Supplementary Online Content

Pingault J-B, Viding E, Galéra C, et al. Genetic and environmental influences on the developmental course of attention-deficit/hyperactivity disorder symptoms from childhood to adolescence. *JAMA Psychiatry*. Published online May 6, 2015. doi:10.1001/jamapsychiatry.2015.0469.

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix 1. Comparison of Study Sample With Initial Sample and UK Characteristics

Twin births between 1994 and 1996 were identified through birth records by the United Kingdom Office for National Statistics. A total of 16,810 families responded to the office to acknowledge their interest in taking part in the study. A first contact with these families was made by TEDS team when the twins were about 18 months. A total of 13,722 families returned data for this first contact. eTable 1 below shows the characteristics of these respondents. Of note is that these characteristics mirror closely data from the U.K. census data, so that TEDS families appear reasonably representative of the UK population. The characteristics of the 8,395 families included in the current study sample are presented in eTable1 and match closely with first contact characteristics. The biggest difference is maternal education, with 4.6 percentage points more of mothers with A-levels of education or higher in the study sample compared to the first contact sample. Further details on the sample and measures can be found in earlier publications.[1–4]

eAppendix 2. Estimation and Interpretation of Genetic and Environmental Components in ACE and ADE Models

This section presents some background for readers unfamiliar with behavioural genetics regarding the estimation of genetic and environmental influences in the classical twin design (CTD), as well as the interpretation and limitations of twin models. There is a vast literature on the topic that cannot be summarize here. We refer the readers to the following key references for additional consideration.[7–12]

Estimations of ACE and ADE univariate models

Behavioural genetics take advantage of differential genetic relationships between family members to estimate genetic and environmental influences on the variance of phenotypes. In the CTD, data of monozygotic twins (assumed to share 100% of their genes) and dizygotic twins (assumed to share 50% of their segregating genes on average) are used. The estimation of the genetic and environmental components of variance is based on the relative differences of the MZ and DZ twin within pair correlations (rMZ and rDZ) for a particular phenotype.

The different components that can be estimated in the CTD are:

- A, which represents additive genetic effects: the sum of the effects of the individual alleles at all loci that influence the trait. The genetic factors are assumed to correlate 1 in monozygotic twins and 0.5 in dizygotic twins (corresponding to 100% and 50% of shared genes).

- C, which represents the common or shared environment, defined as the environmental influences that make the twins reared in the same family similar. Because, by definition, the shared environment is the environment that is common between the two twins, shared environment factors correlate 1 between the two twins. C is assumed to have the same importance in monozygotic and dizygotic twins (i.e. similarity of MZ twins is not due to the fact that they share a higher proportion on environmental factors due to e.g. a more equal treatment based on their physical similarity).

- E, which represents the non-shared environment, i.e., the environment that make twins different. By definition, non-shared environmental factors are not correlated between the two twins. It is important to notice that E also includes measurement error.

- D, which represents non-additive genetic effects: the effects of the interactions of individual alleles either on the same or different loci that influence the trait. These genetic factors are assumed to correlate 1 in monozygotic twins (corresponding to 100% genes sharing) and 0.25 in dizygotic twins, based on the 25% chance of two siblings sharing 2 alleles Identical-by-Descent (alleles from the same parental origin).

A model with only additive genetic effects and E will predict an rDZ which is half the rMZ. The effects of C will increase the rDZ such that it is higher than .5*the rMZ, whereas the effects of D will decrease the expected rDZ such that it is lower than .5*rMZ. Since these situations are mutually exclusive, we cannot estimate both C and D in the CTD, but we fit the model (ACE or ADE) which is consistent with the observed correlations. When the correlations are consistent with an ACE model, we can use simple formulas to get an approximate value for A, C and E. For instance, let's assume that, for phenotype X, rMZ = .80 and rDZ = .50 the effects of A, C and E, referred to as a^2, c^2, and e^2 are calculated as follows:

a^2 = 2*(rMZ – rDZ) = 2*(.80 - .50) = 2*.30 = .60

The heritability of phenotype X is said to be .60, meaning that 60% of the variance can be explained by additive genetic effects. The formula can be understood as follows: because MZ and DZ twins only differ by the proportion of genes they share, the difference between MZ and DZ correlations is due to genetic effects. This difference (.30 here) corresponds to the difference in genetic relationships in MZ and DZ (100% & 50%). Therefore, multiplying this difference by 2 approximates the whole additive genetic influence on phenotype X.

c^2 = rMZ - a^2 = .80 - .60 = .20

Shared environmental influences on phenotype X equal 20%. The formula can be understood as follows: the reason why MZ twins are similar, which is manifested by the within pair correlation of .80, is because of genetic influences and shared environmental influences. Because we already calculated genetic influences, the shared environmental influences are therefore the rest of the correlation: rMZ minus genetic influences.

e^2 = 1 - rMZ = 1 - .80 = .20

Non-shared environmental influences are, by definition, what makes twin dissimilar. E is therefore calculated by subtracting the observed correlation (.80) to a perfect correlation of 1. Importantly, even if the true correlation in monozygotic twins was 1, the observed correlation

would be less than 1 because of measurement error. The CTD does not enable the distinction between true non-shared environment influence and measurement error.

These simple formulas only represent an initial approximation of the genetic and environmental effects. Structural Equation Modelling software is used to compute maximum likelihood estimates of the components based on the covariances and variances of the data, taking into account missing data.

The principle of the univariate models can be extended to include two or more phenotypes such that the MZ and DZ ratio of the cross-twin cross-phenotype correlations are used to estimate the genetic and environmental origins of the correlation between two phenotypes.

ACE and ADE models and the latent growth model.

In the present study, as described in the manuscript, the first step was to estimate a phenotypic latent growth model. This model yielded two important parameters: the intercept and the slope, which represent respectively the baseline level of the phenotype and the systematic change in that phenotype over time. Each of this parameter can be decomposed in a similar fashion to what was explained above using the within pair correlations in MZ and DZ twins. For instance, for the intercept, the MZ correlation was largely superior to the DZ correlation. The same set of rules can be applied with the only difference that they are applied to a latent factor (the intercept) rather than to a simple observed variable.

In addition, as mentioned above, the univariate model can be extended to a bivariate model, which is what we did here. Instead of two observed phenotypes, we used the two latent factors (intercept and slope) but the principles remain the same. The main aim was to verify whether some of the genetic influences detected for the slope were unique to the slope or if they were shared with the intercept (i.e. the same genes underlie both the baseline level and the systematic change in symptoms with age). Again, this is similar to a model with simple observed phenotypes, for instance to verify whether the genetic factors that underlie conduct disorders also influence addictions.

Assumptions and limitations of twin models

Several assumptions of twin models, if not met, may bias the estimates of genetic and environmental influences. These are discussed in detail elsewhere.[8,10,11] We mention below only the ones that may influence the present results.

A first set of assumptions, if not met, can bias the size of the different components, for instance:

- *Equal environment assumption.* The shared environment is assumed to have the same effect in MZ and DZ twins, i.e. to make MZ and DZ twins similar to the same extent. However, it is possible that parents, for instance, will treat MZ twins in a more similar fashion than DZ twins. If this was the case, this phenomenon would unduly increase the estimate of genetic effects, i.e. shared environmental effects would be counted as genetic effects. This assumption has been tested a number of time and seems to hold reasonably well for most traits.[7]

- *MZ twins share 100% of their genes.* Although the genetic material of MZ twins is almost identical, residual genetic differences still exist. Some important structural

differences have been found in some pairs of monozygotic twins.[13,14] If MZ twins are not 100% identical it means that some genes may contribute to differences in MZ twins. This, in turn, may underestimate genetic effects and overestimate the role of the non-shared environment. Although this is unlikely to have an effect at a population level, it cannot be excluded as yet. Of note is that genetic effects were already high in the present study, whereas the non-shared environmental influences were more limited.

Gene-environment correlations and interactions if present and not modelled may also add complexity to the interpretation of the estimates derived from twin models.

- The concept of *gene-environment correlation* relates to the fact that people are not randomly allocated to environment but, to a certain extent, their exposure to the environment varies according to the genetic make-up of their parents (passive rGE) or their own genetic make-up (active rGE). For instance, the behaviour of a hyperactive child may evoke harsh reactions from his/her parents. As such, the adverse environment – here harsh parenting – is partly dependent on the child's behaviour, which is itself partly dependent on his genetic make-up. Active rGE also explains why a lot of environmental measures are found to be heritable in genetically informative studies (e.g. life events like divorce). This does not obviously mean that genes have a direct physiological effects on these environments but rather that exposure to these environment is influenced by genes. The effects of un-modelled gene-environment correlations are included in the A parameter in the ACE model. However, this does not mean that genetic effects are overestimated but rather that some genetic effects included in the A are direct whereas others are indirect. For instance, some genetic variants have been shown to be associated with smoking severity. As such, these genetic variants have an effect on an environmental exposure, i.e. smoking severity. In turn, smoking causes lung cancer. When examining the heritability of lung cancer, the effect of those genes will be taken into account in the genetic estimate. However, these genetic variants do not affect the physiology of lung cells in a direct fashion, which is shown by the fact that they are not associated with cancer in non-smoking patient. Instead, they increase an environmental exposure that, in turn, causes cancer. In this case, the effect of these genetic variants is correctly classified in genetic effects as the ultimate origin of the cascade of events leading to cancer is genetic. However, the proximal cause of the cancer is rather environmental in origins. The twin model cannot distinguish between these genetic effects without measuring environmental factors that lie on the pathway from genes to phenotype. As such, it is important to keep in mind that part of the genetic effects detected in the present study may follow this indirect pattern of influence.

- *Gene-environment interactions* relate to the fact that sometimes the expression of genetic effects depend on the environmental exposure. For instance, susceptibility to skin cancer can be partly genetic in origin. However, these genetic variants will not express if there is no sufficient exposure to the sun, i.e. if the environmental exposure is not present. In other words, if only one factor – genes or the environment – is present, the disease will not manifest but it will if both of them are. Gene-environment interactions, if present and un-modelled, end up being represented in either genetic (A) or non-shared environment (E) estimates. If the interaction of genetic effects is with the non-shared environment (E), this will inflate the E estimate. Conversely, interactions of genetic influences with the shared environment (C), will inflate the genetic (A) estimate. Although these interaction effects

cannot be modelled in the classical twin design, other methods can give us a sense of their existence and if the assumption of additivity of A, C and E is reasonable or not.

In spite of their limitations, twin models, together with other designs in quantitative genetics, can still give us invaluable information on the genetic and environmental architecture of phenotypes and represent a helpful starting point for limiting the problem space when searching for likely developmental mechanisms (for a recent review, see[8]). This is all the more the case as many of the aforementioned assumptions and limitations can be addressed by using complementary research designs (including the DNA-based estimates of heritability based on Genome-Wide Complex Trait Analysis[15]).

eTable 1. Sample Characteristics

	Returned data (N families)	% White	% Mothers with A-levels or higher	% Mother employed	% Father employed	% Female	% MZ
UK census[1]	-	*93%*	*32%*	*49%*	*89%*	-	-
TEDS first contact	13,722	91.7%	35.5%	43.1%	91.6%	50.1%	33.2%
TEDS study sample	8,395	93.1%	40.1%	46.2%	93.0%	51.4%	34.7%

Note. [1]UK data from the 2000 General Household Survey[5] are used rather than more recent data because they provide more appropriate comparisons for TEDS twins who were born 1994-96. The % MZ data are from Imaizumi[6] because they are not available from UK census data. A-levels are the national educational exam taken at 18 years of age in the UK. MZ=monozygotic twins.

eTable 2. Fit Indices

For each model, we report the Akaike-Information Criterion (AIC) and the chi-square, as well as additional approximate fit indexes and three approximate fit indexes: CFI (Comparative Fit Index) for which values close to 1 indicate better fit; RMSEA (Root Mean Square Error of Approximation, and 90% Confidence Interval) and SRMR (Standardized Root Mean Square Residual) for which values close to 0 indicate better fit. No consensus exists on cut-off values for these indexes, but values close to .95 for CFI, 0.06 for RMSEA and 0.08 for SRMR have been suggested[16] to conclude that there is a relatively good fit between the model and the data.

		AIC	Chi Square	DF	pvalue	CFI	SRMR	RMSEA	RMSEA lower	RMSEA upper
Hyperactivity Impulsivity	Cholesky ACE	215218	85.377	54.000	0.004	0.997	0.046	0.012	0.008	0.015
	Cholesky ADE	215230	83.286	54.000	0.006	0.997	0.044	0.011	0.008	0.015
	LGC phenotypic	215447	210.593	65.000	0.000	0.986	0.055	0.023	0.021	0.026
	LGC ACE	215455	200.700	65.000	0.000	0.987	0.055	0.022	0.020	0.025
	LGC ADE	215475	211.385	65.000	0.000	0.985	0.055	0.023	0.021	0.026
Inattention	Cholesky ACE	232708	156.603	54.000	0.000	0.990	0.065	0.021	0.018	0.024
	Cholesky ADE	232619	112.280	54.000	0.000	0.995	0.048	0.016	0.013	0.019
	LGC phenotypic	233097	429.262	65.000	0.000	0.966	0.068	0.037	0.034	0.039
	LGC ACE	233199	458.148	65.000	0.000	0.963	0.077	0.038	0.035	0.040
	LGC ADE	233112	406.384	65.000	0.000	0.968	0.067	0.035	0.033	0.038

Note. The degrees of freedom (DF) for ACE and ADE models are equal. Therefore, the choice between ACE and ADE models for Cholesky models on the one hand and for LGC models on the other hand was based on the Akaike-Information Criterion (AIC), with lower values indicating better fit. Approximate fit indexes were useful to assess whether the phenotypic Latent Growth Model (LGC) fitted the model adequately. All indexes were in line with recommended cut-offs, showing that a linear growth model fitted the data adequately.

eTable 3. Correlation Table, Hyperactivity/Impulsivity Score

	DZ / MZ	TWIN 1				TWIN 2					
		4 years	7 years	12 years	16 years	4 years	7 years	12 years	16 years	MEAN (DZ)	SD (DZ)
TWIN 1	4 years	-	0.667	0.579	0.453	0.424	0.289	0.224	0.191	5.748	5.038
	7 years	0.703	-	0.681	0.581	0.280	0.452	0.271	0.254	4.281	4.318
	12 years	0.546	0.663	-	0.698	0.237	0.284	0.378	0.259	3.640	4.210
	16 years	0.488	0.596	0.647	-	0.172	0.234	0.267	0.409	2.763	3.692
TWIN 2	4 years	0.870	0.640	0.516	0.465	-	0.672	0.560	0.499	5.781	5.131
	7 years	0.632	0.868	0.595	0.506	0.698	-	0.694	0.595	4.357	4.492
	12 years	0.469	0.589	0.838	0.568	0.543	0.670	-	0.692	3.671	4.329
	16 years	0.439	0.524	0.595	0.784	0.479	0.566	0.674	-	2.662	3.595
	MEAN (MZ)	6.100	4.385	3.357	2.508	6.044	4.331	3.435	2.483		
	SD (MZ)	5.162	4.393	3.756	3.339	5.017	4.272	3.918	3.257		

Note. The table shows the observed Pearson pairwise correlations within and across time, within and between twins for MZ (lower part of the table) and DZ (upper part), as well as the observed means and standard deviations (SD). Values in grey are twin correlations at each time point: heritability and environmental estimates at each time point are based on the comparison of each pair of MZ and DZ correlations. Values in yellow represent the across time correlations for each twin, showing the phenotypic continuity in hyperactivity/impulsivity scores. Values in blue represent cross-twin cross-time correlations: the respective role of genes and the environment in explaining hyperactivity/impulsivity is estimated based on the comparison of these correlations between MZ and DZ.

eTable 4. Correlation Table, Inattention Score

	DZ / MZ	TWIN 1				TWIN 2					
		4 years	7 years	12 years	16 years	4 years	7 years	12 years	16 years	MEAN (DZ)	SD (DZ)
TWIN 1	4 years	-	0.656	0.562	0.493	0.291	0.211	0.169	0.168	5.529	5.238
	7 years	0.674	-	0.668	0.606	0.184	0.330	0.232	0.208	5.578	5.054
	12 years	0.535	0.704	-	0.697	0.194	0.231	0.333	0.231	5.157	5.045
	16 years	0.471	0.605	0.706	-	0.147	0.188	0.230	0.326	4.515	5.082
TWIN 2	4 years	0.786	0.565	0.448	0.425	-	0.642	0.560	0.491	5.515	5.195
	7 years	0.553	0.752	0.547	0.457	0.633	-	0.699	0.594	5.712	5.247
	12 years	0.438	0.600	0.775	0.539	0.523	0.684	-	0.708	5.424	5.389
	16 years	0.390	0.475	0.544	0.711	0.475	0.577	0.668	-	4.459	4.930
	MEAN (MZ)	5.455	5.683	4.946	3.867	5.423	5.500	4.838	3.831		
	SD (MZ)	5.133	5.028	4.803	4.542	5.008	4.815	4.770	4.372		

Note. See eTable2.

eTable 5. Characteristics of Children With Increasing Inattention vs Others

	Increasing inattention % or Mean (SD)	Other participants % or Mean (SD)	Effect sizes[1]	P-values
% Female	46.4	52.7	0.05	< .001
% White	93.6	92.9	0.01	= .17
% Mothers with A-levels or higher	39.4	41.4	0.02	=.03
% Mother employed	44.7	46.6	0.02	=.15
GCSE Maths	8.73 (1.47)	9.00 (1.47)	0.19	< .001
GCSE English	8.89 (1.34)	9.18 (1.29)	0.22	< .001
GCSE Science	8.88 (1.52)	9.17 (1.51)	0.20	< .001
G: 7 years	0.03 (0.98)	0.03 (0.99)	0.00	= .94
G: 9 years	-0.02 (1.02)	0.01 (0.99)	0.03	= .35
G: 10 years	-0.04 (1.01)	0.02 (0.99)	0.07	= .10
G: 12 years	-0.14 (1.05)	0.04 (0.98)	0.18	< .001
G: 14 years	-0.22 (1.05)	0.05 (0.98)	0.27	< .001
G: 16 years	-0.16 (0.96)	0.04 (1.00)	0.20	< .001
Conners Inattention Teachers				
14 Years	6.22 (6.66)	3.93 (5.41)	-0.40	< .001

Note. [1]Hedges g (an equivalent of Cohen d for unequal sample sizes) was used for socioeconomic status. For other variables a phi-coefficient is presented. See eTable 1 for an explanation of the first four variables. The General Certificate of Secondary Education (GCSE) is taken by more than 99% of pupils at the end of compulsory education, which is typically around the age of 16 years. English, mathematics and science are compulsory subjects. The grades were coded from 11 (the highest grade) to 4 (the lowest pass grade). G is a standardized score of general cognitive ability that was assessed at each age using a battery of parent-administered and phone- and web-based tests. At each testing age, individuals completed at least two ability tests, which assessed verbal and non-verbal intelligence. For additional details on these measures, see[17,18]. Results for the Conners inattention scale rated by teachers instead of mothers are also presented.

Comment: eTable 5 provides additional information on children with increasing inattention represented in eFigure 1 above. These children differed very little from other children on sex, ethnicity, maternal education and maternal employment, although they were slightly more likely to be males and to have less educated mothers. However, at the end of the follow-up, they had significantly lower educational achievement, as measured by GCSE at age 16 years. In addition, while they did not differ on general cognitive abilities at baseline, differences widened with age and very significant differences between the two groups were observed after age 12 years. These differences seem consistent with the existence of specific developmental processes that become evident with advancing age and concern inattention as well as other cognitive characteristics. Finally, the score on the Conners inattention scale rated by teachers at age 14 years was higher for children with increasing inattention.

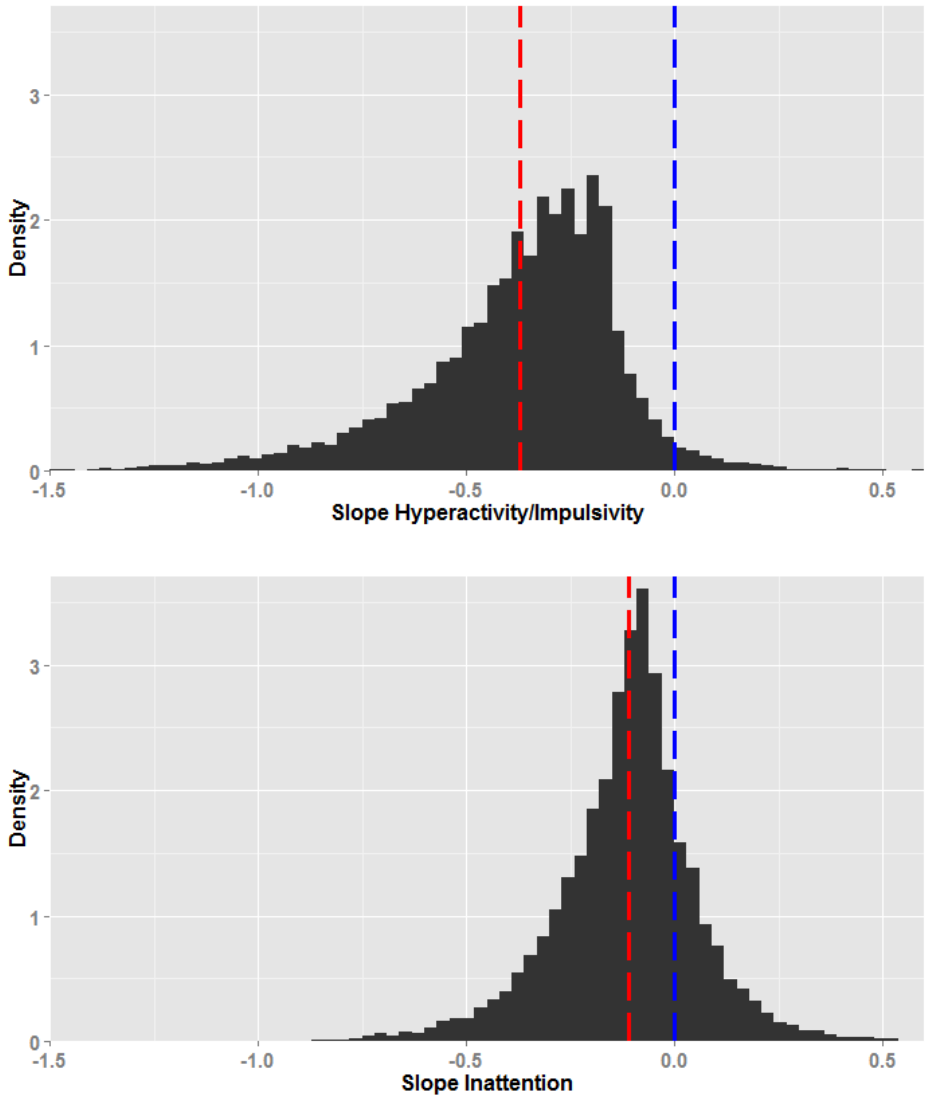
The results presented in eTable 5 are presented for descriptive purposes and caution should be applied when interpreting them. A latent growth model was used to model systematic change in a dimensional score of inattention. Systematic change – here the slope of inattention – was therefore modelled in a continuous fashion as represented in eFigure 1. Consequently, separating the sample into two sub-groups following different developmental trajectories based on the slope value is arbitrary as it cuts in a continuous distribution (here below and above 0). In addition, such a cut-off based on the slope does not take into account the baseline values (i.e. the intercept) to constitute the groups. Although outside the scope of the present study, more appropriate methods can be applied to examine sub-groups following different developmental trajectories.[19–21]

eTable 6. Genetic and Environmental Influences on the Residuals From the Latent Growth Models

RESIDUALS HYPERACTIVITY	Age	7.9	11.3	14.1	16.3
	Total	0.24	0.29	0.32	0.16
	A	.73 (.52-.85)	.63 (.49-.77)	.80 (.70-.85)	.41 (.00.-.70)
	C	.07 (.00-.23)	.15 (.04-.27)	.00 (.00-.28)	.14 (.00-.45)
	E	.21 (.15-.29)	.22 (.18-.27)	.21 (.15-.27)	.45 (.28-.66)
RESIDUALS INATTENTION	Total	0.26	0.33	0.30	0.22
	A	.66 (.49-.77)	.61 (.55-.66)	.72 (.58-.79)	.70 (.60-.79)
	D	.00 (.00-.52)	.00 (.00-.22)	.00 (.00-.36)	.00 (00-.49)
	E	.34 (.25-.45)	.39 (.34-.45)	.28 (.22-.35)	.30 (.21-.40)

Note. Latent growth factors – intercept and linear slope – explained a large amount of variance at each age, with remaining total residual variances between .16 and .33 (lines Total). These residual variances were decomposed into genetic and environmental influences: ACE decomposition for hyperactivity/impulsivity and the ADE decomposition for inattention.

eFigure. Distributions of the Slope for Hyperactivity/Impulsivity and Inattention



The Figure represents the distributions of the predicted linear slopes of hyperactivity/impulsivity and inattention symptoms. The red line represents the mean *yearly* slope, which is negative for hyperactivity/impulsivity (-0.37), meaning that, for each year during the 8 years follow-up, the mean score of Hyperactivity/impulsivity decreased by 0.37 point (the slope being -0.11 for inattention),. The blue line is zero so that participants on the right side of the blue line have a positive slope. As can be seen, this is much more frequent for inattention (≈ 20%) than for hyperactivity/Impulsivity (≈ 3%).

eReferences

1. Kovas Y, Haworth CMA, Dale PS, Plomin R. The genetic and environmental origins of learning abilities and disabilities in the early school years. *Monogr Soc Res Child Dev*. 2007;72(3):vii, 1-144. doi:10.1111/j.1540-5834.2007.00439.x.

2. Oliver BR, Plomin R. Twins' Early Development Study (TEDS): a multivariate, longitudinal genetic investigation of language, cognition and behavior problems from childhood through adolescence. *Twin Res Hum Genet*. 2007;10(1):96-105.

3. Trouton A, Spinath FM, Plomin R. Twins early development study (TEDS): a multivariate, longitudinal genetic investigation of language, cognition and behavior problems in childhood. *Twin Res Off J Int Soc Twin Stud*. 2002;5(5):444-448. doi:10.1375/136905202320906255.

4. Haworth CMA, Davis OSP, Plomin R. Twins Early Development Study (TEDS): a genetically sensitive investigation of cognitive and behavioral development from childhood to young adulthood. *Twin Res Hum Genet Off J Int Soc Twin Stud*. 2013;16(1):117-125. doi:10.1017/thg.2012.91.

5. ONS. *Living in Britain: Results from the 2000 General Household Survey.* London: Stationery Office; 2001.

6. Imaizumi Y. A comparative study of zygotic twinning and triplet rates in eight countries, 1972-1999. *J Biosoc Sci*. 2003;35(02):287-302. doi:10.1017/S0021932003002876.

7. Plomin R, DeFries JC, Knopik VS, Neiderhiser JM. *Behavioral Genetics*. Sixth Edition. New York, NY: Worth Publishers; 2013.

8. Tenesa A, Haley CS. The heritability of human disease: estimation, uses and abuses. *Nat Rev Genet*. 2013;14(2):139-149. doi:10.1038/nrg3377.

9. Martin N, Boomsma D, Machin G. A twin-pronged attack on complex traits. *Nat Genet*. 1997;17(4):387-392. doi:10.1038/ng1297-387.

10. Johnson W, Penke L, Spinath FM. Heritability in the era of molecular genetics: some thoughts for understanding genetic influences on behavioural traits. *Eur J Personal*. 2011;25(4):254-266. doi:10.1002/per.836.

11. Rijsdijk FV, Sham PC. Analytic approaches to twin data using structural equation models. *Brief Bioinform*. 2002;3(2):119-133.

12. Neale MC, McArdle JJ. Structured latent growth curves for twin data. *Twin Res Off J Int Soc Twin Stud*. 2000;3(3):165-177.

13. Ehli EA, Abdellaoui A, Hu Y, et al. De novo and inherited CNVs in MZ twin pairs selected for discordance and concordance on Attention Problems. *Eur J Hum Genet EJHG*. 2012;20(10):1037-1043. doi:10.1038/ejhg.2012.49.

14. Machin G. Non-identical monozygotic twins, intermediate twin types, zygosity testing, and the non-random nature of monozygotic twinning: a review. *Am J Med Genet C Semin Med Genet*. 2009;151C(2):110-127. doi:10.1002/ajmg.c.30212.

15. Plomin R. Genetics: How intelligence changes with age. *Nature*. 2012;482(7384):165-166. doi:10.1038/482165a.

16. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Model Multidiscip J*. 1999;6(1):1-55.

17. Krapohl E, Plomin R. Genetic link between family socioeconomic status and children's educational achievement estimated from genome-wide SNPs. *Mol Psychiatry*. In Press.

18. Krapohl E, Rimfeld K, Shakeshaft NG, et al. The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proc Natl Acad Sci U S A*. 2014;111(42):15273-15278. doi:10.1073/pnas.1408777111.

19. Pingault J-B, Côté SM, Vitaro F, Falissard B, Genolini C, Tremblay RE. The developmental course of childhood inattention symptoms uniquely predicts educational attainment: A 16-year longitudinal study. *Psychiatry Res*. 2014;219(3):707-709. doi:10.1016/j.psychres.2014.06.022.

20. Larsson H, Dilshad R, Lichtenstein P, Barker ED. Developmental trajectories of DSM-IV symptoms of attention-deficit/hyperactivity disorder: genetic effects, family risk and associated psychopathology. *J Child Psychol Psychiatry*. 2011;52(9):954-963. doi:10.1111/j.1469-7610.2011.02379.x.

21. Genolini C, Pingault J-B, Driss T, et al. KmL3D: a non-parametric algorithm for clustering joint trajectories. *Comput Methods Programs Biomed*. 2013;109(1):104-111. doi:10.1016/j.cmpb.2012.08.016.