cpsy

an open access journal

Citation: Mihali, A., Young, A. G., Adler, L. A., Halassa, M. M., Ma, W. J. (2018). A low-level perceptual correlate of behavioral and clinical deficits in ADHD. Computational Psychiatry, 2, 141163. https://doi.org/10.1162/cpsy_a_00018

DOI: https://doi.org/10.1162/cpsy_a_00018

Supporting Information: https//doi.org/10.1162/cpsy_a_00018

Received: 23 October 2017

Accepted: 10 July 2018

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

Corresponding Author: Andra Mihali, Wei Ji Ma, Michael M. Halassa
alm652@nyu.edu, weijima@nyu.edu, mhalassa@mit.edu

The MIT Press

Appendix

DEMOGRAPHIC AND CLINICAL INFORMATION

	Demographics		
Group	Gender	Age	# White
Control	$11F, 9M$	32.5 ± 6.1	11
ADHD	$12F, 8M$	35.3 ± 10.0	9

Table A1: Demographic information of participants.

Values represent mean and standard deviation.

	ADHD scale scores					
Group	ACDS	ACDS2	ASRS	AISRS	MCI	GEC
Control	25.1 ± 4.1	1.3 ± 1.7	19.3 ± 8.8	6.9 ± 4.7	45.7 ± 8.1	45.6 ± 8.5
ADHD	52.8 ± 6.8	14.4 ± 2.5	49.7 ± 6.6	36.5 ± 7.9	73.7 ± 9.0	71.2 ± 6.8
Wilcoxon rank-sum p values	$< 10^{-7}$	$< 10^{-7}$	$< 10^{-5}$	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$

Table A2: Psychiatric scores of participants.

Values represent mean and standard deviation. ACDS denotes ACDS B1-B18, and ACDS2 B22-B39.

	ASRS	AISRS	ACDS	ACDS2	GEC
ASRS					
AISRS	0.85***				
ACDS	0.90***	0.96***			
ACDS2	0.85***	0.96***	0.94***		
GEC	0.85***	0.77***	0.82***	0.81***	
MCI	0.81***	0.81***	0.84***	0.83***	0.95***

Table A3: Spearman correlations across the scores for all diagnosis scales.

10 participants were excluded from this table because not all records were available. However, ACDS, ACDS2, GEC and MCI were available for all participants.

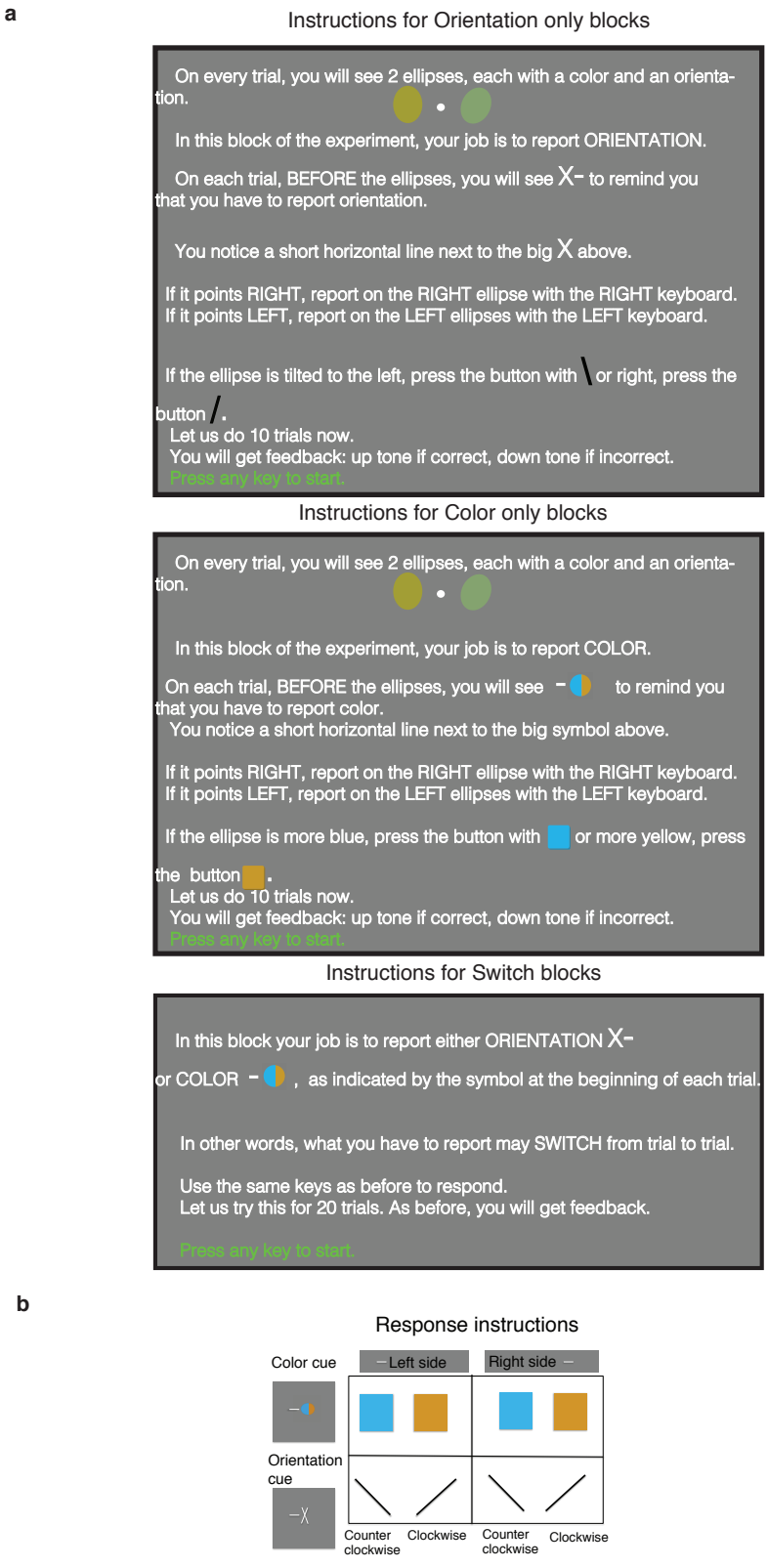

Figure A1. Training information. A) Training instructions for the 3 different types of blocks: Orientation only, Color only and Switch. B) Reminder of the stimulus-response pairings. A sheet containing this information was present on the wall of the psychophysics room within participants' sight.

FURTHER CHARACTERIZATION OF RESPONSES

Accuracy was maintained approximately constant across participants and conditions (mean $\pm$ sem: 0.811 ± 0.007) due to the Psybayes method of adaptive thresholding (Figure A2a). We further characterized the TIMO responses, first with a breakdown by error type, available for 32 participants (Figure A2B) and then according to the type of the previous trial (Figure A2c). In the Switch trials, the majority of TIMO seem to be feature errors, possibly because the mapping from left/right visual field to left/right keyboard was more intuitive than from feature dimension to top/down of a keyboard (Figure A2B). Lastly, we saw no clear pattern from breaking down proportion of TIMO by the type of the previous trial (Figure A2C).

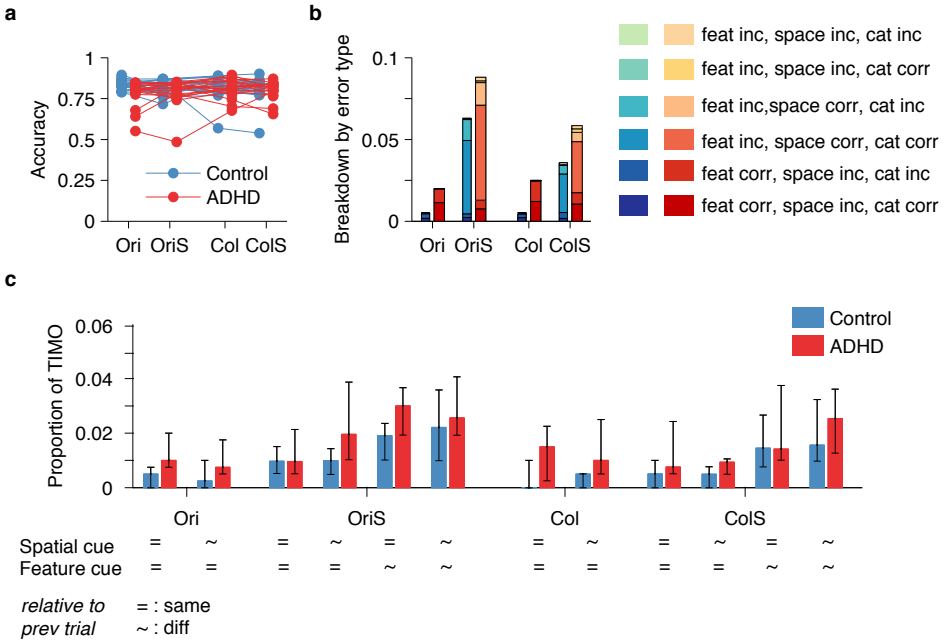

Figure A2. Further characterization of responses. A) Accuracy as proportion correct on the trials when participants selected one of the 2 relevant keys. B) The task irrelevant motor output from Figure 1, broken down by error type. For every condition, the first bar is controls, and the second one ADHD. C) Same, broken down by the type of the previous trial.

FURTHER INFORMATION ON REACTION TIMES

ex-Gaussian model

ex-Gaussian distributions are commonly fitted to reaction time data and are defined by adding 2 random variables, a Gaussian with parameters μ and σ and an exponential with parameter τ. While in our data τ showed an effect of group, neither log μ_{RT} ($F(1,38) = 0.05, p = 0.83, \eta_p^2 = 0.001$) nor log σ_{RT} ($F(1,38) = 0.27, p = 0.61, \eta_p^2 = 0.007$) did, consistent with other studies that showed significant effects of group on τ but not on μ_{RT} (see (?) for a meta-analysis).

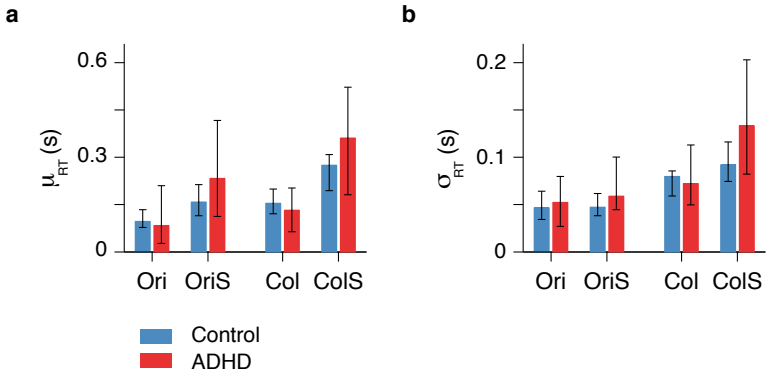

Figure A3. ex-Gaussian parameters fitted to the reaction time distributions across conditions and groups. A) Gaussian mean μ_{RT}, B) Gaussian standard deviation σ_{RT}, both for each task condition.

Alternative models

While ex-Gaussian distributions are routinely used to fit reaction times, they are rarely compared to alternative distributions. We used the corrected Akaike Information Criterion (AICc) and the Bayesian Information criterion (BIC) to compare the ex-Gaussian fits with the fits of 2 other distributions on the positive real line: log-Normal and Gamma. These metrics are defined as AICc $= -2\mathrm{LL}^* + 2k + \frac{2k(k+1)}{n_{\mathrm{trials}}-k-1}$ (?) and BIC $= -2\mathrm{LL}^* + k \log n_{\mathrm{trials}}$ (?), respectively, where LL* is the maximum log likelihood, k is the number of free parameters, and n_{trials} is the number of trials. We found that indeed the ex-Gaussian distribution was a better fit than both the log-Normal (in median by 611 according to AICc and by 607 according to BIC) and the Gamma distribution (in median by 50 according to AICc and by 45 according to BIC); see Figure A4 for individual subjects.

Non-parametric measure of RT variability

We complemented the results about RT τ (Figure 3) with a non-parametric robust measure of intra-individual reaction time variability, the reaction time inter-quartile range (iqr) (Figure A5). Three-way mixed-design ANOVA on log RT iqr's revealed a significant effect of group ($F(1,38) = 5.13, p = 0.029, \eta_p^2 = 0.12$), load ($F(1,38) = 18.84, p < 0.001, \eta_p^2 = 0.33$), feature ($F(1,38) = 21.38, p < 0.001, \eta_p^2 = 0.36$), and a significant load × feature interaction ($F(1,38) = 22.12, p < 0.001, \eta_p^2 = 0.37$). No other two-way interaction nor the three-way interaction were significant ($p > 0.36$). After Sidak correction ($\alpha = 0.0043$), none of the

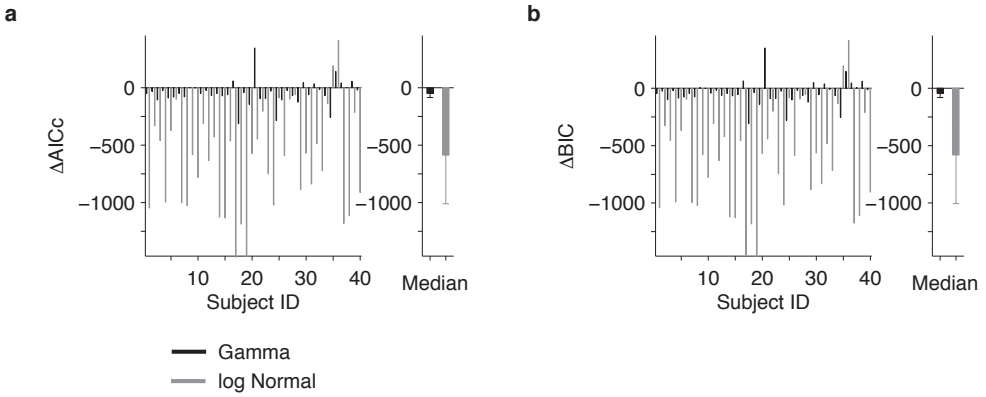

Figure A4. Model comparison justifies the parametrization of reaction times with the ex-Gaussian distribution. A) The ex-Gaussian model has the lowest AICc across the population (Right) and for almost all individual subjects (Left). B) Same result for BIC.

between-groups comparisons were significant. Within controls, the effects of load and feature on log RT iqr were significant for Ori vs OriS ($p = 0.0015$) and Ori vs Col ($p < 0.001$); within ADHD, the only significant effect was of load for Ori vs OriS ($p = 0.002$).

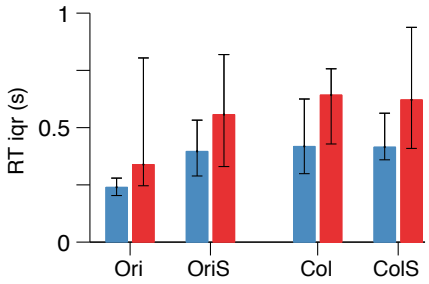

Figure A5. Reaction time variability is higher in ADHD also according to a non-parametric metric, RT iqr.

FURTHER INFORMATION ON PSYCHOMETRIC CURVES AND PARAMETERS

Stimuli sets

Figure A6 depicts the histograms of selected stimuli for each condition and each participant, optimized with the Bayesian stimulus selection method. As a consequence of this method, proportion correct is largely stable across conditions and participants (see Results ??), and the differences between participants were quantified through the psychometric curve parameters. In line with ADHD participants having higher perceptual variability, we see here that the collapsed histograms across all participants within a group show that controls received a higher proportion of more difficult stimuli (higher bump around 0).

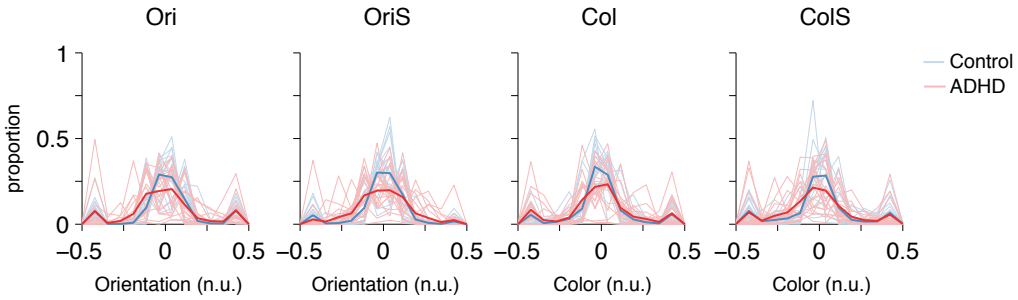

Figure A6. Distributions of stimuli across conditions and participants. Thin lines: individual participants. Thick lines: proportion of stimuli collapsed across all participants within a group.

PSE

Figure A8d shows the estimates of μ (PSE) in the "shared" (main) model. Two-way mixed-design ANOVA on μ with within-group factor feature showed an effect of group ($F(1,38) = 9.47, p = 0.004, \eta_p^2 = 0.2$), but no significant effect of feature ($F(1,38) = 1.17, p = 0.28, \eta_p^2 = 0.03$) and not a significant interaction. After Sidak correction ($\alpha = 0.0253$), no effects were significant. We chose to interpolate color values between blue and yellow since the S-cone pathway is of special interest in ADHD (?). While we found an overall group effect on μ, after Sidak correction the post hoc effect for color failed to reach significance, thus making our results at this point inconclusive about whether ADHD participants have different S-cone dependent color processing.

"Full" model, 12 parameters

While in the main or "shared" model with 8 parameters (Figure 4) we assumed that μ and λ were shared within a feature across load conditions, in the "full" model we did not constrain any parameters, yielding 12 parameters total.

As expected, the "full" model captured the data at least as well as the "shared" model. However, the "shared" model provided either a comparable (in median better by -1.5 according to AICc) or better (in median by -20 according to BIC) than the "full" model (Figure A7). This confirmed the plausibility of the shared-parameters assumption in the main model.

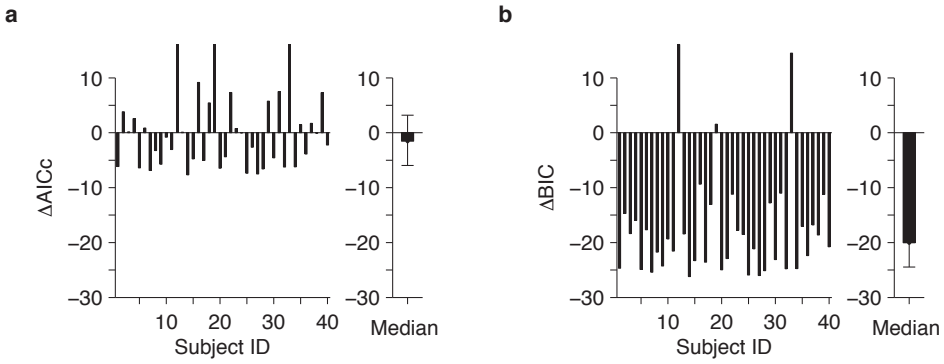

Figure A7. Model comparison justifies using the "shared" model. A) AICc of the "shared" model minus AICc of the "full" model for Left) individual subjects and Right) Group - median and 95% bootstrapped confidence intervals. B) Same for BIC.

We also performed three-way mixed-design ANOVA on the parameter estimates from the "full" model (Figure A8B). Just like in the "shared" model, we found a significant effect of group for log perceptual variability (σ) ($F(1,38) = 5.21, p = 0.028, \eta_p^2 = 0.12$), a significant effect of feature ($F(1,38) = 37.11, p < 0.0001, \eta_p^2 = 0.49$), but no significant effect of load ($F(1,38) = 0.03, p = 0.87, \eta_p^2 = 0.001$). Neither of the two-way interactions nor the three-way interaction were significant ($p > 0.06$). After Sidak correction ($\alpha = 0.0065$, 8 comparisons, since, as in the main model, we excluded across feature comparisons due to their different units) we found a between-group effect for Ori with $p < 0.0001$ and OriS ($p < 0.0025$), but no significant effects of group for neither Col ($p = 0.0063$) nor ColS ($p = 0.47$) (Figure A8).

For the log lapse λ, we found a significant effect of feature ($F(1,38) = 25.88, p < 0.0001, \eta_p^2 = 0.40$) and a significant feature × group interaction ($F(1,38) = 6.01, p = 0.02, \eta_p^2 = 0.14$); nothing else was significant ($p > 0.09$). After Sidak correction ($\alpha = 0.0043$, all 12 comparisons make sense since λ is unitless), no between-group comparisons were significant ($p > 0.02$). Within controls, the feature comparisons Ori vs Col and OriS vs ColS were significant ($p < 0.001$), but not the load ones. Within ADHD, neither the feature nor the load comparisons reached significance ($p > 0.02$).

For the PSE μ, like in the "shared" model, we found a significant effect of group ($F(1,38) = 10.85, p = 0.002, \eta_p^2 = 0.22$) and also a significant group × load × feature interaction ($F(1,38) = 8.42, p = 0.006, \eta_p^2 = 0.18$), nothing else reaching significance ($p > 0.09$). After Sidak correction ($\alpha = 0.0065$, as for σ), we found a significant difference between ADHD and controls for ColS ($p = 0.002$), but not for Col ($p = 0.18$) and not for Ori or OriS ($p > 0.12$). Again, these results cannot provide robust support for ADHD participants having different S-cone dependent color processing.

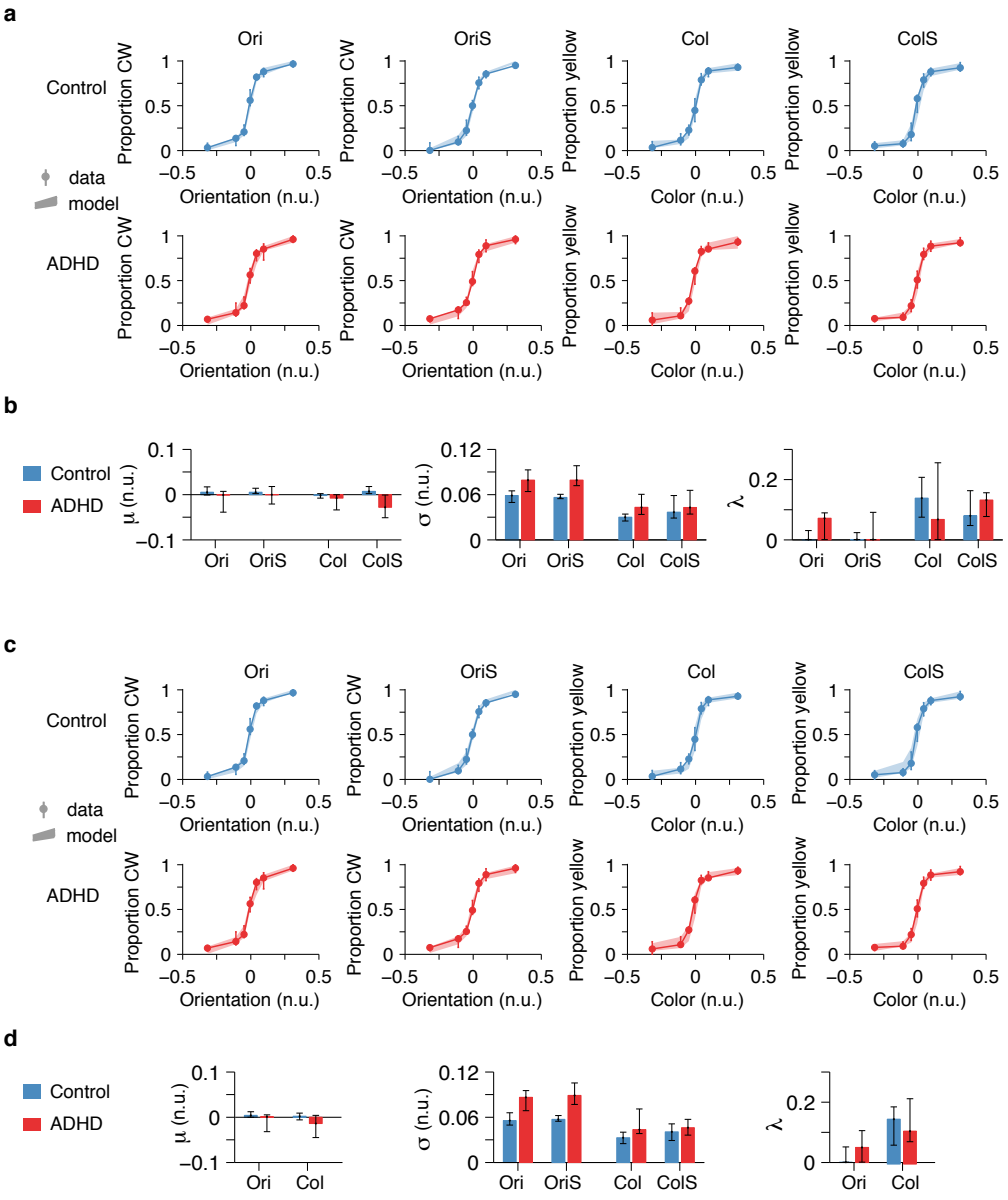

Figure A8. Psychometric curves for both models: data, model fits and parameter values. A) "Full model" (12 parameters total): data and fitted psychometric curves. Solid circles with error bars show median and 95% bootstrapped confidence intervals, while shaded areas show the same for model predictions. The data was binned into 7 quantiles. Since the Bayesian adaptive method presented each participant in each condition with an unique set of stimuli, the midpoint stimulus values of the quantile bins differed for each. However, for ease of visualization, here we place the midpoints stimulus values for each bin as the midpoints obtained from binning into 7 quantiles the entire stimulus set concatenated across participants and conditions. B) "Full model": MLE parameter fits, C) "Shared model" (8 parameters total): data and fitted psychometric curves. D) "Shared model": MLE parameter fits.

Effect of learning

To assess learning across the experiment, we looked at the parameter estimates from the first half of the trials versus the second. Figure A9 shows that the perceptual variability parameters improved slightly on the second versus first half of trials, sign that there might be some learning. As reported in main, four-way mixed-design ANOVA on log σ confirmed a significant effect of time: $F(1,38) = 12.7, p = 0.001, \eta_p^2 = 0.25$. However, we note that these parameter estimates are not as reliable as the ones in Figure 3B, since they were obtained by fitting on only half the data.

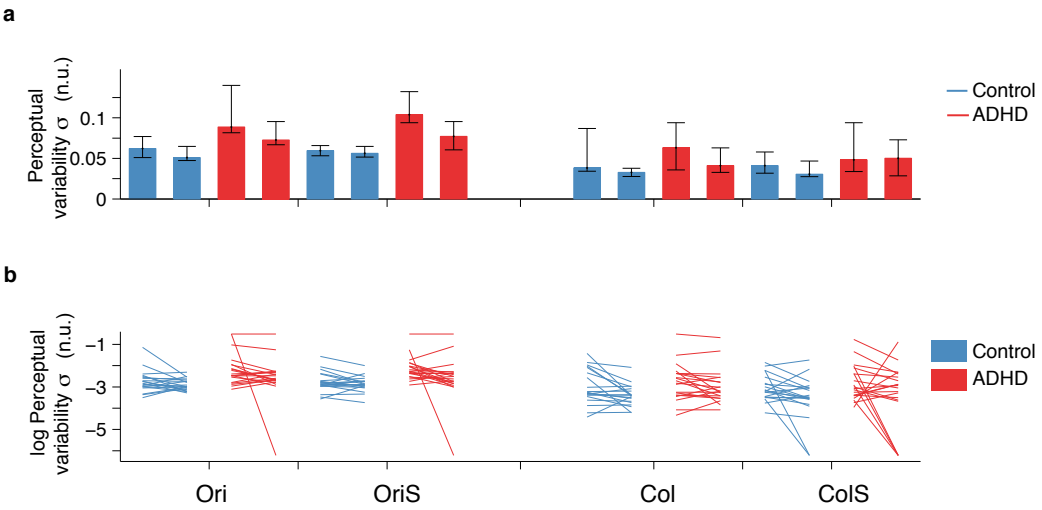

Figure A9. Perceptual variability parameter fits across time. A) Medians across participants and bootstrapped 95% confidence intervals. B) log Perceptual variability values of individual participants. Lines show how the perceptual variability differs from the first half of trials to the second half.

EFFECT OF EYE TRACKING

A possible concern is that half of the participants in each group were eye-tracked, while half were not. If an eye-tracked participant broke fixation, they had to redo the trial. As a result, the eye-tracked participants started more trials (mean and SD for eye-tracked: 1047 $\pm$ 201 trials; non-eye-tracked always completed 800 trials). Thus, a concern could be that differences in task metrics could simply arise due to the experiment being longer and as a result more tiresome. We examined each of the average task metrics within each group, separately for the eye-tracked participants and the non-eye-tracked ones and found no significant differences (Wilcoxon rank-sum $p > 0.13$) (Figure A10).

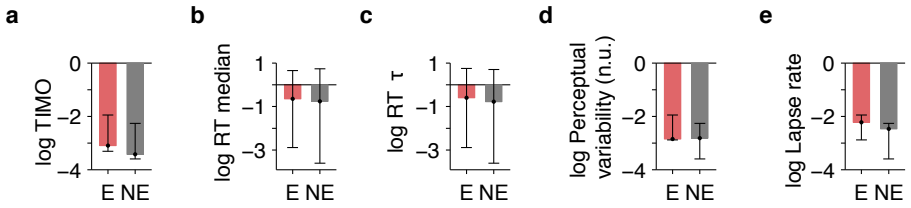

Figure A10. No significant difference between eye tracked (E) and non-eye tracked (NE) participants on behavioral task metrics: A) TIMO, B) RT median, C) RT τ, D) Perceptual variability, E) Lapse rate. Bars represent medians and error bars bootstrapped 95% confidence intervals.

BREAKDOWN OF CORRELATIONS

By group

In Figure A11, we show the points that make up the correlations from Table 1, color coded by group. Of note, the two ADHD participants who had visibly lower orientation discrimination performance (Figure 4A), did not also have outstandingly reduced performance on other metrics; more detailed ophtamological examination could have provided more insight into the possible sources of their reduced orientation discrimination performance.

In Table A4 we show the pairwise correlations across task metrics separately within the Control group and within the ADHD group. Here, the only group specific correlations that survive the multiple-comparisons correction are RT with RT τ and GEC with ACDS.

In addition, we attempted to determine whether for a given pair of task metrics, their correlation within the ADHD group is different from their correlation within the Control group. To do this, we compared the difference between the actual correlations to a distribution of differences between correlations obtained by shuffling the ADHD and Control labels. We did not find significant differences between the ADHD and Control correlations for any pair of task metrics ($p > 0.04$).

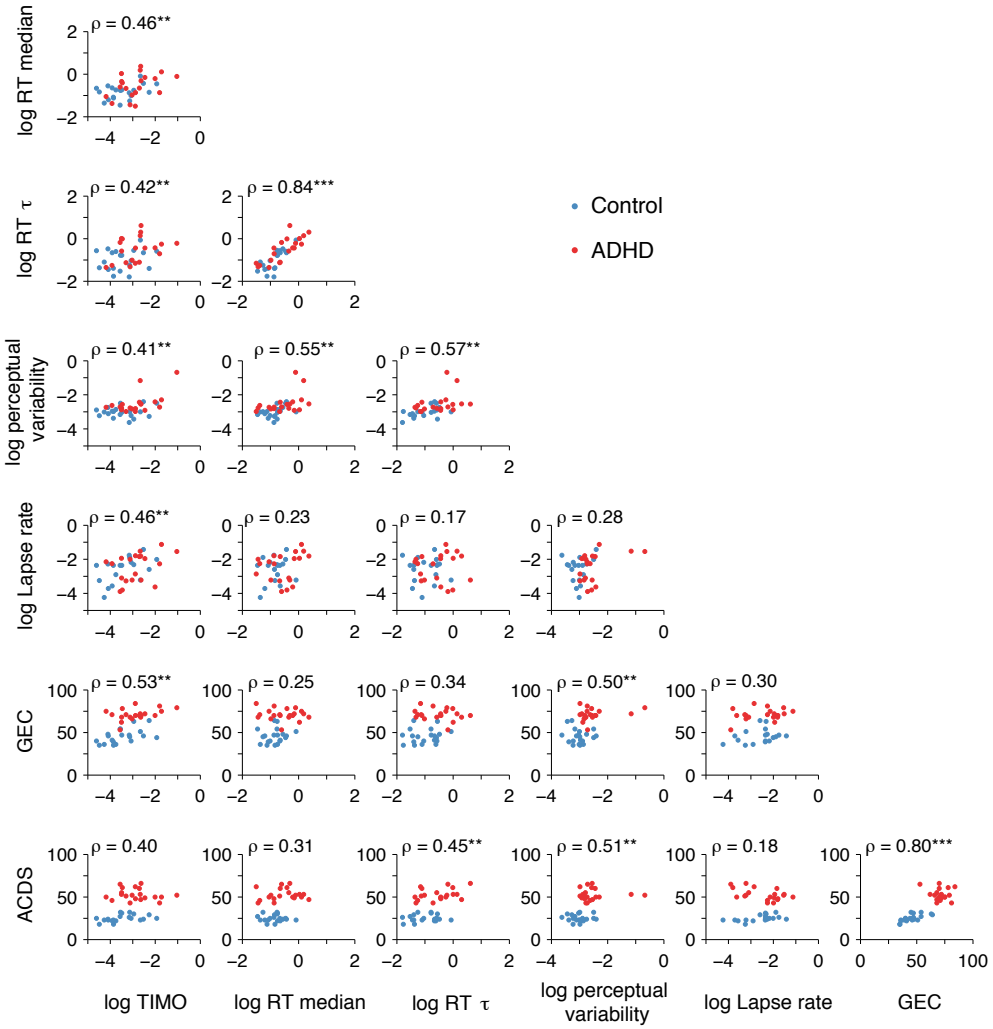

Figure A11. Dots: pairwise task metrics, color coded by group. We also show here the Spearman correlations collapsed across groups, as in Table 1. ** depicts $0.001 < p < 0.0089$ (since $\alpha_{\text{Sidak}} = 0.0089$ after multiple-comparisons correction) and *** depicts $p < 0.001$.

Table A4: No evident pattern of group specific correlations. Pairwise Spearman correlations across task metrics (both behavioral and clinical), as in Table 1, but divided by group. Boldfaced is significant after multiple-comparisons correction, $\alpha = 0.0083$ for Control and $\alpha = 0.0082$ for ADHD. (see Methods)

(a) Control

	TIMO	RT	RT τ	Perceptual variability (σ)	Lapse (λ)	GEC
TIMO						
RT	$\rho = 0.25$ $p = 0.27$					
RT τ	$\rho = 0.13$ $p = 0.59$	$\boldsymbol{\rho = 0.81}$ $p < 0.0001$				
Perceptual variability (σ)	$\rho = -0.02$ $p = 0.92$	$\rho = 0.52$ $p = 0.02$	$\rho = 0.50$ $p = 0.03$			
Lapse rate (λ)	$\rho = 0.49$ $p = 0.03$	$\rho = 0.22$ $p = 0.34$	$\rho = 0.19$ $p = 0.42$	$\rho = -0.05$ $p = 0.84$		
GEC	$\rho = 0.52$ $p = 0.02$	$\rho = 0.16$ $p = 0.49$	$\rho = 0.15$ $p = 0.52$	$\rho = -0.12$ $p = 0.61$	$\rho = 0.25$ $p = 0.29$	
ACDS	$\rho = 0.49$, $p = 0.03$	$\rho = -0.09$ $p = 0.67$	$\rho = 0.03$ $p = 0.89$	$\rho = -0.09$ $p = 0.69$	$\rho = 0.46$ $p = 0.04$	$\boldsymbol{\rho = 0.70}$ $p < 0.0001$

(b) ADHD

	TIMO	RT	RT τ	Perceptual variability (σ)	Lapse (λ)	GEC
TIMO						
RT	$\rho = 0.48$ $p = 0.03$					
RT τ	$\rho = 0.36$ $p = 0.12$	$\boldsymbol{\rho = 0.82}$ $p < 0.0001$				
Perceptual variability (σ)	$\rho = 0.46$ $p = 0.04$	$\rho = 0.52$ $p = 0.02$	$\rho = 0.46$ $p = 0.04$			
Lapse rate (λ)	$\rho = 0.40$ $p = 0.08$	$\rho = 0.38$ $p = 0.10$	$\rho = 0.12$ $p = 0.61$	$\rho = 0.32$ $p = 0.16$		
GEC	$\rho = 0.37$ $p = 0.11$	$\rho = -0.10$ $p = 0.68$	$\rho = 0.19$ $p = 0.42$	r = 0.15 $p = 0.54$	$r = 0.18$ $p = 0.44$	
ACDS	$\rho = -0.18$ $p = 0.45$	$\rho = 0.14$ $p = 0.56$	$\rho = 0.36$ $p = 0.11$	$\rho = 0.03$ $p = 0.87$	$\rho = -0.35$ $p = 0.12$	$\rho = -0.09$ $p = 0.07$

By symptom type

For this analysis, 2 ADHD participants were excluded due to missing AISRS records. A breakdown of the AISRS scores into inattentive and hyperactive shows that their correlations with task metrics recapitulate the correlations seen with ACDS. This is not unexpected, given the high correlation between ACDS and AISRS scores, as well as the fact that the AISRS inattentive and AISRS hyperactive scores were highly correlated ($\rho = 0.89, p < 10^{-13}$).

Table A5: No evident pattern of differential correlations by symptom type. AISRS inattentive and hyperactive correlations with behavioral task metrics are almost identical and largely recapitulate the ACDS correlations. Boldfaced represents $p < 0.0089$.

	TIMO	RT	RT τ	Perceptual variability (σ)	Lapse (λ)	GEC
ACDS	$\rho = 0.40$ $p = 0.01$	$\rho = 0.31$ $p = 0.05$	$\boldsymbol{\rho = 0.45}$ $p = 0.004$	$\boldsymbol{\rho = 0.51}$ $p = 0.0008$	$\rho = 0.18$ $p = 0.26$	$\boldsymbol{\rho = 0.80}$ $p < 0.0001$
AISRS inattentive	$\rho = 0.41$ $p = 0.01$	$\boldsymbol{\rho = 0.43}$ $p = 0.006$	$\boldsymbol{\rho = 0.46}$ $p = 0.003$	$\boldsymbol{\rho = 0.51}$ $p = 0.001$	$\rho = 0.22$ $p = 0.17$	$\boldsymbol{\rho = 0.65}$ $p < 0.0001$
AISRS hyperactive	$\rho = 0.41$ $p = 0.01$	$\boldsymbol{\rho = 0.48}$ $p = 0.002$	$\boldsymbol{\rho = 0.49}$ $p = 0.001$	$\boldsymbol{\rho = 0.63}$ $p < 0.0001$	$\rho = 0.18$ $p = 0.27$	$\boldsymbol{\rho = 0.65}$ $p < 0.0001$

By condition

In Table 1, for each participant, we averaged each behavioral metric across all four conditions. In Figure A12, we present the correlations of perceptual variability with TIMO, RT and RT τ broken down by condition. The correlations that survived after multiple-comparisons correction are between perceptual variability (σ) and TIMO in the Ori condition, between σ and RT in the OriS and σ and RT τ in Ori and Col. Overall, we cannot conclude much from these patterns of results.

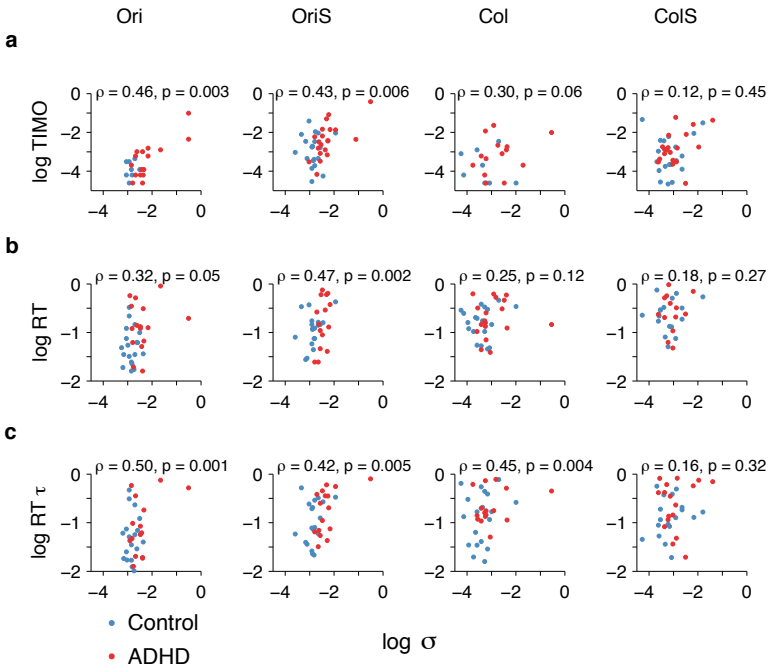

Figure A12. Spearman correlations of perceptual variability with other behavioral metrics broken down by conditions show no conclusive pattern. We show the correlations of log perceptual variability with A) log TIMO, B) log RT median and C) log RT τ.

PREDICTION OF CLINICAL VARIABLES

Logistic regression: prediction of diagnosis from behavioral metrics

Table A6: Logistic regression coefficients, mean ± sem.

(a) Diagnosis ∼ log perceptual variability

	coefficient	t value	p value
intercept	13.3 ± 4.8	2.78	0.0055**
log perceptual variability (σ)	4.7 ± 1.7	2.80	0.0051**

(b) Diagnosis ∼ log TIMO

	coefficient	t value	p value
intercept	1.9 ± 0.9	2.07	0.038*
log TIMO	1.1 ± 0.5	2.25	0.024*

(c) Diagnosis ∼ log perceptual variability + log TIMO

	coefficient	t value	p value
intercept	13.3 ± 4.9	2.70	0.0069**
log perceptual variability (σ)	4.3 ± 1.7	2.47	0.013*
log TIMO	0.63 ± 0.59	1.07	0.28

(d) Diagnosis ∼ log perceptual variability + log TIMO + log RT median + log RT τ + log lapse rate

	coefficient	t value	p value
intercept	14.0 ± 5.4	2.57	0.010*
log TIMO	0.49 ± 0.64	0.76	0.44
log RT median	-1.4 ± 1.7	-0.81	0.42
log RT τ	1.4 ± 1.6	0.88	0.38
log perceptual variability (σ)	4.2 ± 1.9	2.23	0.025*
log lapse rate (λ)	0.38 ± 0.37	1.02	0.30

Linear regression: prediction of clinical metrics GEC and ACDS from behavioral metrics

Table A7: Linear regression coefficients, depicted as mean ± sem, for GEC and ACDS with task metrics.

(a) GEC ∼ log perceptual variability + log TIMO + log RT + log RT τ + log lapse rate.

	coefficient	t value	p value
log intercept	103 ± 12	8.50	$< 10^{-9}$***
log TIMO	8.0 ± 3.2	2.53	0.016*
log RT median	-10.3 ± 7.6	-1.35	0.18
log RT τ	8.1 ± 6.1	1.32	0.19
log σ	6.1 ± 4.9	1.23	0.22
log λ	1.3 ± 1.8	0.75	0.46

(b) ACDS ∼ log perceptual variability + log TIMO + log RT + log RT τ + log lapse rate.

coefficient	t value	p value
73 ± 13	5.77	$< 10^{-5}$***
3.4 ± 3.3	1.04	0.31
-11.3 ± 8.0	-1.41	0.15
14.4 ± 6.4	2.26	0.031*
6.8 ± 5.1	1.33	0.19
0.5 ± 1.9	0.29	0.78