0.000
		N-CC	2.147	0.260	0.599	8.257	0.000

	N-VC	-0.491	0.113	-0.319	-4.359	0.000
	N-PVOEXS	0.139	0.040	0.479	3.455	0.001
	F-PVOEXS	-0.072	0.035	-0.282	-2.045	0.049
Wrong	Constant	80.698	10.617		7.601	0.000
scores	**G-MC**	-1.784	0.564	-0.327	-3.164	0.003
	N-CC	-2.474	0.294	-0.507	-8.406	0.000
	G-EMWOMC	3.671	0.916	0.412	4.006	0.000
	N-NOSIA	0.873	0.307	0.170	2.841	0.007

Discussion

1.5 Differences between ADHD and HC children

AD is one of the core symptoms of ADHD, and our results of three cancellation tests showed that children with ADHD had poorer attention than the HC [36, 37]. This difference is more significant in colored fruit test, indicating that colored fruits, compared to monotonous numbers and graphics, could more arouse the interest of HC children and promote their concentration. In contrast, this stimulating effect on ADHD children may not be enough to offset the difficulties caused by their own AD. It is also possible that there is a color perception defect in ADHD patients[38, 39], which is caused by impaired CNS and retinal dopamine function [40]. Previous studies demonstrated that color allocated attention [41] was significantly decreased in ADHD children [42], and the color tracking test was a promising tool for diagnosing AD [43]. Our study also suggests that the colored fruit cancellation test has a higher sensitivity to AD detection, which is consistent with the previous reports [37, 44].

This study also found significant differences in ET indicators between ADHD children and HC children of the three cancellation tests. The TWOEM of the ADHD group was significantly decreased, indicating that ADHD children could see more goals and had a sufficient eye scanning. The EMWOMC (number test) of the ADHD group was significantly higher than that of the HC group, and the ADHD group had fewer TWOEMs in the third and fifth subtests of the number, colored fruit, and geometry tests compared to the HC group. We speculate that ADHD children had sufficient eye scanning during the cancellation test, but due to lack of concentration, they did not engage in mouse clicking on the target area. It has also been indirectly

verified that there are certain obstacles to EF in ADHD children. Furthermore, the ADHD group had less MIDOF (number test) than the HC group, indicating a weakened checking mechanism and impulsive in ADHD children. The heatmap analysis of the effective target fixation duration also indicated that the selectivity of attention in ADHD was limited, including the time was short, and the distribution was uneven, and there were shortcomings in the breadth and distribution of attention. This provides certain guidance for attention training for children with ADHD, such as strengthening the duration of attention to goals and expanding exploration of regional scope.

1.6 The correlations between the results of cancellation test and ET indicators in children with ADHD

There is a correlation between the results of ADHD children's cancellation test and ET indicators. We found that both whole-test and subtest results showed a positive correlation between correct scores and TWEM, MC, and CC, and a negative correlation between correct scores and TWOEM and TFFFTC, while the wrong scores were the opposite; The longer the time from first fixation to click (TFFFTC), the slower the reaction and the lower the correct score. EMWOMC means that the more eyes see but miss points, the lower the correct score, and it has good correlation in all three tests and their subtests. The longer the MADOF and MADOV, the longer the maximum duration of fixation and access in the target area, and the longer the state of confusion, the lower the correct score in the total results of the three tests, which is also reflected in the fifth subtest of the number and geometry tests. Although there are significant differences among the three tests, the analysis of the correlation in this study shows a good consistency. At the same time, each test has its own characteristics in the correlation between ET indicators and the test results. In terms of directionality and concentration of attention, the more FC and VC, the higher the correct score after repeated confirmation by the subject. In terms of attention, there is more correlation between the total score and the ET indicators of the geometry test; In terms of attention transfer, there is more correlation between the total score and the ET indicators of the number test; In terms of attention selection, the longer the DOFV and FFD, the lower the wrong score of the number test for subjects who observed more carefully at first glance. In terms of attention span, the longer TVD and TD, the longer the subject observes the target, and the lower the total wrong score of the geometry test. In terms of attention allocation, the longer the AFD, the more “free-time” it is speculated to have for the identified goals, and the lower the correct total score on the geometry test.

1.7 The correlations between the results of cancellation test and clinical scale indicators in children with ADHD

There is a certain correlation between the results of the ADHD children's cancellation tests and the clinical scale evaluation results, and there are many factors that are negatively correlated with the correct scores, such as PSQ-anxiety (number third

subtest), ADHD rating score (fruit first subtest), EF-organization of materials (fruit second subtest), PSQ-behavior (geometry first subtest), EF-inhibit (geometry fifth subtest), EF-emotional control (geometry fifth subtest) and EF-BRI (geometry fifth subtest); Among them, EF is the most closely related, indicating that the attention testing process itself cannot be separated from EF. In brain dynamic reconstruction, EF needs to be supported by flexibility[39], while children with ADHD often have significant EF deficit, insufficient flexibility and adaptability, and more obvious attention problems. The results of multiple linear regression analysis also indicate that the results of the colored fruit test are also related to factors such as EF- plan-organize and EF-monitor.

2. Figure S1.

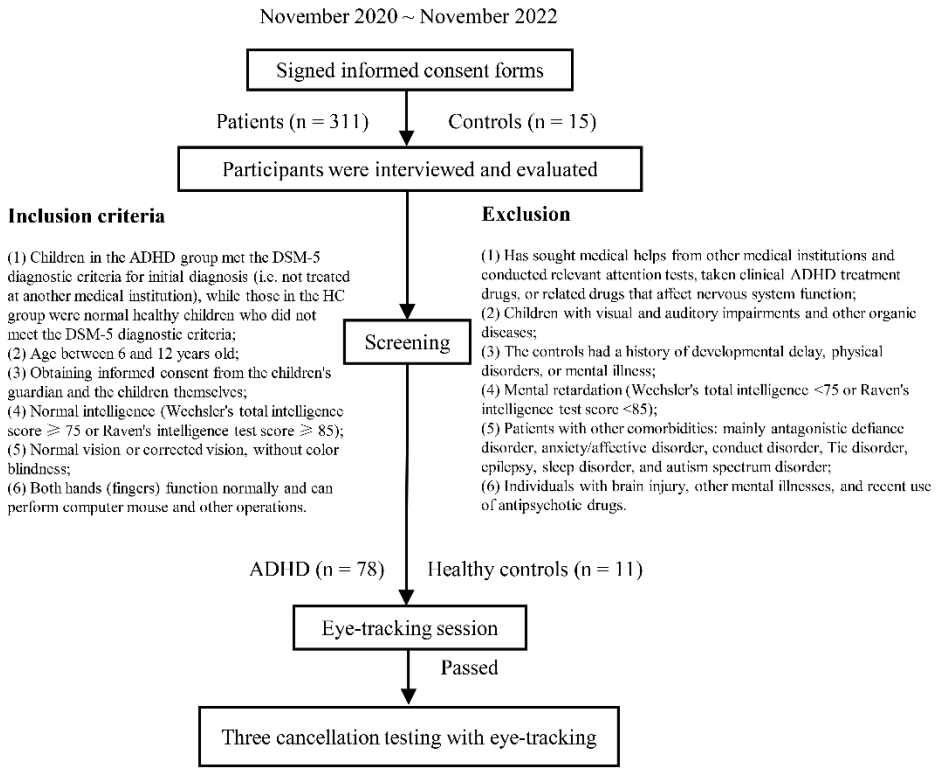

Figure S1. Study flowchart of three cancellation testing with ET in children with or without ADHD

3. Figure S2.

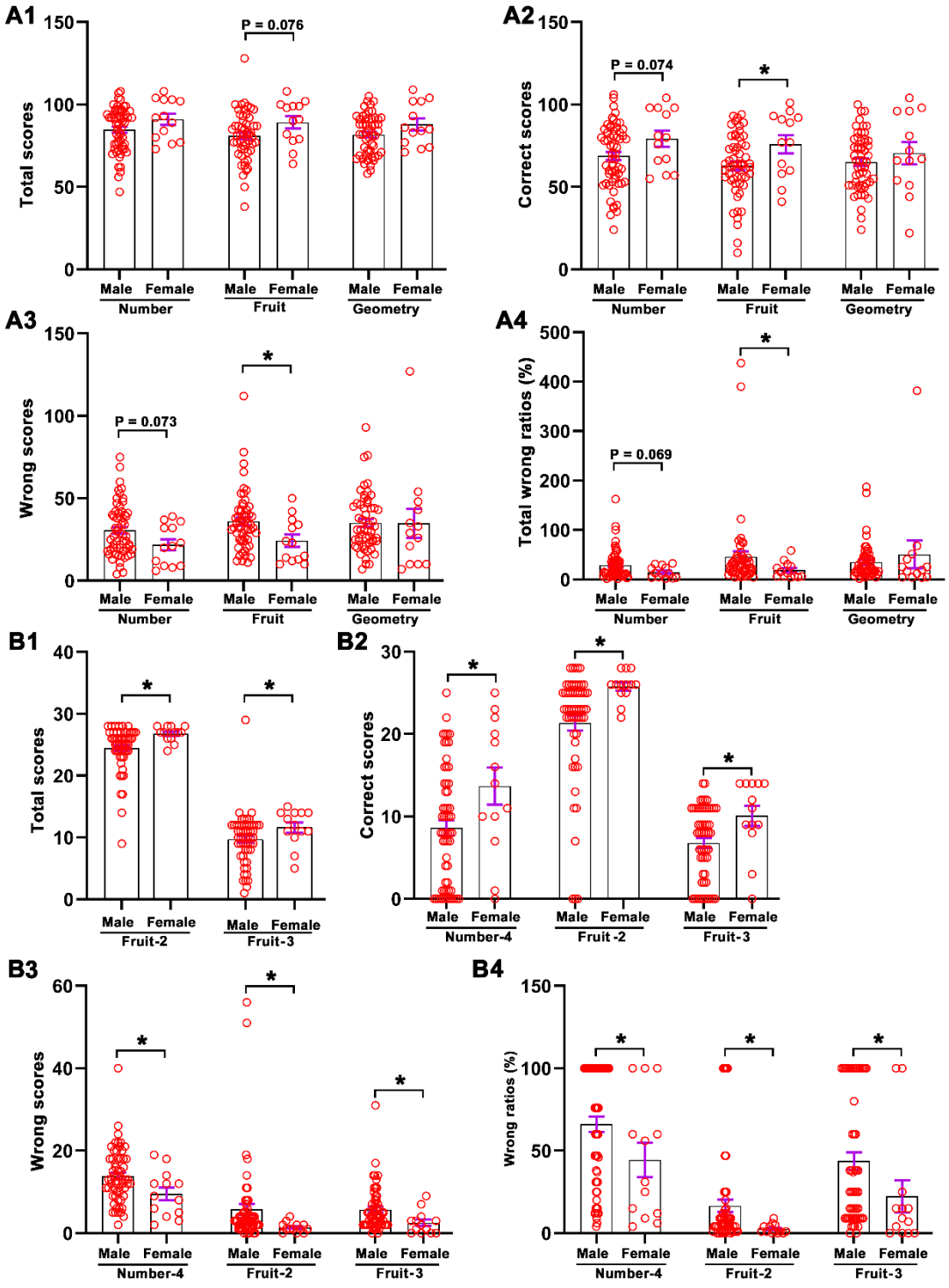

Figure S2. Results of three cancellation tests in male and female ADHD children. **A1.** Comparisons of the total scores between the male and female ADHD children. **A2.** Comparisons of the correct scores between the male and female ADHD children. **A3.** Comparisons of the wrong scores between the male and female ADHD children. **A4.** Comparisons of the total scores between the male and female ADHD children. **B1.** Comparisons of the total scores of the colored-fruit second and third subtests between the male and female ADHD children. **B2.** Comparisons of the correct scores of the number fourth subtest, and colored-fruit second and third subtests between the male

and female ADHD children. **B3.** Comparisons of the wrong scores of the number fourth subtest, and colored-fruit second and third subtests between the male and female ADHD children. **B4.** Comparisons of the wrong ratios of the number fourth subtest, and colored-fruit second and third subtests between the male and female ADHD children. Data are presented as mean ± SEM. * $p < 0.05$.

4. Figure S3.

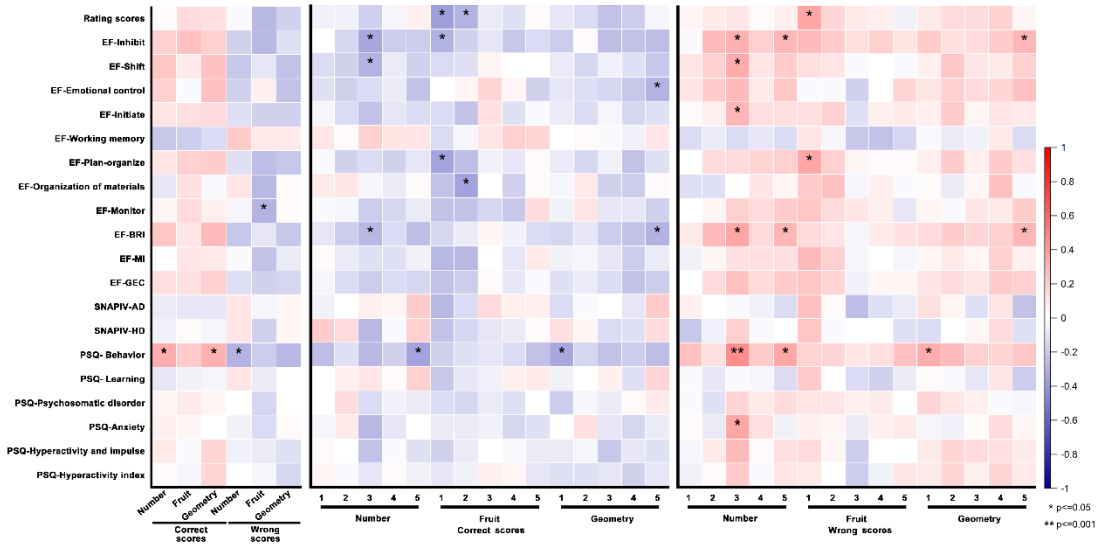

Figure S3. Correlations between the correct/wrong scores and clinical scale evaluation indicators in male ADHD children.

5. Figure S4.

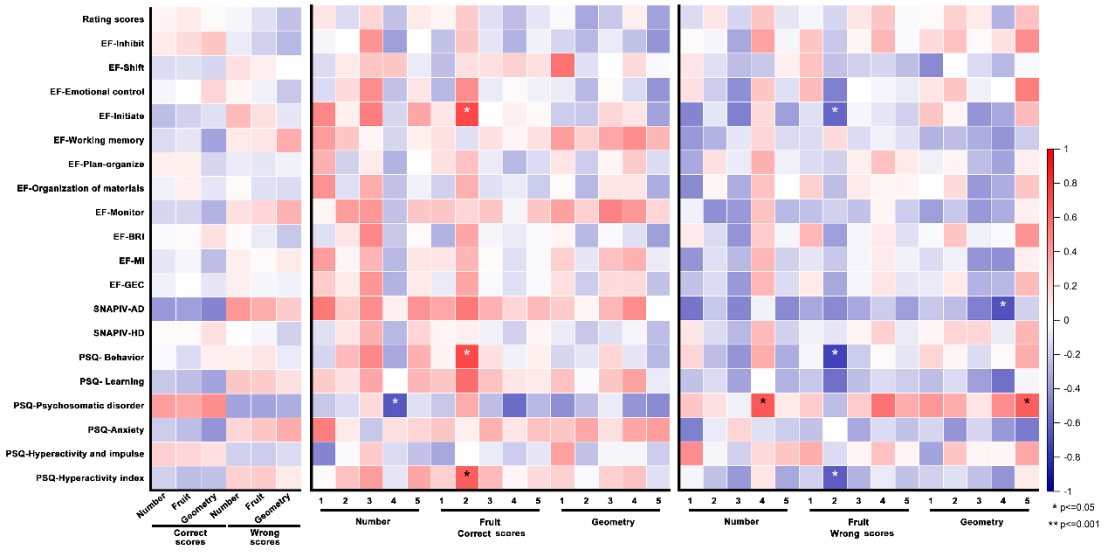

Figure S4. Correlations between the correct/wrong scores and clinical scale evaluation indicators in female ADHD children.

6. Table S1.

Demographic and clinical characteristics

Variables	ADHD (n = 78)	HC (n = 11)	t/χ^2/Mann- Whitney U	*p*
Age (years)	8.192±1.359	8.273±1.191	0.186	0.853
Gender (M/F), n	64/14	9/2	0.000	0.985
DBDRS-AD scores	7.795±1.121	4.3636±0.809	9.779(df=87)	0.000
DBDRS-HD scores	6.423±2.123	2.9091±2.023	5.167(df =87)	0.000
Rating scores	48.295±9.563	/		

7. Table S2.

ET indicators

ET indicators	Abbreviations	Annotations
Average pupillary	AP	The average pupil diameter of all fixation points during the whole testing
Total duration	TD	Time interval for staying on the testing interface
Targets with eye movements	TWEM	Number of target points with eye movement data in AOI
Targets without eye movements	TWOEM	Number of target points without eye movement data in AOI
Mouse clicks	MC	Number of mouse clicks in AOI
Mouse clicks without eye movements	MCWOEM	Have mouse clicks but no eye movements in AOI
Eye movements without mouse clicks	EMWOMC	With eye movements but no mouse clicks in AOI
Total fixation duration	TFD	Duration of fixation points in AOI
Average fixation duration	AFD	The average duration of fixation points in AOI
Total time of interest fixation count	TTIFC	The total time spent on interest fixation points during the testing
Fixation count	FC	The total number of fixation points in AOI
Time to first fixation	TFF	The time from the first fixation point in AOI to start testing
First fixation duration	FFD	Duration of the first fixation point in AOI
Total visit duration	TVD	Total duration of AOI vist

Average visit duration	AVD	Average AOI visit time
Visit count	VC	The total number of visits in AOI
Click count	CC	The total number of mouse clicks in AOI
Time to first click	TTFC	The time from the first mouse click in AOI to start testing
Time from first fixation to click	TFFFTC	The time interval during which the first mouse click is made within the AOI
Maximum duration of fixations	MADOF	Maximum fixation duration within AOI
Minimum duration of fixations	MIDOF	Minimum fixation duration within AOI
Minimum duration of visit	MIDOV	Minimum value of visits duration within AOI
Maximum duration of visit	MADOV	Maximum duration of visits within AOI
Time to first visit	TTFV	The time between the first visit and the start of testing in AOI
Duration of first visit	DOFV	Duration of first visit in AOI
Total duration of glances	TDOG	Total duration of glances in AOI
Average duration of glances	ADOG	Average duration of glances in AOI
Minimum duration of glances	MIDOG	Minimum duration of glances within AOI
Maximum duration of glances	MADOG	Maximum duration of glances within AOI
Number of glances	NOG	The total number of glances in AOI
Time to first glance	TTFG	The time when the first glance in AOI starts testing

Duration of first glance	DOFG	Duration of the first glance in AOI
Number of saccades in AOI	NOSIA	The total number of saccades in AOI
Time to entry saccade	TTENS	The time when saccade enters in AOI
Time to exit saccade	TTEXS	Time at which saccades exit in AOI
Peak velocity of entry saccade	PVOENS	The fastest speed for entry saccade in AOI
Peak velocity of exit saccade	PVOEXS	The fastest speed for exit saccade in AOI

In the AOI, where the correct target is located, the entry and exit of the fixation point is a visit. Glance includes information about the first saccade, not including visits. AOI: Area of Interest.

7. Table S3.

Multiple linear regression analysis of correct and wrong scores of three-cancellation tests using ET and clinical scale indicators in male ADHD children

		Variable	B	Standard error	β	t	P
Number	Correct	Constant	-0.224	6.659		-0.034	0.973
	scores	**N-CC**	2.216	0.356	0.536	6.221	0.000
		F-MC	2.092	0.345	0.424	6.058	0.000
		N-TWEM	1.202	0.520	0.211	2.309	0.027
	Wrong	Constant	104.748	13.115		7.987	0.000
	scores	**N-CC**	-2.887	0.362	-0.617	-7.977	0.000

		G-CC	-1.908	0.367	-0.391	-5.202	0.000
		N-TTFC	-0.366	0.098	-0.252	-3.755	0.001
		EF-initiate	-10.464	2.259	-0.301	-4.631	0.000
		EF-organization of materials	5.267	2.035	0.152	2.588	0.014
		N-TWOEM	1.389	0.518	0.216	2.681	0.012
		G-TTENS	0.000	0.000	0.143	2.139	0.040
Color	Correct	Constant	28.344	8.331		3.402	0.002
Fruit	scores	**F-MC**	3.551	0.359	0.709	9.880	0.000
		N-MC	1.527	0.296	0.358	5.155	0.000
		PSQ- learning	-5.402	1.476	-0.258	-3.660	0.001
		N-TTIFC	-0.042	0.008	-0.393	-4.994	0.000
		F-TTFG	0.000	0.000	0.185	2.512	0.017
		N-PVOENS	0.051	0.024	0.159	2.139	0.040
	Wrong	Constant	87.186	8.761		9.952	0.000
	scores	N-CC	-2.028	0.364	-0.498	-5.578	0.000
		F-MC	-2.194	0.439	-0.452	-4.993	0.000
		G-NOSIA	1.392	0.459	0.269	3.035	0.004
Geometry	Correct	Constant	23.986	6.182		3.880	0.000
	scores	**G-MC**	2.206	0.305	0.550	7.245	0.000
		N-CC	2.147	0.260	0.599	8.257	0.000

		N-VC	-0.491	0.113	-0.319	-4.359	0.000
		N-PVOEXS	0.139	0.040	0.479	3.455	0.001
		F-PVOEXS	-0.072	0.035	-0.282	-2.045	0.049
	Wrong	Constant	80.698	10.617		7.601	0.000
	scores	**G-MC**	-1.784	0.564	-0.327	-3.164	0.003
		N-CC	-2.474	0.294	-0.507	-8.406	0.000
		G-EMWOMC	3.671	0.916	0.412	4.006	0.000
		N-NOSIA	0.873	0.307	0.170	2.841	0.007

8. Table S4.

Multiple linear regression analysis of correct and wrong scores of three-cancellation tests using ET and clinical scale indicators in female ADHD children

		Variable	B	Standard error	*β*	t	*P*
Number	Correct	Constant	76.597	0.620		123.576	0.000
	scores	F-MCWOEM	5.910	0.087	1.051	68.145	0.000
		SNAPIV-HD	9.521	0.336	0.437	28.328	0.000
		F-TTFC	-0.194	0.015	-0.199	-12.711	0.000
		EF-emotional control	-2.951	0.512	-0.098	-5.762	0.005
	Wrong	Constant	41.300	2.032		20.330	0.000
	scores	F-MCWOEM	-6.626	0.357	-1.131	-18.559	0.000

		SNAPIV-HD	-11.466	1.680	-0.505	-6.825	0.001
		EF-BRI	10.398	2.515	0.279	4.134	0.009
Colored Fruit	Correct scores	Constant	120.816	0.005		25265.365	0.000
		G-MADOG	-0.069	0.000	-1.176	-14113.506	0.000
		N-TTIFC	-0.025	0.000	-0.343	-5353.519	0.000
		G-PVOENS	0.096	0.000	0.216	4452.678	0.000
		F-FFD	35.344	0.006	0.222	6353.935	0.000
		F-TTIFC	0.004	0.000	0.053	1987.390	0.000
		N-DOFV	-0.003	0.000	-0.030	-546.287	0.001
		N-PVOENS	-0.007	0.000	-0.012	-119.894	0.005
	Wrong scores	Constant	-18.125	0.000		-53938.980	0.000
		G-MADOG	0.076	0.000	1.208	184699.920	0.000
		N-TTIFC	0.022	0.000	0.290	86855.486	0.000
		G-PVOENS	-0.054	0.000	-0.114	-29610.833	0.000
		F-AFD	-29.275	0.001	-0.215	-27618.552	0.000
		PSQ-anxiety	1.225	0.000	0.057	14700.097	0.000
		N-DOFG	-0.003	0.000	-0.031	-6658.821	0.000
		G-AP	0.120	0.000	0.005	1324.274	0.000
Geometry	Correct scores	Constant	193.771	0.004		50505.618	0.000
		G-AFD	-120.759	0.007	-0.831	-17695.094	0.000

	N-AP	-13.586	0.001	-0.463	-12648.156	0.000
	G-TTENS	0.000	0.000	-0.323	-10147.487	0.000
	F-MIDOF	-0.023	0.000	-0.125	-4876.009	0.000
	G-TFD	0.119	0.000	0.065	751.890	0.001
	F-CC	-0.122	0.000	-0.022	-413.797	0.002
	EF-GEC	-0.291	0.002	-0.006	-155.657	0.004
Wrong	Constant	5.213	0.004		1428.568	0.000
scores	**G-AVD**	104.847	0.007	0.737	16051.377	0.000
	G-CC	-1.863	0.000	-0.281	-20474.318	0.000
	N-TWOEM	-2.083	0.000	-0.188	-9964.080	0.000
	G-MIDOF	0.053	0.000	0.154	8655.496	0.000
	G-TTFC	0.054	0.000	0.035	2051.007	0.000
	EF-GEC	1.561	0.001	0.023	1439.673	0.000
	F-ADOG	-0.001	0.000	-0.004	-130.871	0.005

9. Table S5.

Results of identification of patients with ADHD and male/female ADHD using ET indicators

Identification group	Identification criteria	Sensitivity (%)	Specificity (%)	AUC
ADHD or HC	TWEM(G)	82.61	33.33	0.538

	TWOEM(G)	94.20	88.89	0.962
	Test (Combination)	100.00	100.00	1.000
	Subtest (TWOEM- N-3)	100.00	100.00	1.000
Male ADHD or Female ADHD	MCWOEM(G)	83.93	53.85	0.615
	MIDOG(N)	48.33	92.31	0.701
	Test (Combination)	86.79	66.67	0.753
	TFFFTC(N-2)	60.00	76.92	0.700
	DOFG(N-4)	91.67	46.15	0.632
	TTFG(N-5)	56.67	69.23	0.650
	MCWOEM(G-1)	91.07	53.85	0.648
	Subtest (Combination)	94.34	75.00	0.899

AUC: area under the receiver operating characteristic curve