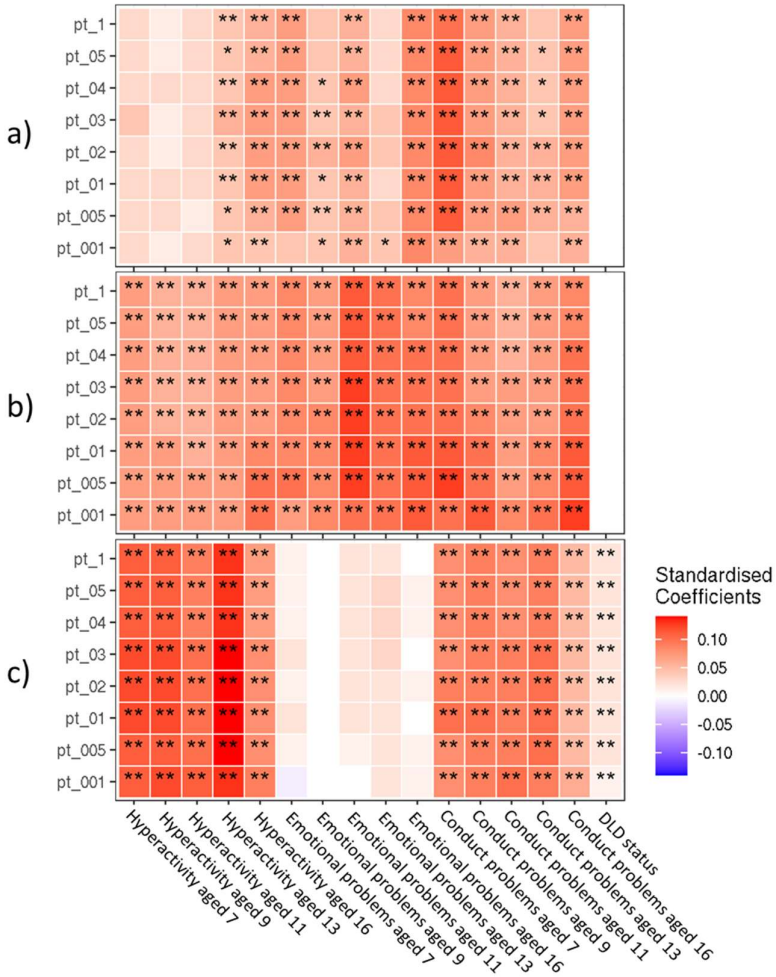
Supplemental Material S1. Details of exploratory analysis used to determine the *p*-value threshold for the polygenic scores. Details of the model fit statistics for RQ2 and analysis with continuous PGS scores for RQ1 and RQ3.

Exploratory Analysis

To select the appropriate polygenic scores (PGS) at specific p-value thresholds for inclusion in downstream analysis, exploratory analysis was conducted. This analysis examined associations between polygenic scores for major depressive disorder (MDD), anxiety disorder (AD), attention deficit hyperactivity disorder (ADHD) and emotional problems, conduct problems, and hyperactivity, as well as Developmental Language Disorder (DLD). Initially, associations with difficulties at each of the time points was assessed in separate models and then these associations were tested in multi-level models accounting for clustering based on age. Associations were assessed at several p-value thresholds ranging from 0.01 to 1 using multiple linear regression models (**Figure S1**).

Figure S1. Illustrating associations between (a) AD, (b) MDD, and (c) ADHD polygenic scores and the SDQ subscales (hyperactivity, emotional problems and conduct problems) at eight *p*-value thresholds (pt) at each of the five timepoints (ages 7, 9, 11, 13, 16). Nominally significant findings are indicated with a single asterisk (*), with a double asterisk (**) highlighting a finding that was significant follow multiple testing corrections.



Anxiety Disorder

Findings relating to the associations between the AD PGS and the SDQ subscales under assessment, whilst highlighting multiple significant associations, also demonstrated inconsistencies in associations across timepoints and *p*-value thresholds. For instance, the first three timepoints (ages 7, 9, 11) showed no significant associations at any *p*-value threshold. However, associations were then shown to increase at timepoint four (age 13), highlighting three nominally significant finding and five that remained significant following multiple testing corrections. Associations then became consistently significant across all thresholds following multiple testing corrections at timepoint 5 (age 16). This increase in association, from earlier to later timepoints may imply that the effect of genetic risk for AD on hyperactivity increases over time.

Inconsistencies were also evident for associations between the AD PGS and emotional problems. Whilst timepoint one (age 7) highlighted significant associations following multiple testing corrections at all but one threshold (0.01), associations were then shown to drop at timepoint two (age 9) with associations at only three thresholds (0.3, 0.2, 0.05) surviving multiple testing corrections. Timepoint three (age 11) then saw significant associations at all *p*-value thresholds, with all associations surviving corrections for multiple testing. Significant associations all but disappeared at timepoint four (age 13) except for one nominally significant association at the lowest *p*-value threshold (0.01) before returning at timepoint five (age 16) with increased significance and effect.

Associations between the AD PGS and conduct problems demonstrated a far more consistent pattern of association, with the first three time points (ages 7, 9, 11) highlighting significant associations following multiple testing corrections at all *p*-value thresholds. This consistent pattern of association then dissipated at timepoint four (age 13) with only three thresholds showing significant effects following corrections for multiple testing (0.2, 0.1, 0.05). Significant associations at all thresholds then returned at timepoint five (age 16) with all associations remaining significant following multiple testing corrections.

Despite the inconsistency in significant associations, all results demonstrated a consistent positive direction of effect suggesting that genetic risk for AD is associated with increases in levels of hyperactivity, emotional problems and conduct problems. However, this was not the case regarding the DLD status as there were no associations, significant or otherwise, between the AD PGS and DLD status at any timepoint or p-value threshold.

Major Depressive Disorder

Results regarding the MDD PGS demonstrated strong and consistent significant associations with all three SDQ subscales. These associations were significant at all timepoints and thresholds. All finding were found to have a consistent positive direction of effect and were significant following corrections for multiple testing. This suggests that genetic risk for depression is associated with increased hyperactivity, emotional problems and conduct problems alike. As with the findings regarding the AD PGS there was no associations, significant or otherwise observed between genetic risk for depression and DLD status at any timepoint or threshold.

Attention Deficit Hyperactivity Disorder

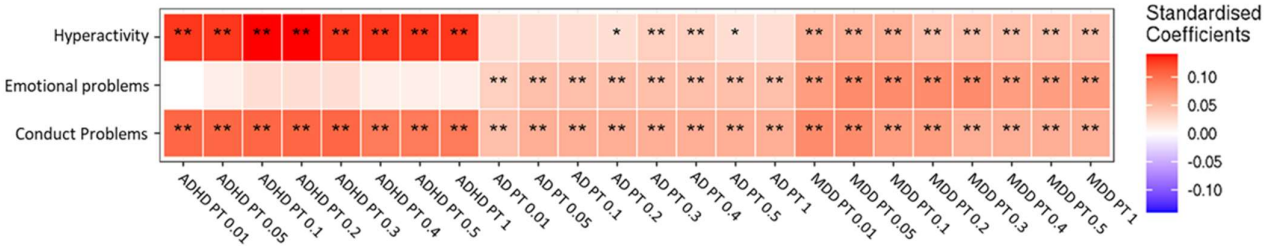
Results relating to the associations between the ADHD PGS and the SDQ subscales, whilst not consistently significant across all subscales, did reveal a specific pattern of associations, with the subscales hyperactivity and conduct problems showing consistently significant positive associations with genetic risk for ADHD. Conversely, no significant associations, nominally or otherwise, were observed regarding the ADHD PGS and emotional

problems suggesting that genetic risk for ADHD is not significantly associated with emotional problems. However, unlike the findings regarding both the MDD and AD PGS, the ADHD PGS demonstrated consistently significant positive associations, at all p-value thresholds, with DLD status. Furthermore, despite effect sizes being small, all association survived correction for multiple testing. This therefore suggests that genetic variant implicated in ADHD are also associated with a positive DLD status within the current sample.

Accounting for Time-Ordered Natures of Data

To further assess these effects, linear mixed effects models were constructed examining associations between the three polygenic scores at the same eight p-value thresholds and each of the SDQ subscales across all timepoints. This approach maximizes the effective power of the sample to assess the consistency and robustness of the previous associations. All models included the fixed effects of age and sex and the random effect age. Results of these analyses are illustrated below in **Figure S2**.

Figure S2. Heatmap illustrating findings from the mixed effects models examining the associations between AD, MDD, and ADHD polygenic scores and the SDQ subscales (hyperactivity, emotional problems and conduct problems) at eight p-value thresholds, and across all timepoints. Nominally significant findings are indicated with a single asterisk (*), with a double asterisk (**) highlighting a finding that was significant follow multiple testing corrections.



Findings from the linear mixed effects models demonstrate that the lack of association seen in **Figure S1** between the ADHD PGS and emotional problems was a robust finding, further confirming that genetic risk for ADHD is not significantly associated with emotional problems within the current sample. Furthermore, results also confirmed strong positive associations, at all thresholds following multiple testing corrections, between the ADHD PGS and both hyperactivity and conduct problems.

It was also revealed that whilst the association between genetic risk for anxiety and hyperactivity may increase over time, as suggested in the previous analysis, it also appears that the significance of this positive effect is limited to specific p-value thresholds. This is unlikely to be the result of a lack of power as the mixed effects approach used increases statistical power by assessing associations across all timepoints. However, it may suggest that the SNPs included at these specific thresholds are those driving the association, and that the reduction or inclusion of further unassociated SNPs are impacting on this relationship. Elsewhere, the AD PGS was shown to be significantly association with increases in emotional problems and conduct problems, all of which survived correction for multiple testing. This likely suggests that the inconsistencies seen between timepoints in the previous analysis were the result of a drop in power at specific timepoint due to missing SDQ data.

Lastly, and in line with the previous results from the regression analysis at each timepoint, the MDD PGS was found to be significant following multiple testing corrections at all thresholds across each of the SDQ subscales.

Taken together, findings highlight .3 and .4 as representing the most consistently significant *p*-value thresholds across each PGS, and the only to remain significant across each of the SDQ measures, with the exception of the ADHD PGS on emotional problems.

Variance Explained

Informed by the previous analyses, variance in the SDQ subscales explained by each PGS at the most consistently significant *p*-value thresholds (.3 and .4) were assessed at each time point. Results of these assessments are illustrated in the bar charts below (**Figure S3)**.

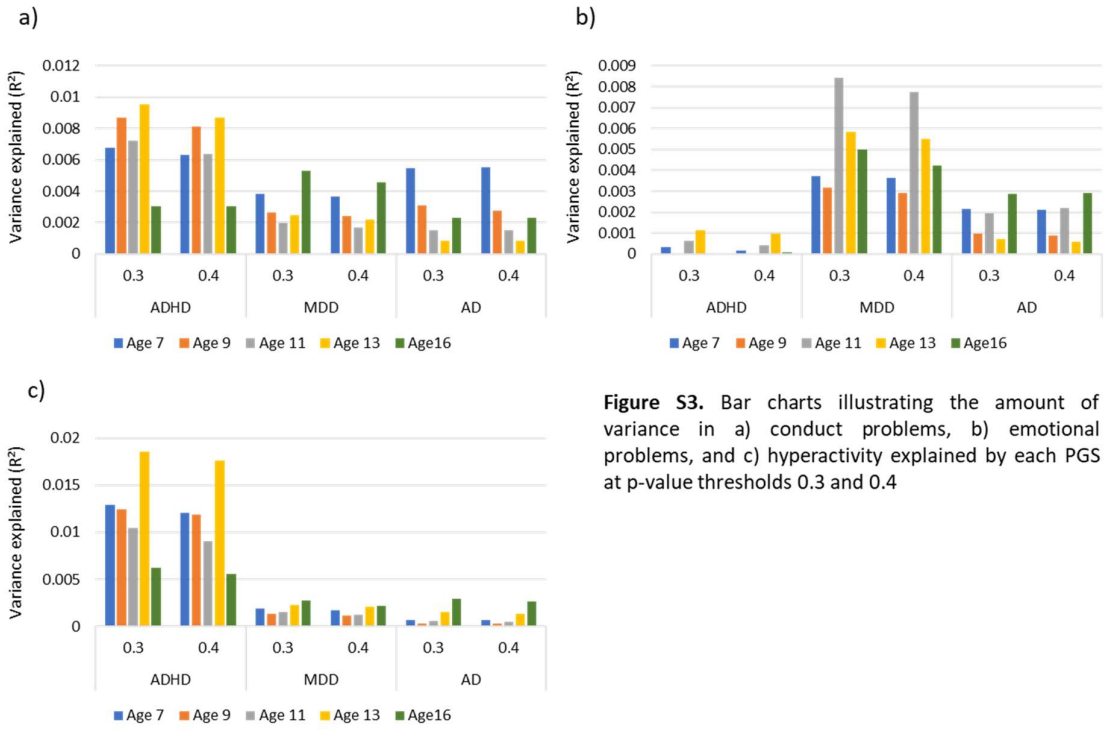


Figure S3. Bar charts illustrating the amount of variance in a) conduct problems, b) emotional problems, and c) hyperactivity explained by each PGS at p-value thresholds 0.3 and 0.4

Results revealed that a p-value threshold of 0.3 more consistently explained the most variance in each of the SDQ subscales compared to 0.4, although the differences were small. Variance explained across the three SDQ subscales by the each PGS at the .3 *p*-value threshold ranged from 0.07% to 0.5% for the AD PRS, 0.2% to 0.8% for the MDD PRS and < .001% to 1.9% for the ADHD PRS. Informed by these findings, and to reduce the multiple testing burden whilst also maximising power, the .3 *p*-value threshold of each PGS was used in further downstream analysis.

Polygenic Scores

PGSs were calculated using the .3 *p*-value threshold. The descriptive statistics for these PGSs are shown in Table S1 and the distribution of PGS is shown in Figures S4 and S5.

Table S1. Descriptive statistics of MDD, AD, and ADHD genome-wide polygenic scores (PT 0.3).

PGS (PT 0.3)	SNPs	*n*	Unstandardized mean (*SD*)	Standardized mean (*SD*)	Skewness	Kurtosis
MDD	139934	5395	-6.6×10^{-4} (1.9×10^{-5})	-5.03×10^{-17} (1)	-.034	3.031
AD	146154	5395	2×10^{-5} (5.6×10^{-5})	-2.17×10^{-17} (1)	-.003	3.122
ADHD	101631	5395	-4.2×10^{-4} (8×10^{-5})	-2.9×10^{-16} (1)	-.019	3.016

Note. SNPS = number of variants included; *n* = number of individuals in the ALSPAC dataset with genetic data; *SD* = standard deviation.

Figure S4. Histograms displaying the distributions of the unstandardized AD, MDD, and ADHD PRS for all individuals within the ALSPAC dataset. The distribution was assessed at a *p*-value thresholds of.3.

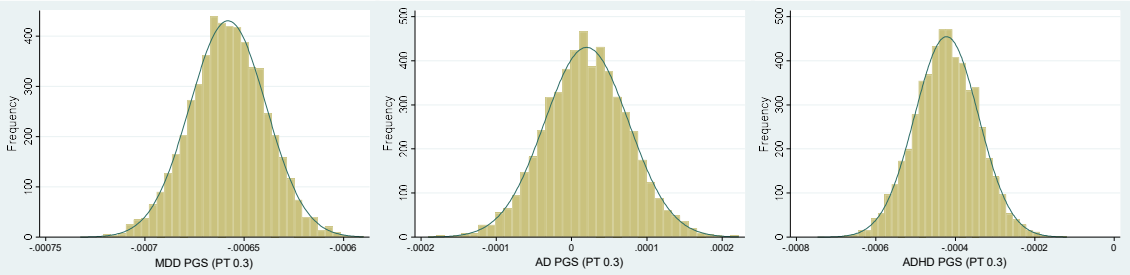


Figure S5. Histograms displaying the distributions of the standardized AD, MDD, and ADHD PRS for all individuals within the ALSPAC dataset. The distribution was assessed at a *p*-value thresholds of .3.

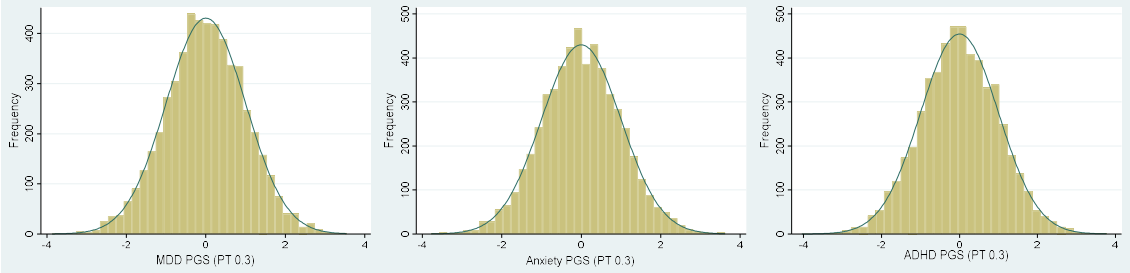


Table S2. Full results from mixed effects models with continuous PGS including main effects and co-variates.

Outcome Variable	Predictor PGS	Covariate Sex			Covariate Time			Main Effect PGS (Continuous)			Main Effect DLD			PGS x DLD Interaction		
		β	95% CI	*p*	β	95% CI	*p*	β	95% CI	*p*	β	95% CI	P	β	95% CI	*p*
Emotional Problems	Model 1: MDD	.18	0.13, 0.22	**< .001**	-.01	-0.02, -0.00	.007	.06	0.04, 0.08	**< .001**	.32	0.24, 0.39	**< .001**	.06	-0.02, 0.14	.113
	Model 2: AD	.18	0.14, 0.22	**< .001**	-.01	-0.02, -0.00	.007	.04	0.02, 0.06	**< .001**	.32	0.24, 0.39	**< .001**	-.02	-0.10, 0.05	.554
	Model 3: ADHD	.18	0.14, 0.22	**< .001**	-.01	-0.02, -0.00	.007	.01	-0.01, 0.03	.293	.31	0.23, 0.39	**< .001**	.01	-0.07, 0.08	.895
	Model 4: Combined	.18	0.13, 0.22	**< .001**	-.01	-0.02, -0.00	.008	.03	0.02, 0.04	**< .001**	.31	0.23, 0.38	**< .001**	.01	-0.03, 0.05	.681
Conduct Problems	Model 5: MDD	-.10	-0.15, -0.06	**< .001**	-.07	-0.08, -0.06	**< .001**	.05	0.02, 0.07	**< .001**	.38	0.30, 0.45	**< .001**	.08	0.00, 0.17	**.048**
	Model 6: AD	-.10	-0.15, -0.06	**< .001**	-.07	-0.08, -0.06	**< .001**	.05	0.02, 0.07	**< .001**	.37	0.30, 0.45	**< .001**	.09	0.01, 0.16	**.029**
	Model 7: ADHD	-.11	-0.15, -0.06	**< .001**	-.07	-0.08, -0.06	**< .001**	.08	0.06, 0.11	**< .001**	.36	0.28, 0.44	**< .001**	0	-0.08, 0.08	.960
	Model 8: Combined	-.11	-0.15, -0.06	**< .001**	-.07	-0.08, -0.06	**< .001**	.04	0.03, 0.06	**< .001**	.36	0.28, 0.44	**< .001**	.04	0.00, 0.08	**.049**
Hyperactivity	Model 9: MDD	-.33	-0.37, -0.28	**< .001**	-.05	-0.06, -0.05	**< .001**	.04	0.01, 0.06	**.002**	.59	0.51, 0.67	**< .001**	.06	-0.02, 0.15	.132
	Model 10: AD	-.32	-0.37, -0.28	**< .001**	-.05	-0.06, -0.05	**< .001**	.03	0.00, 0.05	**.039**	.59	0.51, 0.67	**< .001**	0	-0.08, 0.08	.959
	Model 11: ADHD	-.33	-0.37, -0.28	**< .001**	-.05	-0.06, -0.05	**< .001**	.10	0.08, 0.13	**< .001**	.56	0.48, 0.64	**< .001**	.03	-0.05, 0.11	.498
	Model 12: Combined	-.33	-0.37, -0.28	**< .001**	-.05	-0.06, -0.05	**< .001**	.04	0.03, 0.05	**< .001**	.57	0.49, 0.65	**< .001**	.02	-0.02, 0.06	.327

Figure S7. MDD PGS interaction effects. High MDD PGS slope: β = .44, 95% CI [0.32, 0.55], $p \leq .001$, Low MDD PGS slope: β = .31, 95% CI [0.20, 0.42], $p \leq .001$.

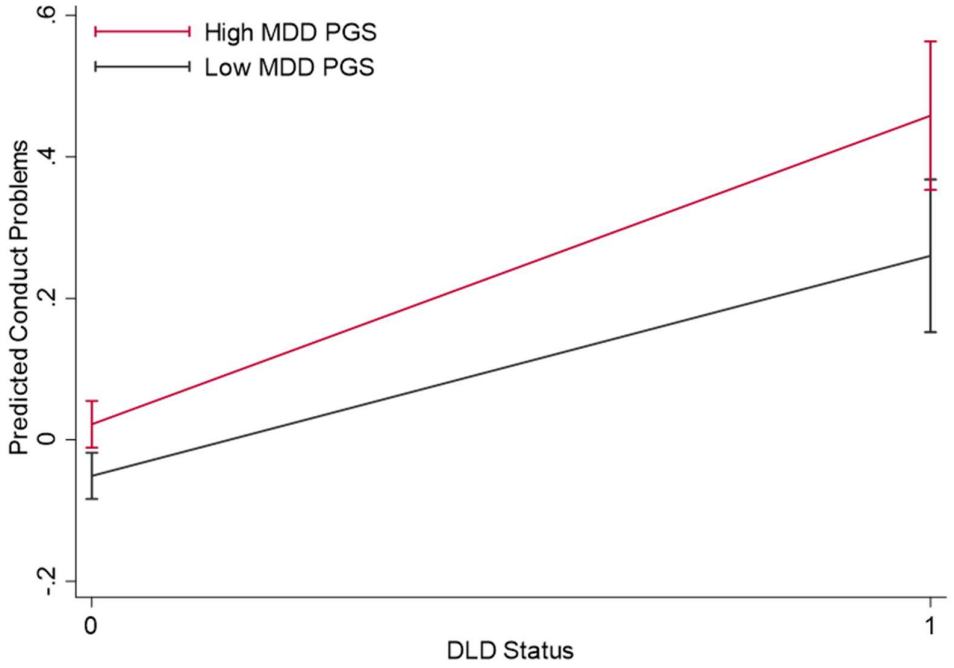


Figure S8. AD PGS interaction effects. High anxiety PGS slope: β = .46, 95% CI [0.34, 0.57], $p \leq .001$. Low anxiety PGS slope: β = .30, 95% CI [0.19, 0.41], $p \leq .001$.

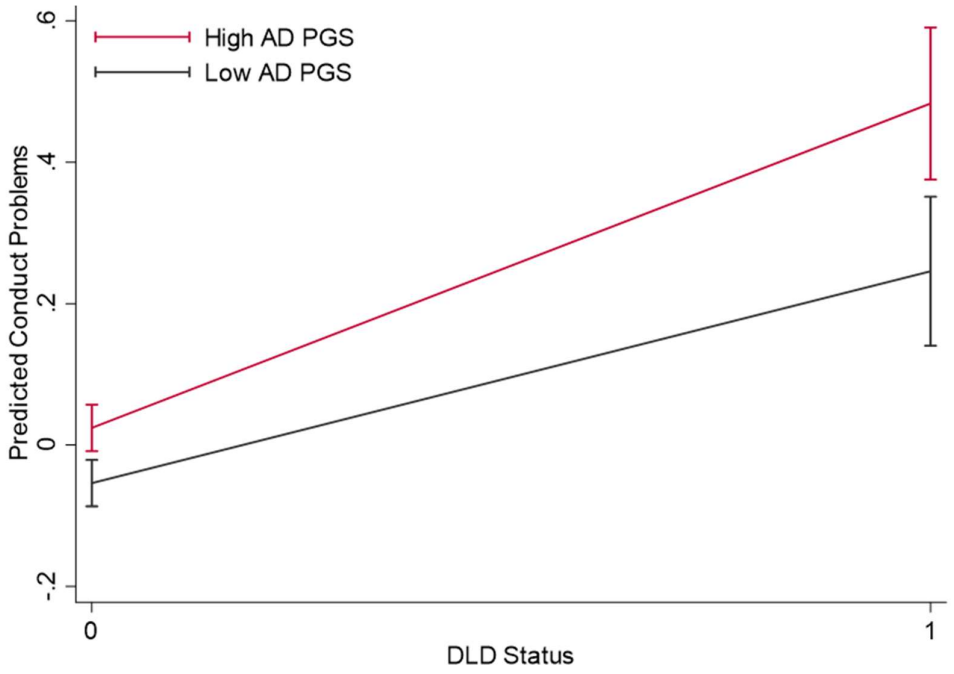
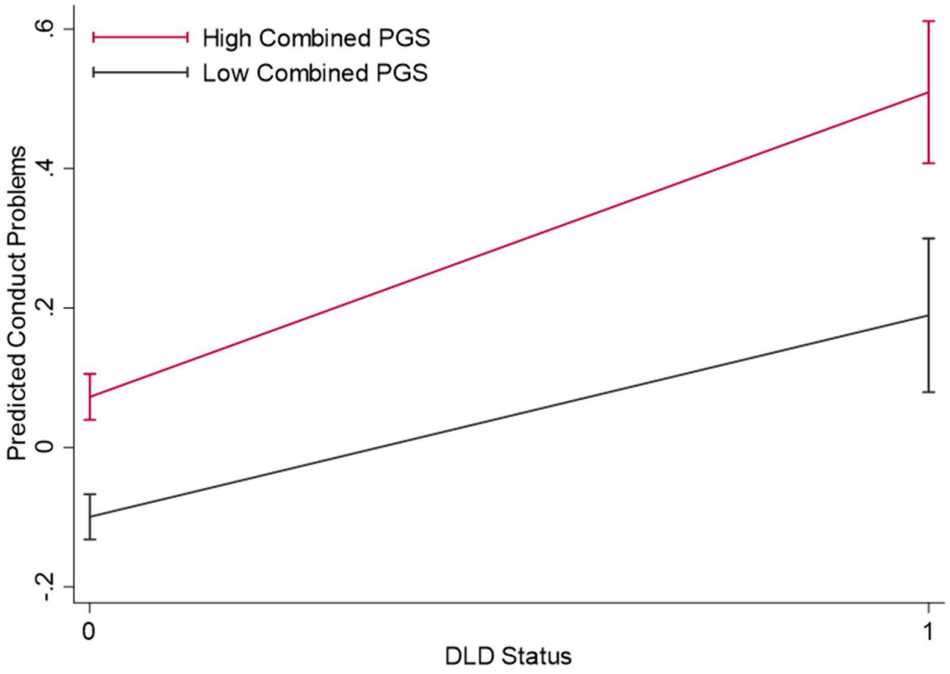


Figure S9. Combined PGS interaction effects. High combined PGS slope: β = .44, 95% CI [0.33, 0.55], $p \leq .001$. Low combined PGS slope: β = .29, 95% CI [0.18, 0.40], $p \leq .001$.



Summary of Latent Subgroups

Table S3. Emotional problems.

Number of Groups	AIC	ssa-BIC	Entropy	Smallest Class	LRT
2	79998.946	80024.036	.760	37%	< .001
3	78255.568	78294.995	.683	18%	< .001
4	**77984.264**	**78038.028**	**.581**	**16%**	**< .001**
5	77748.364	77816.465	.563	5%	.028
6	77655.531	77737.969	.521	6%	.019

Table S4. Conduct problems.

Number of Groups	AIC	ssa-BIC	Entropy	Smallest Class	LRT
2	72191.221	72216.311	.702	42%	< .001
3	**70839.056**	**70878.483**	**.669**	**13%**	**< .001**
4	70622.626	70676.390	.612	5%	.003
5	70567.771	70635.872	.636	1%	< .001
6	70576.358	70658.797	.663	0%	.939

Table S5. Hyperactivity.

Number of Groups	AIC	ssa-BIC	Entropy	Smallest Class	LRT
2	99348.996	99374.082	.765	49%	< .001
3	**97352.178**	**97391.599**	**.698**	**20%**	**< .001**
4	96901.571	96955.328	.683	7%	< .001
5	96798.613	96866.705	.644	3%	< .001
6	96788.661	96871.089	.584	3%	.298

Table S6. Genetic and environmental influences on subgroups of emotional problems (continuous combined PGS).

	Stable Low	Decreasing Within Normal Range		Increasing Within Normal Range		Consistently Raised	
		RRR [95% CI]	*p*	RRR [95% CI]	*p*	RRR [95% CI]	*p*
Model 1	**Reference**						
ELCE		0.99 [0.97, 1.02]	.645	0.99 [0.96, 1.02]	.358	0.95 [0.92, 0.98]	< .001
SES		0.91 [0.84, 0.99]	.021	0.95 [0.87, 1.04]	.292	0.83 [0.76, 0.91]	< .001
DLD		1.33 [1.00, 1.77]	.047	1.53 [1.10, 2.13]	.011	2.22 [1.65, 2.97]	< .001
Sex		1.22 [1.05, 1.41]	.008	2.21 [1.85, 2.64]	< .001	2.13 [1.79, 2.54]	< .001
PGS		1.03 [0.99, 1.07]	.103	1.03 [0.98, 1.07]	.229	1.08 [1.03, 1.13]	.001
PGSs × ELCE		1.00 [0.99, 1.02]	.508	0.98 [0.97, 1.00]	.012	1.01 [0.99, 1.02]	.449
Model 2	**Reference**						
ELCE		0.99 [0.97, 1.02]	.665	0.98 [0.96, 1.01]	.272	0.95 [0.93, 0.98]	.001
SES		0.91 [0.84, 0.99]	.021	0.95 [0.87, 1.05]	.305	0.83 [0.76, 0.91]	< .001
DLD		1.33 [1.00, 1.77]	.048	1.55 [1.12, 2.15]	.009	2.21 [1.65, 2.96]	< .001
Sex		1.22 [1.05, 1.41]	.008	2.20 [1.84, 2.63]	< .001	2.13 [1.79, 2.54]	< .001
PGS		1.03 [0.99, 1.07]	.123	1.03 [0.98, 1.07]	.232	1.07 [1.03, 1.12]	.001
PGSs × SES		1.01 [0.97, 1.05]	.721	1.00 [0.96, 1.05]	.945	1.02 [0.98, 1.07]	.333
Model 3	**Reference**						
ELCE		0.99 [0.97, 1.02]	.640	0.98 [0.95, 1.01]	.265	0.95 [0.93, 0.98]	.001
SES		0.91 [0.84, 0.99]	.021	0.95 [0.87, 1.04]	.298	0.83 [0.76, 0.91]	< .001
DLD		1.33 [1.00, 1.77]	.048	1.56 [1.12, 2.16]	.008	2.19 [1.63, 2.94]	< .001
Sex		1.22 [1.06, 1.41]	.007	2.20 [1.84, 2.63]	< .001	2.13 [1.79, 2.54]	< .001
PGS		1.04 [1.00, 1.08]	.044	1.03 [0.99, 1.08]	.172	1.07 [1.03, 1.12]	.002
PGS × DLD		0.89 [0.78, 1.03]	.111	0.94 [0.80, 1.11]	.470	0.98 [0.85, 1.13]	.813

Note. RRR = relative risk ratio; CI = confidence intervals; ELCE = early language and communication environment; SES = socioeconomic status; PGS = polygenic score; DLD = developmental language disorder.

Table S7. Genetic and environmental influences on subgroups of conduct problems (continuous combined PGS).

	Stable Low	Stable Within Normal Range		Consistently Raised	
		RRR [95% CI]	***p***	**RRR [95% CI]**	***p***
Model 4	**Reference**				
ELCE		0.96 [0.94, 0.98]	< .001	0.91 [0.88, 0.94]	< .001
SES		0.89 [0.83, 0.96]	.002	0.73 [0.66, 0.81]	< .001
DLD		1.55 [1.19, 2.02]	.001	2.4 [1.74, 3.33]	< .001
Sex		0.95 [0.83, 1.08]	.397	0.70 [0.57, 0.85]	< .001
PGS		1.09 [1.05, 1.12]	< .001	1.14 [1.08, 1.19]	< .001
PGS × ELCE		1.00 [0.99, 1.01]	.949	1.00 [0.98, 1.01]	.699
Model 5	**Reference**				
ELCE		0.96 [0.94, 0.98]	< .001	0.91 [0.88, 0.94]	< .001
SES		0.89 [0.83, 0.96]	.001	0.73 [0.66, 0.81]	< .001
DLD		1.55 [1.19, 2.02]	.001	2.41 [1.74, 3.33]	< .001
Sex		0.95 [0.83, 1.08]	.397	0.70 [0.57, 0.85]	< .001
PGS		1.09 [1.06, 1.13]	< .001	1.14 [1.08, 1.20]	< .001
PGS × SES		0.98 [0.95, 1.02]	.332	0.96 [0.92, 1.01]	.159
Model 6	**Reference**				
ELCE		0.96 [0.94, 0.98]	< .001	0.91 [0.88, 0.94]	< .001
SES		0.89 [0.83, 0.96]	.002	0.73 [0.66, 0.81]	< .001
DLD		1.55 [1.19, 2.02]	.001	2.35 [1.69, 3.27]	< .001
Sex		0.95 [0.83, 1.08]	.398	0.70 [0.57, 0.85]	< .001
PGS		1.09 [1.05, 1.13]	< .001	1.13 [1.07, 1.19]	< .001
PGS × DLD		0.98 [0.86, 1.12]	.789	1.06 [0.90, 1.25]	.464

Note. RRR = relative risk ratio; CI = confidence intervals; ELCE = early language and communication environment; SES = socioeconomic status; PGS = polygenic score; DLD = developmental language disorder.

Table S8. Genetic and environmental influences on subgroups of hyperactivity (continuous combined PGS).

	Stable Low	**Stable Within Normal Range**		**Consistently Raised**	
		RRR [95% CI]	***p***	**RRR [95% CI]**	***p***
Model 7	**Reference**				
ELCE		0.94 [0.92, 0.96]	< .001	0.89 [0.87, 0.92]	< .001
SES		0.84 [0.78, 0.91]	< .001	0.73 [0.66, 0.80]	< .001
DLD		2.14 [1.52, 3.00]	< .001	4.69 [3.28, 6.69]	< .001
Sex		0.58 [0.50, 0.66]	< .001	0.29 [0.24, 0.35]	< .001
PGS		1.08 [1.04, 1.12]	< .001	1.15 [1.10, 1.21]	< .001
PGS × ELCE		1.00 [0.99, 1.01]	.910	1.00 [0.99, 1.02]	.737
Model 8	**Reference**				
ELCE		0.94 [0.92, 0.96]	< .001	0.94 [0.92, 0.96]	< .001
SES		0.84 [0.78, 0.91]	< .001	0.84 [0.78, 0.91]	< .001
DLD		2.14 [1.53, 3.00]	< .001	2.14 [1.52, 3.00]	< .001
Sex		0.58 [0.50, 0.66]	< .001	0.58 [0.50, 0.66]	< .001
PGS		1.08 [1.04, 1.12]	< .001	1.08 [1.04, 1.12]	< .001
PGS × SES		0.99 [0.96, 1.03]	.712	1.00 [0.99, 1.01]	.910
Model 9	**Reference**				
ELCE		0.94 [0.92, 0.96]	< .001	0.89 [0.87, 0.92]	< .001
SES		0.84 [0.78, 0.91]	< .001	0.73 [0.66, 0.80]	< .001
DLD		2.12 [1.51, 2.97]	< .001	4.57 [3.20, 6.53]	< .001
Sex		0.58 [0.50, 0.66]	< .001	0.29 [0.24, 0.35]	< .001
PGS		1.08 [1.05, 1.12]	< .001	1.14 [1.09, 1.20]	< .001
PGS × DLD		0.91 [0.77, 1.08]	.291	1.00 [0.83, 1.19]	.986

Note. RRR = relative risk ratio; CI = confidence intervals; ELCE = early language and communication environment; SES = socioeconomic status; PGS = polygenic score; DLD = developmental language disorder.