Supplementary Online Materials

Polygenic score for educational attainment captures DNA variants shared between personality traits and educational achievement

Authors:

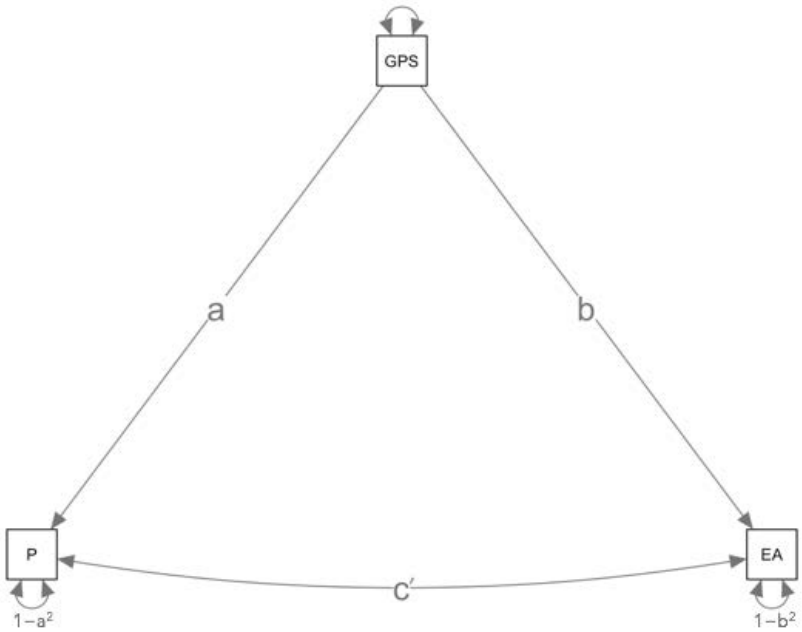
Emily Smith-Woolley[1*], Saskia Selzam[1*], Robert Plomin[1]

*Joint first authors

Supplementary Methods

Methods S1 – Structural equation model of personality domains, GCSE results and polygenic scores

To test the extent to which the covariance between personality domains and GCSE results are explained by the polygenic scores, we used structural equation modelling. Because we assume causality from polygenic score (an aggregate score of DNA variants) to outcome variables, we applied the following model to our z-standardized variables:



Note. P = personality trait, EA = educational achievement defined as GCSE results, GPS = genome-wide polygenic score

Paths a, b and c are the beta effect size parameters. The GPS effect is described by the product of a and b, which is the pathway from the causal variable GPS to P, and GPS to EA. Path c' describes the residual relation between P and EA after accounting for the effects of the causal variable GPS in P and EA, respectively. The total effect can be derived by summing the effects of the residual and the indirect path, described as c' + ab. To calculate the proportion of the total effect that is explained by the causal variable GPS, the effect of the indirect path ab is divided by the total effect c' + ab.

Supplementary Tables

Table S1 – Descriptive statistics of all variables for the full sample

		Mean (SD)			ANOVA of gender		ANOVA of age	
	N	Whole sample	Males	Females	*F*	R^2	*F*	R^2
Academic motivation Composite	**3079**	**0.00 (1.00)**	**-0.02 (0.97)**	**0.02 (1.02)**	**-**	**-**	**-**	**-**
PISA math self-efficacy	3078	17.53 (5.51)	18.94 (5.00)	16.55 (5.64)	99.66***	0.05	0.08	<0.01
PISA math interest	3079	2.53 (0.93)	2.65 (0.90)	2.45 (0.95)	22.9***	0.01	2.89	<0.01
PISA time spent on math	3050	4.48 (1.71)	4.48 (1.83)	4.48 (1.63)	0.00	<0.01	0.42	<0.01
Attitude towards key subjects	3078	2.54 (0.49)	2.52 (0.48)	2.56 (0.49)	2.5	<0.01	1.08	<0.01
Openness Composite	**2881**	**0.00 (1.00)**	**-0.02 (0.96)**	**0.01 (1.03)**	**-**	**-**	**-**	**-**
Academic self-concept	2837	3.55 (0.63)	3.65 (0.59)	3.48 (0.64)	35.53***	0.02	0.14	<0.01
Ambition	2837	3.89 (0.67)	3.9 (0.64)	3.88 (0.69)	0.55	<0.01	0.01	<0.01
Curiosity	3221	4.77 (0.9)	4.88 (0.88)	4.70 (0.91)	23.18***	0.01	0.06	<0.01
Hopefulness	3227	4.69 (0.72)	4.78 (0.69)	4.63 (0.73)	23.94***	0.01	1.96	<0.01
Openness	2803	3.58 (0.58)	3.53 (0.58)	3.62 (0.57)	9.11***	<0.01	1.78	<0.01
GRIT	2887	3.26 (0.59)	3.21 (0.57)	3.30 (0.61)	11.91***	0.01	0.91	<0.01
Optimism	2887	3.23 (0.71)	3.28 (0.69)	3.19 (0.72)	8.89***	<0.01	1.52	<0.01
Conscientiousness Composite	**2713**	**0.02 (0.99)**	**0.02 (0.98)**	**0.02 (0.99)**	**-**	**-**	**-**	**-**
SDQ Hyperactivity	6001	6.45 (2.30)	6.38 (2.30)	6.50 (2.30)	2.96	<0.01	2.30	<0.01
SWAN Hyperactivity	1313	4.78 (0.99)	4.8 (0.94)	4.77 (1.01)	0.19	<0.01	0.00	<0.01
SWAN Inattention	1313	4.64 (0.87)	4.62 (0.88)	4.66 (0.86)	0.31	<0.01	1.74	<0.01
Conscientiousness	2796	3.71 (0.60)	3.62 (0.600)	3.77 (0.59)	27.52***	0.01	0.00	<0.01
Agreeableness Composite	**6612**	**-0.01 (0.99)**	**-0.02 (1.01)**	**0.00 (0.98)**	**-**	**-**	**-**	**-**
Agreeableness	2798	3.66 (0.56)	3.53 (0.55)	3.75 (0.56)	74.16***	0.04	0.13	<0.01
SDQ Prosocial behavior	6001	7.14 (1.93)	6.52 (1.92)	7.63 (1.78)	357.53***	0.08	0.96	<0.01
School engagement	3068	3.00 (0.68)	2.98 (0.68)	3.02 (0.67)	2.16	<0.01	0.26	<0.01
SDQ Conduct scale	6000	8.39 (1.44)	8.29 (1.48)	8.46 (1.41)	13.52***	<0.01	3.09	<0.01
Gratitude	3229	5.8 (0.84)	5.65 (0.84)	5.91 (0.83)	51.36***	0.02	0.10	<0.01

Neuroticism Composite	**6008**	**-0.01 (1.01)**	**-0.02 (0.90)**	**0.00 (1.08)**	**-**	**-**	**-**	**-**
Cognitive Disorganization	5998	3.93 (2.86)	3.37 (2.73)	4.38 (2.88)	126.06***	0.03	1.26	<0.01
CASI anxiety	6004	8.07 (5.96)	6.20 (4.85)	9.53 (6.34)	332.69***	0.08	0.38	<0.01
MFQ	6003	3.59 (4.33)	2.56 (3.29)	4.39 (4.85)	183.36***	0.04	2.32	<0.01
Subjective happiness	5998	2.88 (0.94)	2.89 (0.91)	2.87 (0.97)	0.18	<0.01	3.80	<0.01
Life satisfaction	3224	2.40 (0.62)	2.38 (0.58)	2.41 (0.64)	1.76	<0.01	3.21	<0.01
Peer problems	6001	1.54 (1.51)	1.60 (1.53)	1.48 (1.49)	6.61*	<0.01	4.06*	<0.01
Neuroticism	2808	2.58 (0.68)	2.47 (0.64)	2.66 (0.69)	35.82***	0.02	6.87*	<0.01
Extraversion	**2807**	**0.00 (0.97)**	**-0.01 (0.97)**	**0.01 (0.97)**	**0.32**	**<0.01**	**0.69**	**<0.01**
GCSE	**8322**	**-0.01 (1.00)**	**-0.02 (1.02)**	**0.00 (0.99)**	**0.79**	**<0.01**	**1.71**	**<0.01**
general cognitive ability	**3939**	**0.01 (0.99)**	**0.00 (0.98)**	**0.01 (0.99)**	**-**	**-**	**-**	**-**
EduYears GPS	**10346**	**0.01 (1.00)**	**0.02 (1.01)**	**0.00 (0.99)**	**1.08**	**<0.01**	**0.16**	**<0.01**
Neuroticism GPS	**10346**	**0.01 (1.00)**	**0.01 (1.01)**	**0.01 (0.99)**	**0.02**	**<0.01**	**0.10**	**<0.01**
Wellbeing GPS	**10346**	**0.00 (1.00)**	**0.00 (0.99)**	**0.00 (1.01)**	**0.00**	**<0.01**	**0.01**	**<0.01**

Note: Means and standard deviations for individual measures are calculated based on raw data. Means and standard deviations for domains are calculated with z-standardised age and sex regressed data. Values of standard deviation are given in parentheses. $^{+}$ = standardization of the individual cognitive scales assessed at age 7, 12 and 16 was required to form this composite. N= sample size after exclusions. For DZ twin pairs, ANOVA performed on one randomly selected twin per pair to test the effect of sex and age. Results = F statistic; * = p<.05; ** = p<.01; R^2= proportion of variance explained.

Table S2 – Sensitivity analysis of missingness of personality/motivation composites on socio-economic status, general cognitive ability, and GCSE grades

GCSE					
	Present M (SD)	**Missing M (SD)**	***z***	***p***	***R^2***
Academic motivation	0.14 (0.97)	-0.09 (1.01)	8.01	6.63E-16	0.009
Openness Composite	0.15 (0.96)	-0.09 (1.01)	8.46	1.31E-17	0.010
Conscientiousness Composite	0.17 (0.98)	-0.09 (1.00)	8.69	1.73E-18	0.011
Agreeableness Composite	0.10 (0.98)	-0.30 (1.01)	12.73	5.77E-39	0.026
Neuroticism Composite	0.12 (0.98)	-0.26 (1.00)	13.07	6.06E-41	0.025
Extraversion	0.17 (0.95)	-0.09 (1.01)	9.23	1.00E-20	0.013

General cognitive ability					
	Present M (SD)	**Missing M (SD)**	***z***	***p***	***R^2***
Academic motivation	0.03 (0.98)	-0.05 (0.99)	1.96	0.05	0.001
Openness Composite	0.03 (0.99)	-0.05 (0.99)	2.09	0.04	0.001
Conscientiousness Composite	0.09 (0.97)	-0.11 (1.00)	5.16	1.98E-07	0.007
Agreeableness Composite	0.03 (0.98)	-0.16 (1.03)	3.22	1.24E-3	0.005
Neuroticism Composite	0.06 (0.98)	-0.17 (1.00)	5.26	1.18E-07	0.009
Extraversion	0.05 (0.97)	-0.08 (1.01)	3.41	6.16E-4	0.003

SES					
	Present M (SD)	**Missing M (SD)**	***z***	***p***	***R^2***
Academic motivation	0.2 (0.98)	0.04 (1.01)	5.91	3.09E-09	0.004
Openness Composite	0.21 (0.98)	0.04 (1.00)	6.23	3.96E-10	0.005
Conscientiousness Composite	0.22 (0.98)	0.04 (1.00)	6.26	3.20E-10	0.005
Agreeableness Composite	0.21 (0.98)	-0.14 (1.00)	13.47	7.88E-43	0.022
Neuroticism Composite	0.23 (0.98)	-0.12 (1.00)	13.85	3.02E-45	0.022
Extraversion	0.22 (0.98)	0.04 (1.01)	6.67	2.07E-11	0.006

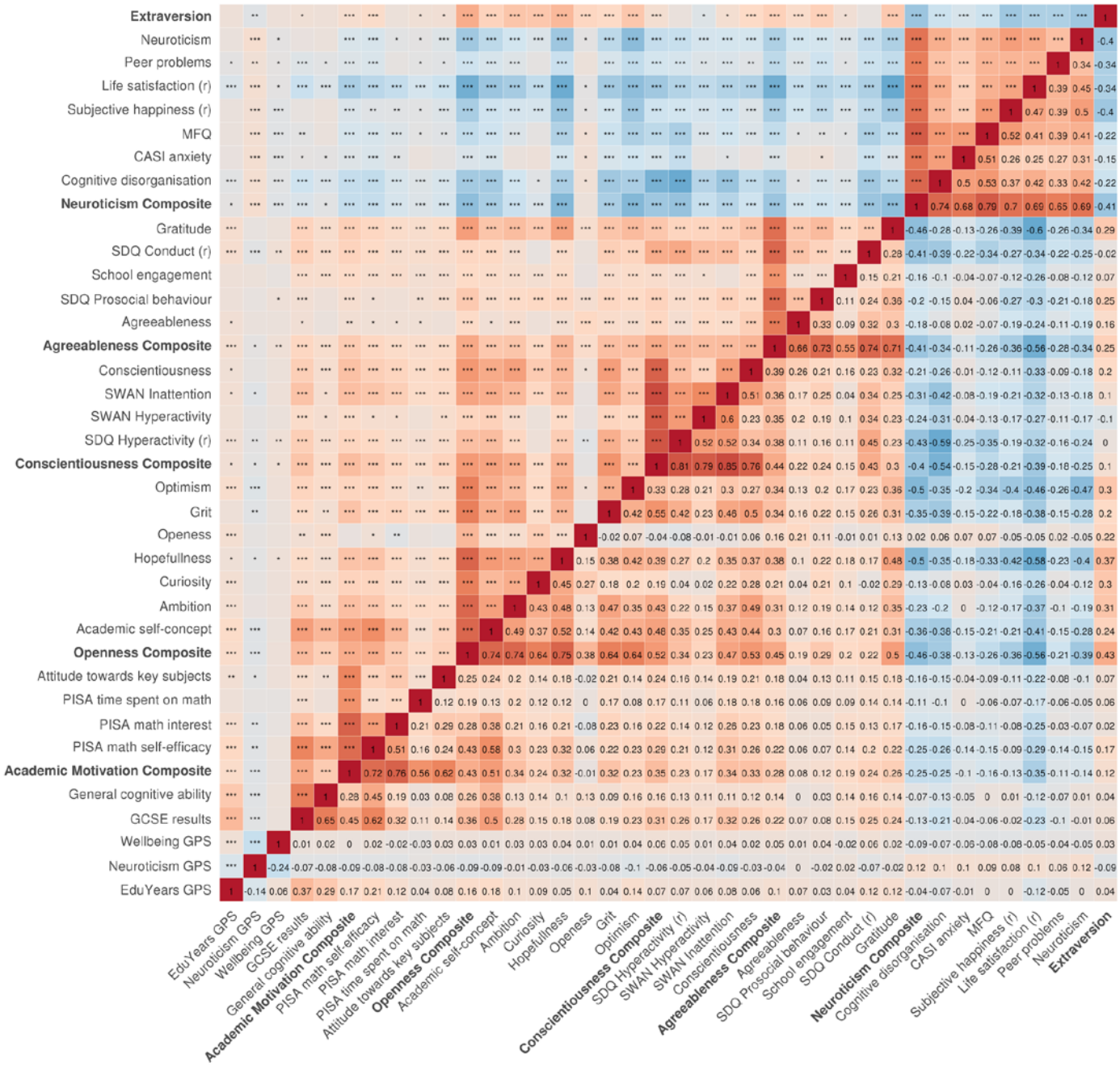
Note: SES = socio-economic status; GCSE = General Certificate for Secondary Education; present = data is non-missing for the respective composite; missing = data is missing for the respective composite. Analyses were performed on unrelated individuals only, where one twin per twin pair was randomly selected if data was available on dizygotic twin pairs. Degrees of freedom (df) for GCSE analyses = 5,600; df for general cognitive ability analyses = 2,676; df for SES analyses = 6,569.

Table S3 – Descriptive statistics of all variables for the sample of unrelated individuals

	N	*Mean*	*SD*	*Skew*	*Min*	*Max*
Academic motivation Composite	2084	0.00	1.00	-0.57	-4.4	2.49
Openness Composite	1958	0.00	1.00	-0.21	-4.48	3.12
Conscientiousness Composite	1838	0.02	0.99	-0.24	-3.52	2.6
Agreeableness Composite	4415	-0.01	0.99	-0.62	-4.56	2.62
Neuroticism Composite	4005	-0.01	1.01	0.89	-2.14	5.37
Extraversion	1909	0.00	0.97	-0.42	-3.86	2.26
GCSE	5602	-0.01	1.00	-0.43	-3.63	1.87
EduYears GPS	7026	0.01	1.00	0.08	-3.49	3.71
Neuroticism GPS	7026	0.01	1.00	0.04	-5.39	4.34
Wellbeing GPS	7026	0.00	1.00	0.01	-3.96	3.78

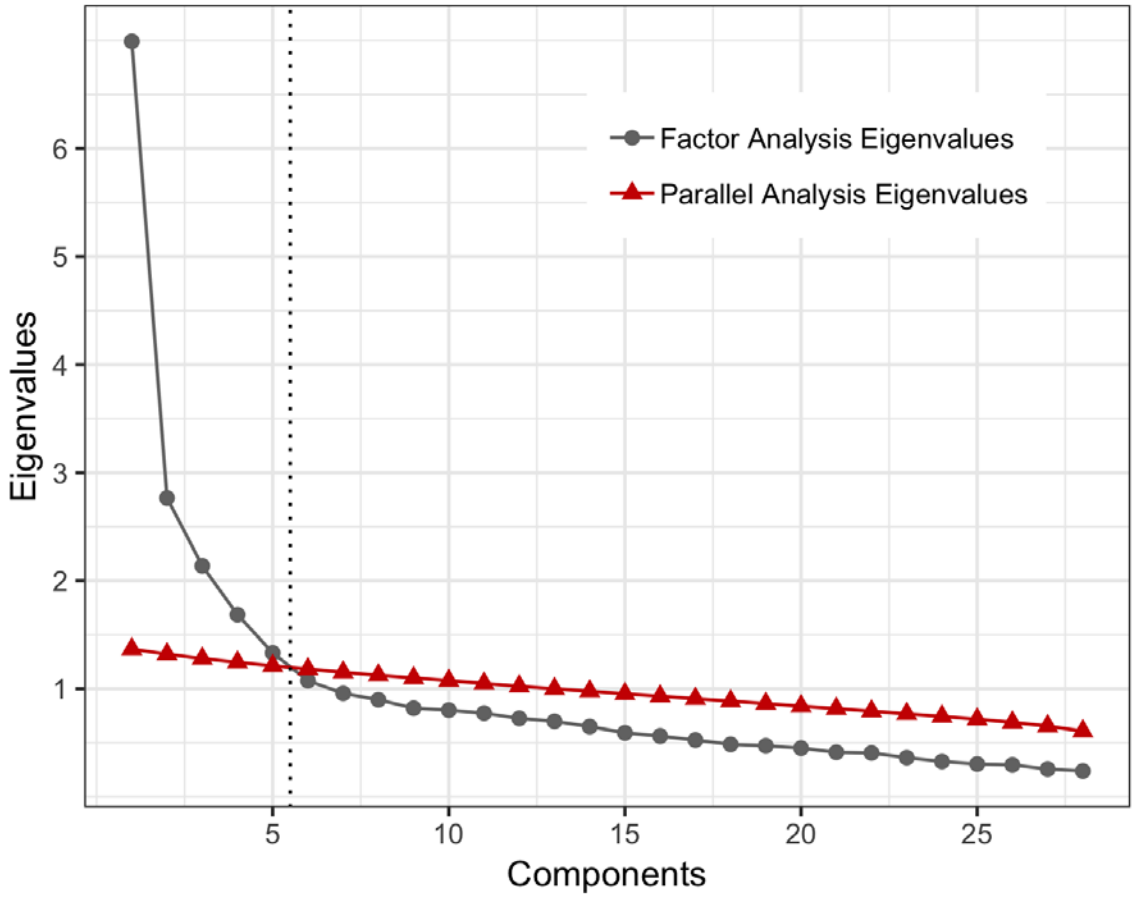
Supplementary Figures

Figure S1 – Correlations across all individual measures of personality and motivation, the personality/motivation composites and polygenic scores



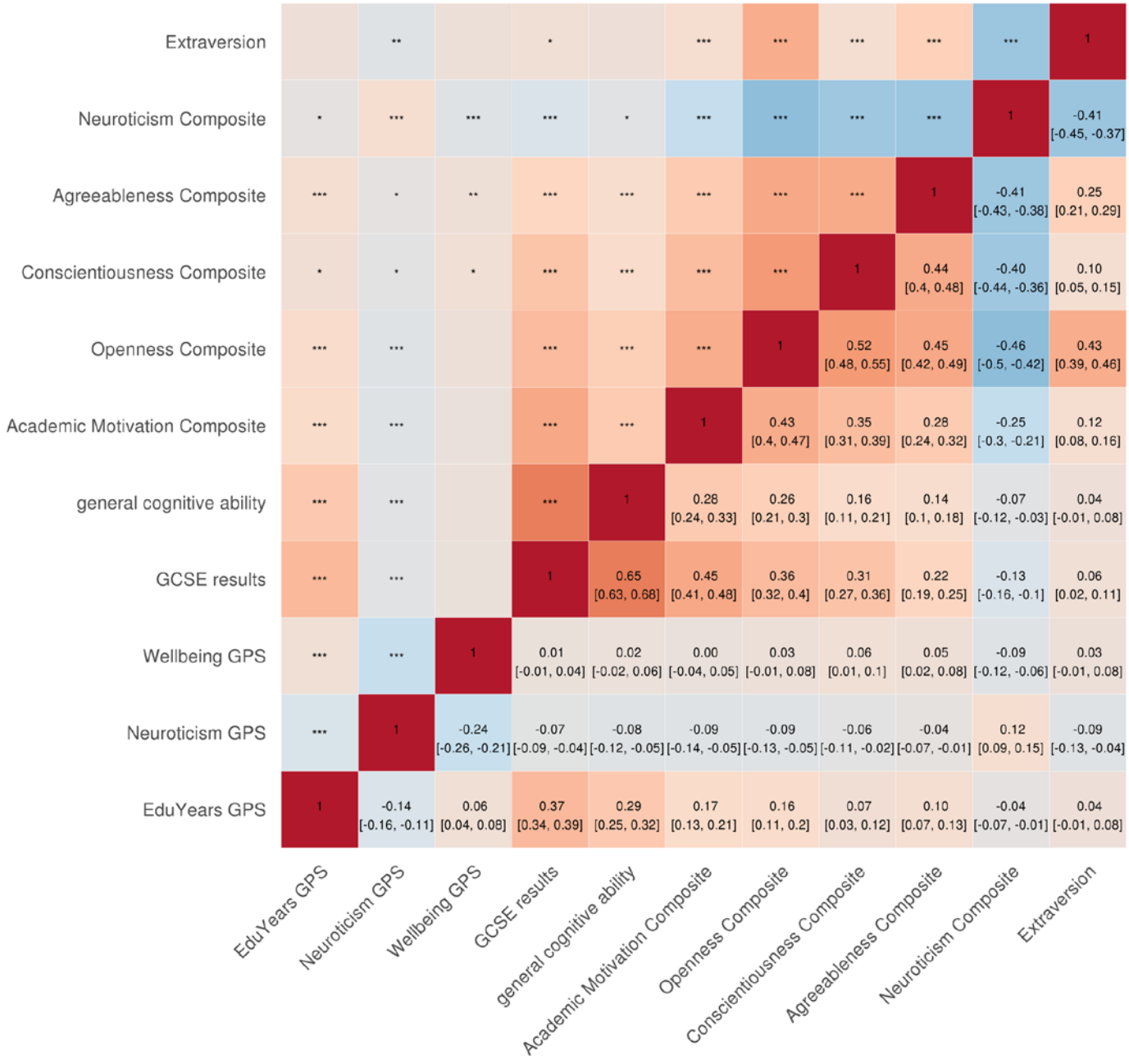
Note: (r.) = recoded so that higher scores were positive, i.e. less conduct problems. Variable labels in bold represent composites made up of the succeeding individual scales. * = p < 0.05; ** = p < 0.001; *** = p < 0.0001.

Figure S2 – Results from the parallel analysis.



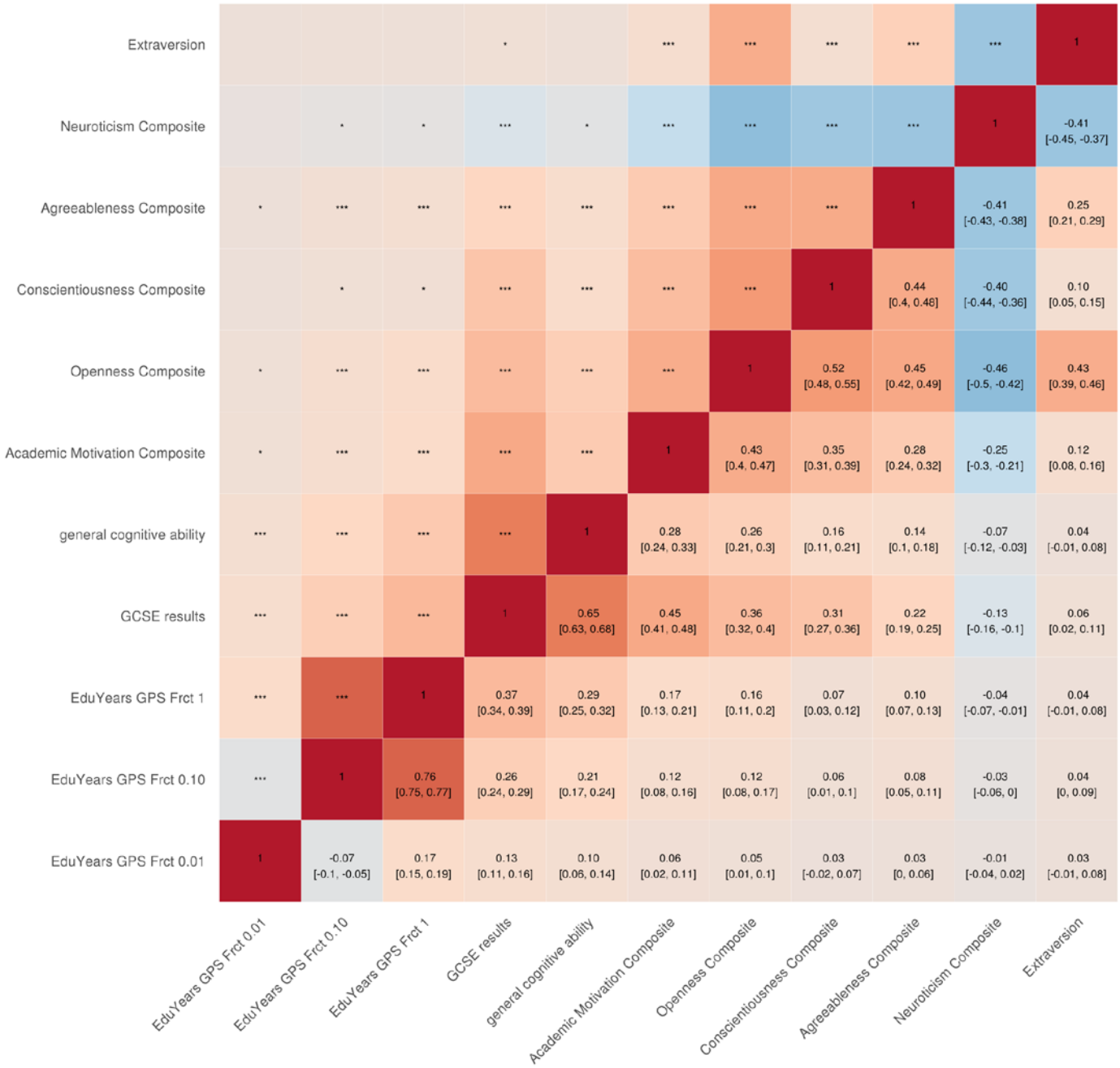
Note: The dotted line represents the point at which the factor eigenvalue in the study becomes smaller than the 95th percentile of the simulated eigenvalues from parallel analysis

Figure S3 – Correlations across all polygenic scores and personality/motivation domains



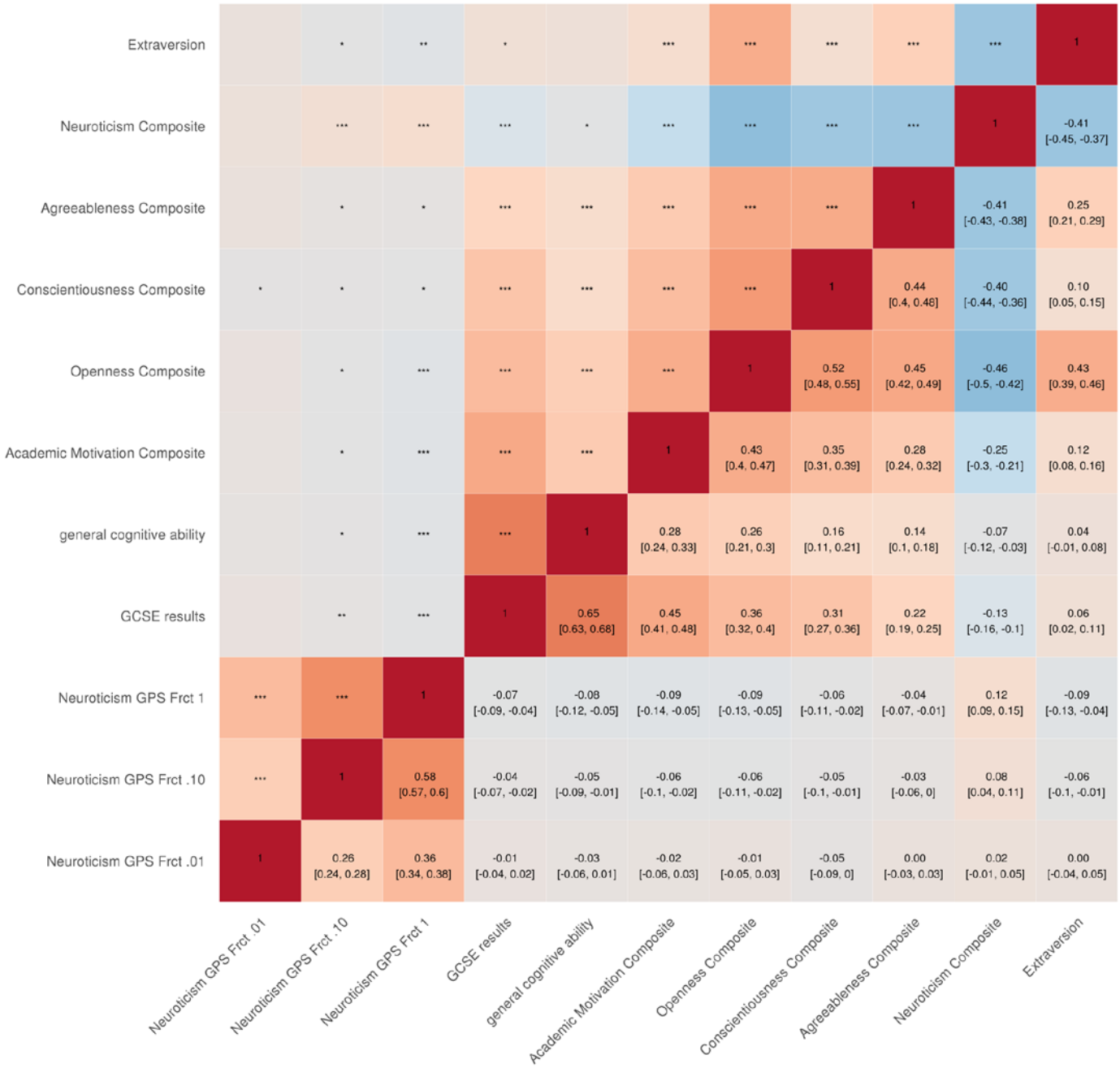
Note: 95% confidence intervals of the Pearson's correlation coefficients shown in square brackets. * = p < 0.05; ** = p < 0.001; *** = p < 0.0001.

Figure S4 – Correlations across all *EduYears* GPS thresholds and personality/motivation domains



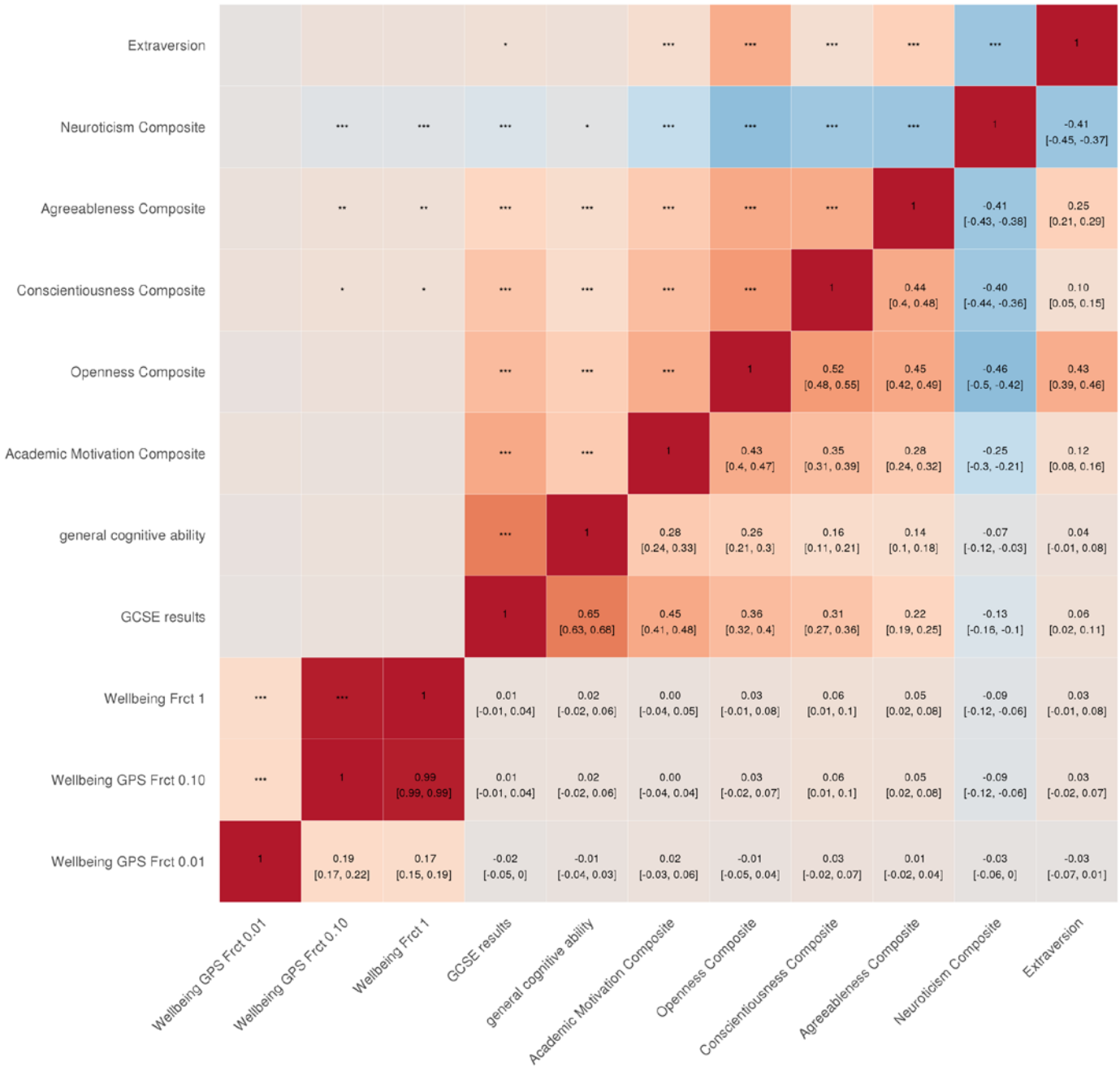
Note: 95% confidence intervals of the Pearson's correlation coefficients shown in square brackets. * = p < 0.05; ** = p < 0.001; *** = p < 0.0001.

Figure S5 – Correlations across all Neuroticism GPS thresholds and personality/motivation domains



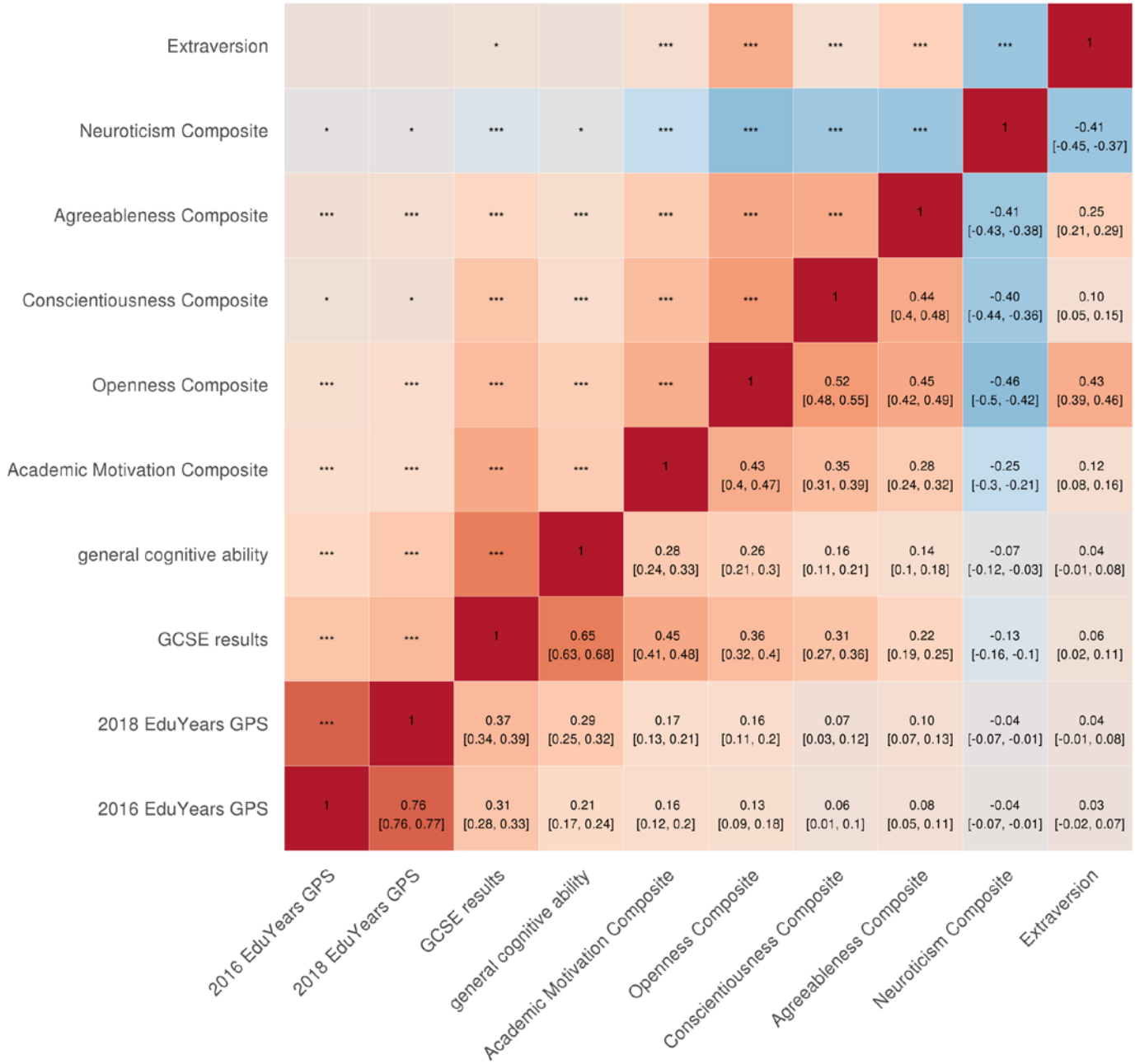
Note: 95% confidence intervals of the Pearson's correlation coefficients shown in square brackets. * = p < 0.05; ** = p < 0.001; *** = p < 0.0001.

Figure S6 – Correlations across all Wellbeing GPS thresholds and personality/motivation domains



Note: 95% confidence intervals of the Pearson's correlation coefficients shown in square brackets. * = p < 0.05; ** = p < 0.001; *** = p < 0.0001.

Figure S7 – Correlations between the 2016 *EduYears* GPS and outcome measures



Note: 95% confidence intervals of the Pearson's correlation coefficients shown in square brackets. * = p < 0.05; ** = p < 0.001; *** = p < 0.0001.