Supplementary Material

Selecting the number of classes

To select the number of classes for the two growth mixture models (GMMs), we initially modelled a single k-class solution, modelling subsequent k+1 solutions until the optimum solution was reached. Residual variances for repeated measures of emotional problems were fixed to equality over time. Each model was run with 4000 random starting values and 400 optimisations (STARTS = 4000 400 in Mplus) (1). As only three time-points were available for both the childhood and adolescent GMMs, change was modelled as linear. Fit statistics for the GMMs are presented in Table ST2. For childhood models, intercept variance was fixed to zero because for models with more than 2-classes the classes explained the variation in the intercept (i.e. the intercept variance was close to zero). Similarly, for adolescent models slope variance was fixed to zero because for models with more than 2-classes, the classes explained the variation in the slope (i.e. the slope variance was close to zero). For both the childhood and adolescence GMMs, model fit improved from 1- to 4-class solutions, as indicated by the sample size adjusted Bayesian information criterion (ssaBIC), Vuong-Lo-Mendell-Rubin Likelihood Ratio Rest (VLMR-LRT) and bootstrap likelihood ratio test (BLRT). However, both the childhood and adolescent 4-class solution included a small class (<5%), which was likely to lead to estimation problems / model instability in subsequent analyses. The 3-class solutions in both childhood and adolescence were therefore selected (entropy = 0.844 and 0.877 respectively).

To determine whether the classes were theoretically meaningful, we first interpreted the trajectories in light of the suggested cut-points for the emotional problems measure (scores ≥5 considered abnormal/high and those =4 borderline/slightly raised) (2) and then assessed differences in 'validators' of class membership.

We assessed early temperament and later depressive disorder in early adulthood as class 'validators'. Early temperament was measured at age 3 years using the parent-rated EAS Temperament Survey five-item emotionality subscale (3) (cries easily; tends to be somewhat emotional; often fusses and cries; gets upset easily; reacts intensely when upset) (individual item range 1-5). Depressive disorder was assessed at age 17.5 years using the Computerised Interview Schedule – Revised (CIS-R) (4) to generate depression diagnoses according to ICD10-F32 criteria.

Early temperament was associated with both the childhood and adolescent emotional problem classes (Wald $\chi^2(2)=483.58$, $p<0.001$ and Wald $\chi^2(2)=142.73$, $p<0.001$ respectively). For the childhood classes, higher early emotionality was associated with increased likelihood of being in both the decreasing (OR=1.26 (1.23-1.28)) and increasing classes (OR=1.16 (1.13-1.18)) compared to both the low class. It was also associated with the decreasing compared to the increasing class (OR=1.09 (1.05-1.12)). For the adolescent classes, higher early emotionality was again associated with both the decreasing (OR=1.13 (1.10-1.16)) and increasing classes (OR=1.11 (1.09-1.14)) compared to the low class. Early temperament did not differentiate the decreasing and increasing adolescent classes (OR=1.01 (0.98-1.05)). Both the childhood emotional problem classes and the adolescent emotional problem classes were associated with depressive disorder in early adulthood (Wald $\chi^2(2)=8.49$, $p=0.014$ and Wald $\chi^2(2)=52.22$, $p<0.0001$ respectively). Specifically, membership in the increasing childhood class (OR=1.89 (1.23-2.90), but not the decreasing class (OR=0.93 (0.53-1.63), was associated with an increased likelihood of MDD (major depressive disorder) compared to the low class. Membership in the increasing childhood class was not associated with an increased likelihood of MDD compared to the decreasing class (OR=2.04 (0.97-4.28)). In adolescence, the increasing class was associated with MDD compared to the low class (OR=3.87 (2.67-5.61)) but the decreasing class was not (OR=1.41 (0.78-2.55)). The increasing class was also associated with MDD compared to the decreasing class (OR=2.75 (1.37-5.52)).

Including covariates

The potential confounders of child sex, social class and maternal depression, home ownership, education and marital status were assessed. Social class was assessed as the average household income band, including social benefits, each week on a ten-point scale when the child was 134 months old. In-line with previous definitions, low income was categorised as the bottom quintile (5). Maternal severe depression, home ownership, education (O level or higher vs. below O level) and marital status (married vs. not married) were assessed my maternal report, assessed in pregnancy. Including these covariates revealed a similar pattern of results (see Supplementary Table ST3).

Excluding individuals exposed to prior peer-victimization

Sensitivity analyses were conducted excluding individuals who - according to an earlier parent-report measure - were exposed to peer victimization prior to the self-reported victimization assessment at age 8.5 years. Prior peer victimisation was assessed using a single item of the Strengths and Difficulties Questionnaire (SDQ) (2) peer problems subscale "picked on or bullied by other children" in childhood (ages 47, 81 and 97 months). Children were coded as having been exposed to prior victimization if parents rated this to have been "certainly true" at any of the three ages (N=129: 3.8%) and not having been exposed to prior victimization if parents rated this as "not true" or "somewhat true" at all three time-points (N=3296). Excluding individuals who were exposed to (parent-reported) prior victimization prior to the (self-reported) victimization assessment at age 8.5 years revealed a similar pattern of results, with the exception that the impact of victimization on transitioning from the low to the increasing trajectory was reduced (Supplementary Table ST3).

Inverse probability weighting

Inverse probability weighting (IPW) was used to address missing data. IPW has been recommended over alternative methods for dealing with missing data (such as multiple

imputation) in situations where whole blocks of data are missing for a large proportion of individuals (6). Weights were derived from a logistic regression analysis between a set of measures assessed in pregnancy (child sex, maternal severe depression and maternal financial difficulties (difficulties affording food, clothing, heating, accommodation and things for the baby)) that were independently predictive of variables in the analysis and/or inclusion in the main sample (N = 3988 / 9297). Minimal missing data on indicators used to derive weights were singly imputed as the modal or mean value (all indicators had <9% of values missing). The Hosmer-Lemeshow test was used assess the fit of the missingness model, with results showing no indication of poor fit (Hosmer-Lemeshow $\chi^2(6)=6.210$, $p=0.400$). Weights ranged from 2.13 to 3.83. The analyses were rerun using IPW to address any potential bias caused by participant dropout. Using IPW revealed a similar pattern of results (see Supplementary Table ST3).

References.

1. Muthén LK, Muthén BO. Mplus User's Guide. Seventh ed. Los Angeles, CA: Muthén & Muthén; 1998-2012.
2. Goodman R. The Strengths and Difficulties Questionnaire: A research note. Journal of Child Psychology and Psychiatry. 1997;38(5):581-6.
3. Buss AH, Plomin R. Temperament: Early Developing Personality Traits. Hillsdale, NJ: Lawrence Erlbaum; 1984.
4. Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. Psychol Med. 1992;22(02):465-86.
5. Langton EG, Collishaw S, Goodman R, Pickles A, Maughan B. An emerging income differential for adolescent emotional problems. Journal of Child Psychology and Psychiatry. 2011;52(10):1081-8.
6. Seaman SR, White IR, Copas AJ, Li L. Combining Multiple Imputation and Inverse-Probability Weighting. Biometrics. 2012;68(1):129-37.

ST1: Descriptive statistics for emotional problems

Age	N	Mean (SE)	Female mean (SE)	Male mean (SE)	Mean gender difference (SE)	
47 months	8083	1.44 (0.02)	1.46 (0.02)	1.41 (0.02)	0.06 (0.03)	p=0.095
81 months	7954	1.50 (0.02)	1.57 (0.03)	1.42 (0.03)	0.15 (0.04)	p<0.001
91 months	7191	1.65 (0.02)	1.71 (0.03)	1.60 (0.03)	0.11 (0.04)	p=0.008
140 months	6427	1.45 (0.02)	1.60 (0.03)	1.29 (0.03)	0.31 (0.04)	p<0.001
157 months	7018	1.44 (0.02)	1.66 (0.03)	1.58 (0.03)	0.45 (0.04)	p<0.001
198 months	5563	1.49 (0.02)	1.87 (0.04)	1.08 (0.03)	0.78 (0.05)	p<0.001

Possible N=8425 in childhood (ages 47, 81 and 91 months) and N=7018 in adolescnece (ages 140, 157 and 198 months).

ST2: Fit Information for the Growth Mixture Models

Number of classes	LL	Number of free parameters	ssaBIC	VLMR-LRT p-value	BLRT p-value	Entropy	Smallest class
Childhood*							
1	-43505.261	4	87033.967	-	-	-	N=8425 (100%)
2	-41532.872	7	83106.772	<0.0001	<0.0001	0.827	N=1496 (18%)
3	-41121.069	10	82300.750	0.0017	<0.0001	0.844	N=826 (10%)
4	-40668.793	13	81413.780	<0.0001	<0.0001	0.829	N=309 (4%)
Adolescence**							
1	-35181.780	4	70386.273	-	-	-	N=7018 (100%)
2	-34230.458	7	68500.666	<0.0001	<0.0001	0.888	N=887 (13%)
3	-33676.819	10	67410.423	<0.0001	<0.0001	0.877	N=446 (6%)
4	-33307.880	13	66689.580	0.0001	<0.0001	0.856	N=248 (4%)

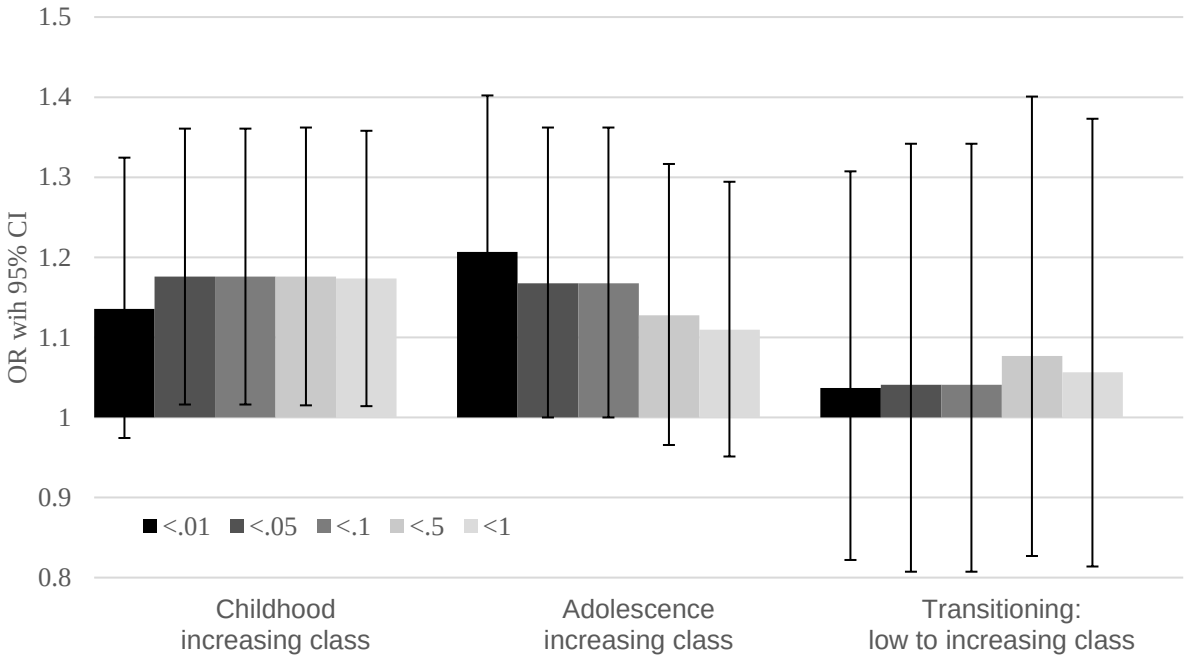
*Intercept variance was constrained to zero **Slope variance was constrained to zero. LL=Loglikelihood; ssaBIC= sample size adjusted Bayesian Information Criteria; VLMR-LRT=Vuong-Lo-Mendell-Rubin Likelihood Ratio Rest; BLRT= Bootstrapped Likelihood Ratio Rest. Smallest class based on the estimated model, rounded to nearest person/percent.

ST3: Sensitivity analyses

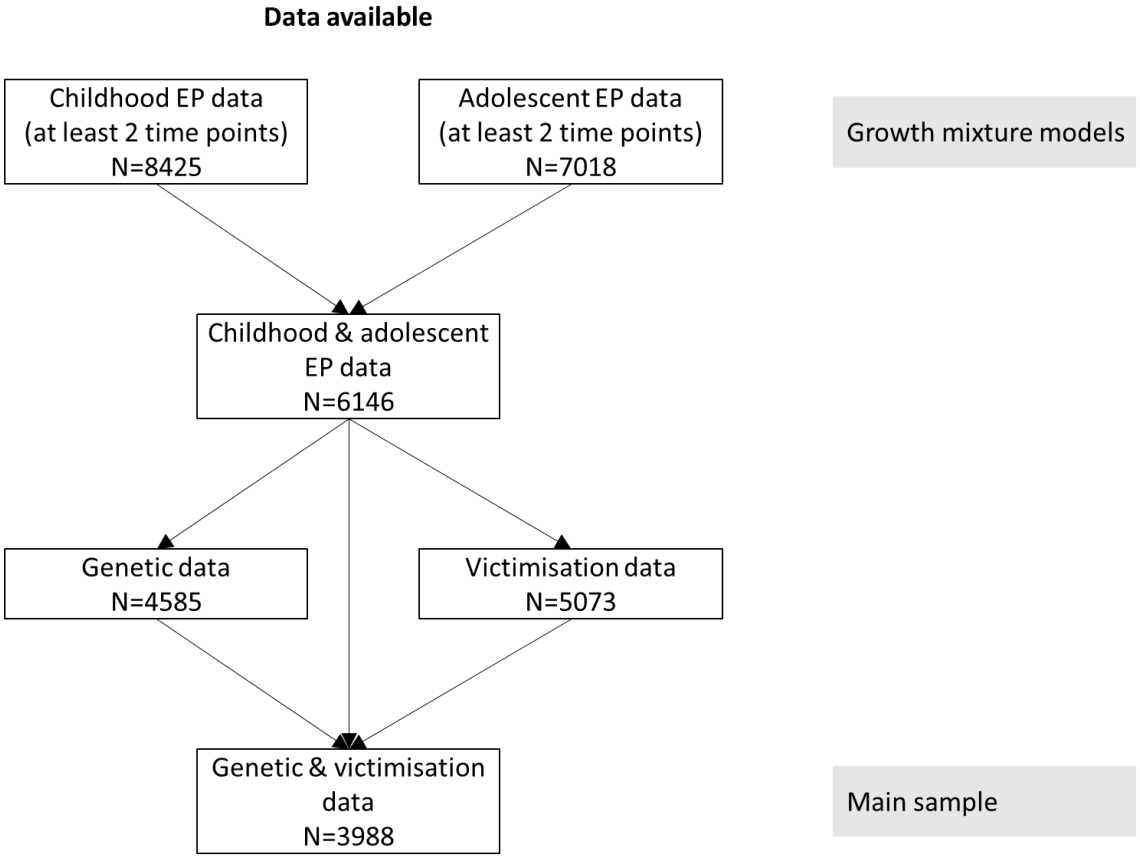
	Original results (N=3988)			Including covariates (N=3118)			Excluding prior victims (N=3296)			Using IPW (N=3988)		
	OR	(95% CI)	p	OR	(95% CI)	p	OR	(95% CI)	p	OR	(95% CI)	p
Schizophrenia PRS and increasing emotional problem class[*]												
Childhood	1.18	(1.02-1.36)	0.030	1.18	(1.00-1.40)	0.048	1.17	(0.97-1.40)	0.094	1.18	(1.02-1.36)	0.026
Adolescence	1.17	(1.00-1.36)	0.050	1.21	(1.01-1.46)	0.040	1.21	(1.01-1.44)	0.043	1.17	(1.00-1.37)	0.052
Transitioning from low childhood to increasing adolescent class[**]												
Schizophrenia PRS	1.04	(0.81-1.34)	0.758	1.13	(0.87-1.48)	0.349	1.12	(0.83-1.51)	0.466	1.06	(0.82-1.37)	0.668
Chronic victimisation	2.59	(1.48-4.53)	0.001	2.54	(1.26-5.11)	0.009	1.90	(0.90-4.00)	0.093	2.53	(1.44-4.45)	0.001

IPW= inverse probability weighting. [*]Low as the reference class, covariates predicting childhood and adolescent class [**]Covariates predicting victimisation and adolescent class

SF1. Associations for schizophrenia polygenic risk using a range of p-value thresholds from the discovery sample



SF2. Participation rates and missing data



EP=emotional problems

SF3: The probabilities of transitioning from each of the childhood trajectory classes to each of the adolescent classes