Prospective associations between internalising symptoms in youth and educational achievement: A monozygotic twin differences study.

Introduction

How well a child does at school has implications for the individual, their family, and society more widely (Mirowsky & Ross, 2017). Educational achievement is associated with a range of life outcomes, including well-being, physical and mental health, occupational status and even life expectancy (Cutler & Lleras-Muney, 2012). It is therefore important to understand the causes and correlates of individual differences in academic achievement.

One candidate influence is internalising symptoms (encompassing anxiety and depressive symptoms): children experiencing higher levels of anxiety and depression may go on to do less well in school. For example, children who worry may find it difficult to concentrate; young people experiencing depressed mood may struggle to motivate themselves to complete homework tasks; and youth with social anxiety may be inhibited in the classroom and avoid asking questions and participating. It is also plausible that some individuals may find the school setting so aversive that they are deterred from continuing onto higher education or fail to get the necessary grades required to transition. A competing hypothesis is that detrimental effects are tempered by the heightened conscientious behaviours that are associated with some forms of anxiety and related symptoms (Scher & Osterman, 2002). For example, a child who worries about failing a test may revise more assiduously than their less anxious peer.

The association between internalising disorders and internalising symptoms with academic underachievement has been examined in numerous studies. In clinical samples, anxiety and depressive disorders have typically been found to be associated with self-reported underachievement, more days of school missed, lower objective school grades, and lower likelihood of transitioning to tertiary education (see de Lijster et al. (2018) and Wickersham, Sugg, et al. (2021) for reviews). However, most studies that have examined this relationship have not included measurement of potential confounding factors. For example, certain shared environmental factors (i.e., aspects of the environment that are common to members of a twin pair and give rise to phenotypic similarity), which could include socioeconomic status, are important determinants of both educational achievement (Sirin, 2005) and internalising problems (Bromberger et al., 2017), and so may account for some of the association between the two phenotypes. Similarly, child-specific traits such as baseline

cognitive ability, predict both achievement (Hegelund, Flensborg-Madsen, Dammeyer, & Mortensen, 2018) and internalising problems (Flouri et al., 2019) and therefore may also explain the relationship between the two. Furthermore, it has been established that both internalising symptoms (Trzaskowski, Zavos, Haworth, Plomin, & Eley, 2012) and educational achievement (Rimfeld et al., 2018) are heritable, with overlapping genetic influences on these phenotypes (Krapohl et al., 2014). Genetic factors may thus partially explain the association between internalising symptoms and educational outcomes. In order to understand a potential causal relationship between internalising symptoms and educational outcome, there is a need to control for potential phenotypic and genetic confounding and adopt longitudinal designs.

A handful of studies have overcome these limitations to examine the prospective association between internalising disorders and later academic achievement in treatment-seeking samples, over and above potential confounds. For example, a population-based cohort study of over two million individuals using data from the Swedish national health registery examined the association between a clinical diagnosis of social anxiety disorder and a number of markers of educational achievement, such as external examination grades at 16 years and transition to higher education (Vilaplana-Pérez et al., 2020). Individuals with a diagnosis of SAD achieved poorer outcomes across all measures of educational achievement compared to individuals without a diagnosis of SAD. The findings remained unchanged when controlling for unmeasured environmental and genetic factors shared between siblings and when controlling for psychiatric morbidities. This study indicates that clinical levels of internalising problems are prospectively related to objective markers of lower educational achievement, and the association is not an artefact of socioeconomic status, individual factors such as cognitive ability, or shared genetic liability.

However, the focus on treatment-seeking individuals limits the generalisability of these findings, and because these are dimensional disorders there is value in examining the association between internalising symptoms and educational achievement more broadly. Studies with community samples have provided evidence of an association between internalising symptoms and objective educational achievement across the severity spectrum. For example, a meta-analysis of 26 studies found a small but significant effect of variation in anxiety, depression, and internalising symptoms (defined in the review as a combination of anxiety and depression symptoms) on later school failure (defined in the review as failure to complete compulsory education). Depression and internalising symptoms but not anxiety symptoms were associated with later school grades (Riglin, Petrides, Frederickson, & Rice,

2014). Unfortunately, the review was unable to test the effect of confounds, such as socioeconomic status or baseline cognitive ability, leaving open the question of whether internalising symptoms are independently related to academic problems in community samples. However, certain community studies have examined the role of specific confounds. For example, in a study with a New Zealand cohort, the association between anxiety disorders at age 14-16 years and transition to tertiary education was found to persist over and above socioeconomic and individual factors, including substance abuse, neuroticism, and cognitive ability at age 9 (Woodward & Fergusson, 2001). Two longitudinal community studies focusing on depression and later attainment (Wickersham, Dickson, et al., 2021; Fergusson & Woodward, 2002) reported divergent findings. A New Zealand-based cohort study of Fergusson and Woodward (2002) found an association between depression diagnosis in mid-adolescence and educational attainment in late-adolescence, however this was explained by the presence of individual and socioeconomic factors. In contrast, a UK-based study of Wickersham and colleagues (2021) found that young people with a depression diagnosis showed a decline in attainment later on, over and above a measure of socioeconomic status. Taken together, there is evidence to suggest that across the severity spectrum, earlier internalising symptoms are associated with later objective academic attainment, and there is some evidence to indicate that this persists over and above phenotypic confounds. As yet, no studies have tested the possibility that across the severity spectrum, the relationship remains after adjusting for genetic as well as phenotypic confounds. If found, this would support the notion that educational underachievement is, in part, a functional consequence of experiencing internalising difficulties. Additionally, little is known about the differential associations of different internalising subtypes and academic achievement. The present study sought to bridge this gap.

The first aim was to examine prospective associations between internalising symptoms and educational achievement after controlling for socioeconomic status and cognitive ability assessed in childhood, using data from a large British twin study. We were also interested in testing differential prospective associations between subtypes of internalising symptoms and educational achievement. In other words, are particular subtypes of internalising symptoms, namely, social anxiety, fear, negative affect, and negative cognition, uniquely associated with academic achievement, over and above socioeconomic status and baseline cognitive ability? The second aim of the study was to use the monozygotic twin design to assess whether the observed associations remain significant after controlling for genetic and shared environmental confounding.

Method

Sample

Participants were drawn from the Twins Early Development Study (TEDS), a longitudinal study of twins born in the United Kingdom between 1994 and 1996. The families in TEDS (Rimfeld et al., 2019) are representative of the British population in their socio-economic distribution, ethnicity, and parental occupation for their birth cohort (Rimfeld et al., 2019). TEDS research was approved by the Institute of Psychiatry, Psychology & Neuroscience Ethics Committee, and all participants gave informed consent. All individuals with major medical, genetic, or neurodevelopmental disorders were excluded from the dataset.

The present study focuses on data collected in a subsample of TEDS twins over four waves: age 7, age 9, age 16, and age 18 (approximate years). Participants were included in the current study if they had anxiety-related symptoms available at two of the three assessment points (7/9/16 years) to compute all subscales. There was a total of 10,791 twins, including 1,981 monozygotic (MZ) twin pairs.

Measures

Internalising symptoms were extracted from the parent-rated Anxiety-Related Behaviours Questionnaire (Eley et al. (2003)). This measure includes items reflecting symptoms of both anxiety and depression. It is not a diagnostic instrument but shares similarities with other widely used screening measures of child anxiety disorders (e.g., the Spence Children's Anxiety Scale; Spence (1998)). The items included are rated on a three-point scale (0 = never, 1 = sometimes, 2 = often).

Previous factor analysis revealed five scales at ages 7 and 9: fear, social anxiety, negative cognition, negative affect, and obsessive-compulsive behaviour (OCB) (Hallett, Ronald, Rijsdijk, & Eley, 2009). In the present study, in line with (Trzaskowski et al., 2012), the OCB scale was excluded from the analysis. This is due to low consistency of OCB item loadings at different assessment points. Specific items for each subscale are shown in Table S1 (Supplementary Material).

In order to generate the composites for the present study, the four subscales were first computed at each time point. Where items were missing, these were calculated using the participant's mean subscale score, with a maximum of 20% missing items allowed per participant per time point. Due to positive skew (see Table S2), subscales at each time point were submitted to a log+1 transformation for analysis. Internal consistency was generally acceptable for the subscales at each time point (see Table S2). Within-subject phenotypic correlations over time were calculated to determine the stability of each subscale. The correlations indicated a moderate degree of stability over time (see Table S2). The subscale composites were computed by creating a mean of the subscales at age 7, 9, and 16, with a maximum of one missing data point allowed. This was done to allow examination of the stable component of these internalising subtypes across childhood and adolescence. Internal consistency of the composites was moderate to good (.74-.84; see Table S2).

The negative affect and negative cognition composite scores ranged from a minimum possible score of 0 to maximum possible of 12; the social anxiety composite score ranged from 0-8; and the fear composite score ranged from 0-10.

General cognitive ability (*g*) was measured at age 7 and used as a baseline measure in the present study. It was calculated as a mean of conceptual grouping (McCarthy, 1972), a Wechsler Intelligence Scale for Children (WISC) similarities test, a WISC vocabulary test, and a WISC picture completion test (Wechsler, 1992), all collected via telephone testing.

Socioeconomic status was assessed as a composite from five variables: mother and father employment levels, mother and father educational levels, and mother's age on birth of first child. This data was collected at first contact.

Educational achievement was measured in three ways: General Certificate of Secondary Education (GCSE) grades; A-Level grades; and transition to university.

GCSE exam results were obtained from twins themselves or from their parents via mailed questionnaires or via telephone. It has been shown previously that self-reported exam results are accurate (Rimfeld, Dale, & Plomin, 2015). GCSEs are UK-wide standardized examinations taken at age of 16 at the end of compulsory education. Children choose from a variety of different subjects, while English, Mathematics, and Science are compulsory. We

used exam grades from English, Mathematics, and Science for the current analyses. The grades were coded from 11 (the highest grade, A*) to 4 (the lowest pass grade, G) (no information about failed results was available). Composite measures were created because the scores on the core subjects are highly correlated. The composite measure was constructed as the mean of English (mean of English language and English literature grades), Science (mean of single or double-weighted Science or, when taken separately Chemistry, Physics, and Biology grade), and Mathematics, with a range of 4-11.

A-level examination grades were obtained from twins themselves or from their parents via questionnaires sent over mail or via telephone when the twins were age 18. A-level examination grades (ranging from A* to E) were coded from 6 (A*) to 1(E). Participants who did not study post-16 years were given a rating of 0. A mean grade was calculated, computed as the average grade achieved across all subjects, with a range of 0-6.

Information on whether or not twins chose to attend university was collected via questionnaire at age 18. Choosing to attend university was treated as a dichotomous variable, where 1 indicated the choice to pursue university and 0 indicated any other post compulsory education destination, such as going into employment, training, or unemployment.

Statistical Analysis

Analyses were pre-specified and registered on Open Science Framework (https://osf.io/j2p7b/). Analyses were conducted in *R* and STATA 14.2. Analyses were undertaken with all the available data. Therefore, number of observations varied across analyses.

Phenotypic analyses

A series of univariate regressions were undertaken in the whole sample (i.e., MZ and DZ twins) to determine whether each internalising subtype was associated with academic outcomes. Linear regressions were used for continuous outcome variables (GCSE and A-Levels) and logistic regressions were carried out for the binary outcome (transition to university education). Analyses were then repeated using multivariable regression models to examine the unique influence of all internalising subtypes on the three educational outcomes in turn. Analyses controlled for natal sex, socioeconomic status, and general cognitive ability at age 7, and accounted for the non-independence of twin pairs. Variance inflation factors

(VIF) were computed to assess multicollinearity in the multivariable regression models. All VIFs were under 2, indicating that multicollinearity was unlikely to influence our results (see Table S3).

MZ twin difference analyses

Phenotypic analyses were repeated using MZ twin difference scores, which control for genetic and shared environmental confounding. Since MZ twins are genetically identical and share their rearing environment, phenotypic differences between members of an MZ twin pair must result from non-shared environmental or child-specific influences. Relative MZ twin difference scores were calculated by randomly assigning twin members as Twin 1 and Twin 2 and subtracting the score of Twin 2 from that of Twin 1. This produced continuous scores for GCSE and A-level variables, and an ordinal score for transition to university; transition to University was a binary outcome (coded as 0 or 1) and therefore calculating the MZ twin difference score produce an score of -1, 0, or 1. Univariate and multivariable linear regression models were undertaken, controlling for general cognitive ability, not sex or socioeconomic status since these variables are identical between members of MZ twin pair. Linear regression models were used in MZ twin difference analyses of GSCE and A-levels (continuous variables) and ordered regressions for transition to university education (ordinal variable).

Results

Missing Data Analysis

After excluding those individuals with major medical, genetic, or neurodevelopmental disorders (n=1,097), the total potential sample size was 19,379 participants. Of these, 10,791 (56%) had data available at two of the three assessment points to compute all four internalising subscales and so were included in the study. Participants who had subscale data available and so could be included in the study were more likely to be female and from a higher socioeconomic status than those who did not (see Table S4). However, the magnitude of the differences was small and not clinically meaningful. Of the 10,791 participants for whom anxiety-related behaviours data was available at two of the three assessment points to compute the four subscales of interest, data was available on GCSE grades for 82.24% of the study dataset, on A-level grades for 58.74%, and on transition to higher education for 79.23%. For all three educational outcomes, missing data was statistically more likely amongst males, those from a lower socioeconomic status, and those scoring higher on the

internalising subscale scores (see Table S5 for details). However, the magnitude of the differences was small (see Table S5) and not clinically meaningful. For example, the average negative affect subscale score was 0.23 (SD=0.19) for those with GCSE data, compared to an average subscale score of 0.31 (SD=0.23) for those whose GCSE data was missing.

Sample Characteristics

Table 1 shows descriptive statistics for the whole sample and gender differences in internalising subscale scores and academic outcomes. Females had significantly higher scores on all ARBQ composites, attained higher GCSE and A-level grades, and were more likely to transition to university education.

MZ twin difference scores for the internalising subscale scores, GCSE and A-level grades are shown in Table 1 (the frequency distributions of the MZ twin difference scores are shown in Supplementary Figure 1). The mean of the MZ difference scores generally approximates zero since the random assignment of twins ensures that cases where 'twin 1' scores higher than 'twin 2' are cancelled out by cases in which 'twin 2' scores higher than 'twin 1'. What is important is the distribution of these scores. Clustering around zero demonstrates similarity between MZ twins whereas deviations from zero demonstrate the presence of differential internalising subscale scores due to the nonshared environment. Although many of the MZ twin difference scores approximate to zero, indicating that they show similar levels of internalising symptoms, a substantial number do not rate similar levels of internalising symptoms as indicated by the standard deviations (see Table 1). For example, the mean difference score for the negative cognition composite was zero, but the SD was 1.56, meaning that approximately 32% of MZ twins differ by at least 1.56 points.

Phenotypic Associations between Internalising Symptom Subtypes and Educational Outcomes

Results of the phenotypic analyses are shown in Table 2. In the univariate regression models, all internalising symptom subtypes were significantly and negatively associated with GSCE outcomes (β coefficients ranging from -.04 [95% CI: -.06, -.01] for social anxiety to -.12 [95% CI: -.15, -.09] for negative affect). However, only negative affect and negative cognition were associated with poorer A-level outcomes (β = -.09 [95% CI: -.12, -.05] and -.05 [95% CI: -.08, -.02], respectively) and reduced likelihood of transitioning to higher education (OR = .81 [95% CI: .75, .87] and .90 [95% CI: .84, .96], respectively).

Multivariable models indicated that negative cognition and negative affect were associated with unique variance in GSCE outcomes (β = -.10 [95% CI: -.13, -.07] and -.04 [95% CI: -.07, -.01], respectively), but only negative affect significantly predicted A-level attainments (β = -.09 [95% CI: -.13, -.05]) and transition to university education (OR = .81 [95% CI: .74, .87]).

Associations between Internalising Symptom Subtypes and Educational Outcomes, controlling for familial effects using MZ twin differences

When controlling for genetic and shared environmental influences using MZ twin difference scores, univariate regression models showed only negative cognition to be associated with GCSE outcomes (β = -.04 [95% CI: -.06, -.01]). None of the internalising symptom subtypes were significantly associated with A-levels attainments or transition to university education (see Table 2). In multivariable analyses of MZ twin difference scores, the association between negative cognition and GCSE attainments remained significant (β = -.03 [95% CI: -.06, -.01]).

Discussion

The current study represents the first investigation of the associations between internalising symptom subtypes in youth and later academic attainments, using a prospective and genetically-informative design. With respect to our first aim, results indicated a robust association of certain internalising symptom subtypes with academic outcomes. Specifically, symptoms of negative cognition and negative affect experienced between mid-childhood and mid-adolescence were negatively associated with attainments in national examinations at age 16. These associations were significant after accounting for the potential confounding effects of intellectual ability and socioeconomic status, and existed independent of the influence of other internalising symptom subtypes. Moreover, our findings indicated that earlier negative affect continued to be associated with academic attainments in late adolescence and transition to higher education in early adulthood. Although we cannot infer a causal relationship based on the current findings, it is nevertheless of potential clinical significance that the relationship of negative affect with academic attainment appears to be relatively persistent, with symptoms experienced in youth continuing to be associated with academic underachievement several years later.

In relation to our second aim, the association between negative affect and academic achievement became non-significant when using MZ twin difference scores to control for

genetic and shared environmental effects. Therefore, these results could suggest that much of the phenotypic association between negative affect and academic underachievement is accounted for by genetic and/or shared environmental effects. This finding is consistent with previous studies showing that common genes account for the majority of the covariance in anxiety and educational outcomes (Tambs et al., 2012). Similarly, certain environmental experiences that are common to members of a twin pair, such as socioeconomic status, have been linked with internalising symptoms (Morrissey & Kinderman, 2020; Ridley, Rao, Schilbach, & Patel, 2020; Gilman, Kawachi, Fitzmaurice, & Buka, 2002) and educational attainments (von Stumm et al., 2020), and may therefore contribute to the relationship between these phenotypes. In the current study, we attempted to control for socioeconomic status using a composite index that included parental educational level, parental occupational status, and mothers' age at birth of their first child. However, socioeconomic status is a broad construct and thus our index is unlikely to have captured all relevant aspects.

Interestingly, in contrast to our findings for negative *affect*, we found that the relationship between negative *cognition* and GCSE grades remained significant in the MZ twin difference analyses. This suggests that a substantial proportion of the link between negative cognition and academic attainment may be accounted for by non-shared environmental factors, consistent with the notion that negative cognition may causally influence academic performance. In this vein, aspects of negative cognition (e.g., excessive worry, preoccupation with negative social evaluation) have been shown to have a negative impact on core aspects of neuropsychological functioning (e.g., concentration and working memory) and it is therefore plausible that they could lead to academic underachievement. Given that our finding of a non-shared environmental association between negative cognition and GCSE grades did not extend to A-level or transition to university education, it should be interpreted cautiously. It is of note that in the UK there are several other educational qualifications that are broadly comparable to A-levels (e.g. National Vocational Qualifications). It is plausible that some participants in the current study did not complete A-level but undertook one of these alternative qualifications instead, which could have obscured the associations of interest.

Our study raises the question of why, in our phenotypic analyses, negative affect and cognition were uniquely associated with certain educational attainments, whereas social anxiety and fear were not. This is broadly in keeping with a previous meta-analysis showing that educational outcomes were more consistently associated with depression than anxiety (Riglin et al., 2014). However, the current findings may also reflect how our constructs were

defined and measured. Negative affect and cognition are broad constructs that include features relevant to a range of anxiety and mood disorders. Negative affect assesses symptoms including anhedonia, low mood, distress, and tension, which could be indicative of depression or certain anxiety disorders (e.g., tension is a diagnostic symptom of generalised anxiety disorder). Negative cognition captures self-criticism, worry, and self-blame, which are common features of depression but also anxiety disorders (e.g., self-criticism is commonly elevated in young people with social anxiety disorder). As such, the negative affect and cognition subscales may capture transdiagnostic processes, whereas the fear and social anxiety were more circumscribed and may have captured less variance.

The findings of this study have several implications for clinical and educational practice. First, experiencing negative affect and cognition (particularly negative affect) between mid-childhood and mid-adolescence appears to be a marker for subsequent academic underachievement. This emphasises the importance of increasing awareness and detection of these symptom profiles. Second, with respect to negative *affect*, our findings suggest that much of its association with academic attainment is confounded by familial (genetic and shared environmental) effects. If replicated, this would suggest that while symptoms of negative affect may warrant intervention in order to improve mental health, negative affect may not by causally related to academic underachievement and therefore treating symptoms will not necessarily lead to improved academic outcomes. Instead, it may be that such children require additional educational support to optimise their academic attainments. Third, our findings provide preliminary evidence for a non-shared environmental association between negative *cognition* and educational outcomes, although we note that the effect sizes are modest. Further research is needed to establish a causal link, but nevertheless this finding raises the possibility that treatments targeting aspects of negative cognition (e.g., worry, self-criticism) will not only benefit mental health, but potentially also educational outcomes.

The current study has several strengths including its longitudinal, genetically informative design, and inclusion of multiple, objective educational milestones. However, results should be considered within the context of several limitations. First, a proportion of participants did not have educational outcome data available. However, while there were some statistically significant differences in those with and without outcome data, the magnitude of the differences was small and unlikely to be clinically meaningful. Furthermore, although selective attrition may have resulted in less variance at the outcome timepoints, previous research has shown that selective attrition effects estimates of prevalence but has less influence on estimates of association, as was the focus in the current

study (Wolke et al., 2009). Second, internalising symptoms were assessed using a parent-rated questionnaire. Research has shown significant informant discrepancies in the reporting of internalising symptoms in youth (Youngstrom et al., 2011), highlighting the importance of including multiple informants in future studies. Third, the social anxiety subscale of the ARBQ comprised four items and maybe therefore have captured a narrow range of symptoms. Of particular note, the items focussed exclusively only on social anxiety related *behaviour* (e.g. being shy or timid) and did not assess related *cognitions* (e.g., negative evaluative concerns), which may have a greater link with educational performance. Fourth, it should be acknowledged that the relationship between internalising symptoms and GCSEs results *could* be capturing concurrent as opposed to prospective associations, since our composite scores of internalising subtypes encompassed symptoms from aged 7 to 16 years, and GSCEs are typically completed at age 16. Fifth, the ARBQ is not a diagnostic instrument and does not have clinical cut-offs. Moreover, as mentioned above, the negative affect and cognition subscales are transdiagnostic constructs and do not relate specifically to one diagnostic category. In the current study, we were interested in examining internalising symptoms across the continuum in a population-based sample. However, future studies, using prospective, genetically-informative designs, are needed to examine the association of diagnosed anxiety and mood disorders with educational outcomes. Sixth, the present study is unable to examine whether the observed associations are specific to internalizing psychopathology or may also be at least partly accounted for by externalizing psychopathology. Future studies testing whether the association holds after controlling for externalising problems will be important, given the known co-occurrence of internalizing and externalising symptoms (Fanti & Henrich, 2010) and association between externalising symptoms and lower educational achievement (Lewis, Asbury, & Plomin, 2017).

In summary, the current study suggests that experiencing certain internalising symptoms across childhood and early adolescence is associated with reduced academic attainment. In particular, negative affect is associated with persistent academic underachievement through to late adolescence and early adulthood, which in turn may have lifelong consequences both for the individual but also society at large. Although further research is needed to understand the aetiology of the relationship, our findings indicate that academic underachievement is not simply a consequence of the disruption caused by symptoms and therefore additional interventions may be required to optimise outcomes.

References

Wickersham, A., Dickson, H., Jones, R., Pritchard, M., Stewart, R., Ford, T., & Downs, J. (2021). Educational attainment trajectories among children and adolescents with depression, and the role of sociodemographic characteristics: longitudinal data-linkage study. *British Journal of Psychiatry, 218*(3), 151-157. doi:10.1192/bjp.2020.160

Wickersham, A., Sugg, H. V. R., Epstein, S., Stewart, R., Ford, T., & Downs, J. (2021). Systematic review and meta-analysis: The association between child and adolescent depression and later educational attainment. *Journal of the American Academy of Child & Adolescent Psychiatry, 60*(1), 105-118. doi:10.1016/j.jaac.2020.10.008

Morrissey, K., & Kinderman, P. (2020). The impact of childhood socioeconomic status on depression and anxiety in adult life: Testing the accumulation, critical period and social mobility hypotheses. *SSM - Population Health, 11*, 100576. doi:10.1016/j.ssmph.2020.100576

Vilaplana-Pérez, A., Pérez-Vigil, A., Sidorchuk, A., Brander, G., Isomura, K., Hesselmark, E., . . . Fernández de la Cruz, L. (2020). Much more than just shyness: the impact of social anxiety disorder on educational performance across the lifespan. *Psychological Medicine*, 1-9. doi:10.1017/S0033291719003908

von Stumm, S., Smith-Woolley, E., Ayorech, Z., McMillan, A., Rimfeld, K., Dale, P. S., & Plomin, R. (2020). Predicting educational achievement from genomic measures and socioeconomic status. *Developmental Science, 23*(3), e12925. doi:10.1111/desc.12925

Ridley, M., Rao, G., Schilbach, F., & Patel, V. (2020). Poverty, depression, and anxiety: Causal evidence and mechanisms. *Science, 370*(6522). doi:10.1126/science.aay0214

Rimfeld, K., Malanchini, M., Spargo, T., Spickernell, G., Selzam, S., McMillan, A., . . . Plomin, R. (2019). Twins Early Development Study: A genetically sensitive investigation into behavioral and cognitive development from infancy to emerging adulthood. *Twin Research and Human Genetics, 22*(6), 508-513. doi:10.1017/thg.2019.56

Flouri, E., Papachristou, E., Midouhas, E., Ploubidis, G. B., Lewis, G., & Joshi, H. (2019). Developmental cascades of internalising symptoms, externalising problems and cognitive ability from early childhood to middle adolescence. *European Psychiatry, 57*, 61-69. doi:10.1016/j.eurpsy.2018.12.005

Hegelund, E. R., Flensborg-Madsen, T., Dammeyer, J., & Mortensen, E. L. (2018). Low IQ as a predictor of unsuccessful educational and occupational achievement: A register-based study of 1,098,742 men in Denmark 1968–2016. *Intelligence, 71*, 46-53. doi:10.1016/j.intell.2018.10.002

de Lijster, J. M., Dieleman, G. C., Utens, E. M. W. J., Dierckx, B., Wierenga, M., Verhulst, F. C., & Legerstee, J. S. (2018). Social and academic functioning in adolescents with anxiety disorders: A systematic review. *Journal of Affective Disorders, 230*, 108-117. doi:10.1016/j.jad.2018.01.008

Rimfeld, K., Malanchini, M., Krapohl, E., Hannigan, L. J., Dale, P. S., & Plomin, R. (2018). The stability of educational achievement across school years is largely explained by genetic factors. *npj Science of Learning, 3*(1), 16. doi:10.1038/s41539-018-0030-0

Mirowsky, J., & Ross, C. E. (2017). *Education, social status, and health* (1st ed.). New York: Routledge.

Lewis, G. J., Asbury, K., & Plomin, R. (2017). Externalizing problems in childhood and adolescence predict subsequent educational achievement but for different genetic and environmental reasons. *Journal of Child Psychology and Psychiatry, 58*(3), 292-304. doi:10.1111/jcpp.12655

Bromberger, J. T., Schott, L. L., Matthews, K. A., Kravitz, H. M., Harlow, S. D., & Montez, J. K. (2017). Childhood socioeconomic circumstances and depressive symptom burden across 15 years of follow-up during midlife: Study of Women's Health Across the Nation (SWAN). *Archives of Women's Mental Health, 20*(4), 495-504. doi:10.1007/s00737-017-0747-4

Rimfeld, K., Dale, P. S., & Plomin, R. (2015). How specific is second language-learning ability? A twin study exploring the contributions of first language achievement and intelligence to second language achievement. *Translational Psychiatry, 5*(9), e638-e638. doi:10.1038/tp.2015.128

Riglin, L., Petrides, K. V., Frederickson, N., & Rice, F. (2014). The relationship between emotional problems and subsequent school attainment: A meta-analysis. *Journal of Adolescence, 37*(4), 335-346. doi:10.1016/j.adolescence.2014.02.010

Krapohl, E., Rimfeld, K., Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Pingault, J.-B., . . . Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences, 111*(42), 15273-15278. doi:10.1073/pnas.1408777111

Trzaskowski, M., Zavos, H. M. S., Haworth, C. M. A., Plomin, R., & Eley, T. C. (2012). Stable genetic influence on anxiety-related behaviours across middle childhood. *Journal of Abnormal Child Psychology, 40*(1), 85-94. doi:10.1007/s10802-011-9545-z

Tambs, K., Kendler, K. S., Reichborn-Kjennerud, T., Aggen, S. H., Harris, J. R., Neale, M. C., . . . Røysamb, E. (2012). Genetic and environmental contributions to the relationship between education and anxiety disorders - a twin study. *Acta Psychiatrica Scandinavica, 125*(3), 203-212. doi:10.1111/j.1600-0447.2011.01799.x

Cutler, D. M., & Lleras-Muney, A. (2012). Education and health: insights from international comparisons. NBER Working Paper No. 17738. In. Cambridge: NBER.

Youngstrom, E. A., Youngstrom, J. K., Freeman, A. J., De Los Reyes, A., Feeny, N. C., & Findling, R. L. (2011). Informants are not all equal: predictors and correlates of clinician judgments about caregiver and youth credibility. *Journal of child and adolescent psychopharmacology, 21*(5), 407-415. doi:10.1089/cap.2011.0032

Fanti, K. A., & Henrich, C. C. (2010). Trajectories of pure and co-occurring internalizing and externalizing problems from age 2 to age 12: findings from the National Institute of Child Health and Human Development Study of Early Child Care. *Dev Psychol, 46*(5), 1159-1175. doi:10.1037/a0020659

Wolke, D., Waylen, A., Samara, M., Steer, C., Goodman, R., Ford, T., & Lamberts, K. (2009). Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. *British Journal of Psychiatry, 195*(3), 249-256. doi:10.1192/bjp.bp.108.053751

Hallett, V., Ronald, A., Rijsdijk, F., & Eley, T. C. (2009). Phenotypic and genetic differentiation of anxiety-related behaviors in middle childhood. *Depression and Anxiety, 26*(4), 316-324. doi:10.1002/da.20539

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of educational research, 75*(3), 417-453. doi:10.3102/00346543075003417

Eley, T. C., Bolton, D., O'Connor, T. G., Perrin, S., Smith, P., & Plomin, R. (2003). A twin study of anxiety-related behaviours in pre-school children. *Journal of Child Psychology and Psychiatry, 44*(7), 945-960. doi:10.1111/1469-7610.00179

Gilman, S. E., Kawachi, I., Fitzmaurice, G. M., & Buka, S. L. (2002). Socioeconomic status in childhood and the lifetime risk of major depression. *International Journal of Epidemiology, 31*(2), 359-367. doi:10.1093/ije/31.2.359

Scher, S. J., & Osterman, N. M. (2002). Procrastination, conscientiousness, anxiety, and goals: Exploring the measurement and correlates of procrastination among school-aged children. *Psychology in the Schools*, *39*(4), 385-398. doi:10.1002/pits.10045

Fergusson, D. M., & Woodward, L. J. (2002). Mental health, educational, and social role outcomes of adolescents with depression. *Archives of General Psychiatry*, *59*(3), 225-231. doi:10.1001/archpsyc.59.3.225

Woodward, L. J., & Fergusson, D. M. (2001). Life course outcomes of young people with anxiety disorders in adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(9), 1086-1093. doi:10.1097/00004583-200109000-00018

Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behaviour Research and Therapy*, *36*(5), 545-566. doi:10.1016/s0005-7967(98)00034-5

Wechsler, D. (1992). *Wechsler intelligence scale for children* (3rd ed.). London: Psychological Corporation.

McCarthy, D. (1972). *Manual for the McCarthy scales of children's abilities*: Psychological Corporation.

Table 1: Sample characteristics and gender differences in ARBQ composite scores and academic outcomes.

	All	Female	Male	Gender differences		
				Test statistic	df	*p*
Total Sample, N (%)	10,791 (100%)	5,775 (53.53%)	5,016 (46.47%)			
Complete Pairs (n)	5,392					
Social Anxiety, M (SD)	1.36 (1.36)	1.47 (1.40)	1.23 (1.30)	*t*=9.49	10791	<.001
Fear, M (SD)	1.57 (1.38)	1.73 (1.40)	1.38 (1.32)	*t*=13.43	10791	<.001
Negative Affect, M (SD)	1.15 (1.18)	1.18 (1.21)	1.11 (1.18)	*t*=2.97	10791	<.001
Negative Cognition, M (SD)	2.15 (1.87)	2.25 (1.90)	2.03 (1.83)	*t*=5.90	10791	<.001
GCSEs, M (SD)	8.90 (1.23)	9.02 (1.19)	8.92 (1.23)	*t*=3.97	8874	<.001
A-levels[1], M (SD)	3.48 (1.64)	3.67 (1.47)	3.21 (1.82)	*t*=9.49	5970	<.001
University Education[2], N (%)	4948 (45.85%)	2870 (49.70%)	2078 (41.42%)	χ^2=52.44	1	<.001
MZ Twin Differences, N (%)	3,969 (36.77%[3])	2,215 (55.81%)	1,754 (44.19%)			
Complete Pairs (n)	1,981	1,108	873			
Social Anxiety, M (SD)	0.00 (0.92)	0.00 (0.96)	0.00 (0.86)			
Fear, M (SD)	0.00 (0.89)	0.00 (0.89)	0.00 (0.90)			
Negative Affect, M (SD)	0.00 (0.83)	0.00 (0.84)	0.00 (0.82)			
Negative Cognition, M (SD)	0.00 (1.56)	0.00 (1.56)	0.00 (1.55)			
GCSEs, M (SD)	0.00 (0.58)	0.00 (0.55)	0.00 (0.63)			
A-levels, M (SD)	0.00 (0.93)	0.00 (0.87)	0.00 (1.00)			

Notes: [1] A-level grade points ranged from 0-6, with a score of 0 designated to participants who did not study post-16. [2] N indicating the number of participants who transition to higher education. [3] Percentage within the whole sample. M = mean; SD = standard deviation; df = degrees of freedom; ARBQ = Anxiety and Related Behaviours Questionnaire.

Table 2: Results of regression models showing the phenotypic association between internalising symptom subtypes and academic attainments in the whole sample.

	Univariate models			Multivariable models		
GCSEs						
	***β* (95% CI)**	***t***	***p***	***β* (95% CI)**	***t***	***p***
Social anxiety	-.04 (-.06, -.01)	-2.80	<.01	.02 (-.01, .05)	1.50	.134
Fear	-.06 (-.08, -.03)	-4.54	<.001	-.02 (-.04, .01)	-1.25	.212
Negative affect	-.12 (-.15, -.09)	-9.20	<.001	-.10 (-.13, -.07)	-6.21	<.001
Negative cognition	-.09 (-.12, -.07)	-7.42	<.001	-.04 (-.07, -.01)	-2.86	<.01
A-Levels						
	***β* (95% CI)**	***t***	***p***	***β* (95% CI)**	***t***	***p***
Social anxiety	-.01 (-.04, .02)	-0.73	.463	.02 (-.02, .05)	1.06	.288
Fear	-.02 (-.05, .01)	-1.25	.213	.01 (-.03, .04)	0.43	.669
Negative affect	-.09 (-.12, -.05)	-5.27	<.001	-.09 (-.13, -.05)	-4.39	<.001
Negative cognition	-.05 (-.08, -.02)	-3.12	<.01	-.01 (-.05, .02)	-0.76	.445
Transition university education						
	Odds Ratio (95% CI)	***z***	***p***	**Odds Ratio (95% CI)**	***z***	***p***
Social anxiety	.98 (.92, 1.05)	-0.60	.550	1.06 (.98, 1.14)	1.40	.161
Fear	.94 (.88, 1.01)	-1.73	.083	.99 (.92, 1.07)	-0.20	.839
Negative affect	.81 (.75, .87)	-6.04	<.001	.81 (.74, .87)	-5.25	<.001
Negative cognition	.90 (.84, .96)	-3.14	<.01	.99 (.91, 1.08)	-0.25	.803

Note: Sample size varied across analyses. For univariate analyses: *n* = 5,686 – 5,694 for GCSEs; *n* = 3,963 – 3,969 for A-levels; *n* = 5,319 – 5, 329 for transition to higher education. For multivariable analyses: *n* = 5,674 for GCSEs; *n* = 3,957 for A-levels; *n* = 5,309 for transition to university education. All analyses adjusted for sex, IQ, SES and relatedness of twin members using robust clustering.

Table 3: Results of univariate and multivariable regression models predicting MZ twin differences in academic attainments from MZ twin differences in internalising symptom subtypes.

	Univariate models			Multivariable models		
GCSEs						
	***β* (95% CI)**	***t***	***p***	***β* (95% CI)**	***t***	***p***
Social anxiety	-.04 (-.08, .00)	-1.89	.059	-.01 (-.06, .04)	-0.38	.703
Fear	-.04 (-.09, -.00)	-1.85	.064	-.02 (-.06, .03)	-0.73	.363
Negative affect	-.03 (-.08, .01)	-1.45	.148	.00 (-.05, .05)	0.18	.859
Negative cognition	-.04 (-.06, -.01)	-3.18	<.01	-.03 (-.06, -.01)	-2.28	<.05
A-Levels						
	***β* (95% CI)**	***t***	***p***	***β* (95% CI)**	***t***	***p***
Social anxiety	-.04 (-.09, .02)	-1.37	.170	-.03 (-.09, .03)	-0.94	.347
Fear	.01 (-.05, .07)	0.29	.774	.03 (-.03, .10)	0.98	.329
Negative affect	-.02 (-.08, .05)	-0.50	.619	.01 (-.06, .09)	0.34	.733
Negative cognition	-.03 (-.06, .01)	-1.51	.131	-.03 (-.07, .02)	-1.26	.208
University education transition	**Odds Ratio (95% CI)**	***Z***	***p***	**Odds Ratio (95% CI)**	***Z***	***p***
Social anxiety	.97 (.82, 1.15)	-0.33	.740	1.01 (.83, 1.23)	0.13	.900
Fear	1.06 (.89, 1.27)	0.64	.522	1.11 (.92, 1.35)	1.09	.278
Negative affect	.88 (.72, 1.09)	-1.19	.234	.91 (.72, 1.15)	-0.81	.418
Negative cognition	.94 (.85, 1.04)	-1.18	.240	.94 (.82, 1.07)	-0.98	.329

Note: Sample size varied across analyses. For univariate analyses: n = 1,103 – 1,108 for GCSEs; n = 692 – 696 for A-levels; n = 1,039 – 1,044 for university transition. For multivariable analyses: n = 1,101 for GCSEs; n = 691 for A-levels; n = 1,035 for university transition. All analyses controlled for MZ difference in IQ (not sex or SES since these variables are identical between members of MZ twin pair).

Supplementary Material

Table S1: ARBQ items and subscales

Subscale	ARBQ Item
Negative cognition	Often critical of him/herself
	Tends to blame him/herself
	Has low self-confidence
	Many worries, often seems worried†
	Anxious that bad things will happen
	Asks for reassurance that s/he is OK
Negative affect	Often unhappy, downhearted, or tearful†
	Does not enjoy him/herself
	Seems keyed up, on edge, tense
	Does something over and over again
	Complains or whines
	Is often extremely upset or distressed when parent leaves
Fear	Is afraid of small closed spaces, heights, water, or the dark
	Many fears, easily scared†
	Is afraid of animals or insects
	Resists sleeping alone†
	Afraid of medical procedures e.g. going to doctor or dentist
Social anxiety	Tends to be shy or timid
	Is afraid in social situations
	Takes a long time to warm to strangers
	Nervous or clingy in new situations, often loses confidence†

Note: The following items were excluded from the analysis due to inconsistent factor loading across assessment points: '*Has tics or twitches'*; *'Complains of stomach-aches, headaches, or sickness*'; '*Tends to check that things are done exactly right*'; '*Fussy or overparticular*'; '*Fussy about keep hands clean*'. † These items were included in the ARBQ at ages 7 and 9 but not at age 16. Abbreviations: ARBQ = Anxiety and Related Behaviours Questionnaire.

Table S2: ARBQ subscales and composites in the whole TEDS sample[1]

	Age 7	Age 9	Age 16	Composite across ages
Negative Affect	N=14,926	N=6614	N=9921	N=10,823
M (SD)	1.66 (1.73)	1.22 (1.55)	0.57 (1.11)	1.15 (1.19)
Internal Consistency (α)	.56	.59	.62	.74
Skew (raw data)	1.50	1.77	2.83	1.80
Skew (log+1 transformed data)	0.15	0.53	1.46	0.68
Stability over time (r_s)				
Age 7	-	.46*	.28*	-
Age 9	-	-	.33*	-
Age 16	-	-	-	-
Negative Cognition	N=14,927	N=6622	N=9913	N=10,834
M (SD)	2.76 (2.50)	2.83 (2.57)	1.01 (1.59)	2.15 (1.87)
Internal Consistency (α)	.74	.77	.74	.84
Skew (raw data)	1.03	1.06	2.14	1.26
Skew (log+1 transformed data)	-0.16	-0.16	0.93	0.29
Stability over time (r_s)				
Age 7	-	.50*	.31*	-
Age 9	-	-	.36*	-
Age 16	-	-	-	-
Social Anxiety	N=14,940	N=6641	N=9935	N=10,852
M (SD)	1.81 (1.86)	1.49 (1.75)	0.74 (1.23)	1.36 (1.36)
Internal Consistency (α)	.75	.74	.78	.84
Skew (raw data)	1.00	1.25	1.84	1.22
Skew (log+1 transformed data)	0.09	0.35	1.07	0.49
Stability over time (r_s)				
Age 7	-	.56*	.35*	-
Age 9	-	-	.37*	-
Age 16	-	-	-	-
Fear	N=14,938	N=6638	N=9944	N=10,854
M (SD)	2.25 (2.01)	1.80 (1.82)	0.64 (0.98)	1.57 (1.38)
Internal Consistency (α)	.57	.58	.41	.74
Skew (raw data)	0.94	1.17	1.80	1.23
Skew (log+1 transformed data)	-0.16	0.09	0.95	0.33
Stability over time (r_s)				
Age 7	-	.54*	.29*	-
Age 9	-	-	.37*	-
Age 16	-	-	-	-

Notes: Subscales at ages 7, 9, 16: Mean, SD, and internal consistency of untransformed scores at ages 7, 9, 16, and composites are presented. Subscale score ranges: negative cognition range 0-12 (age 7, age 9) and 0-10 (age 16); negative affect range 0-12 (age 7, age 9) and 0-10 (age 16); fear range 0-10 (age 7, age 9) and 0-6 (age 16); social anxiety range 0-8 (age 7, age 9) and 0-6 (age 16). Correlation analyses were undertaken on transformed data. * $p<.001$. Abbreviations: M = mean; SD = standard deviation; ARBQ = Anxiety and Related Behaviours Questionnaire.

Table S3: Variance inflation factors for multivariable regression models

		GCSEs		A-levels		Transition higher education	
Phenotypic analyses		VIF	1/VIF	VIF	1/VIF	VIF	1/VIF
	Negative cognition	1.64	.61	1.64	.61	1.64	.61
	Negative affect	1.49	.67	1.48	.67	1.49	.67
	Social anxiety	1.35	.74	1.36	.73	1.35	.74
	Fear	1.35	.74	1.34	.75	1.35	.74
MZ twin difference analyses		VIF	1/VIF	VIF	1/VIF	VIF	1/VIF
	Negative cognition	1.52	.66	1.54	.65	N/A	N/A
	Negative affect	1.31	.76	1.31	.76		
	Social anxiety	1.29	.78	1.32	.76		
	Fear	1.17	.86	1.17	.85		

Note: VIF = variance inflation factor

Table S4. Missing Predictor Data

	Internalising Subscales		
	Present N=10,791 55.68%†	**Missing** N=8,5880 44.32%†	**Test statistic**
Gender			$\chi^2 = 53.28$*
Male	n=5016	n=4446	
Female	n=5775	n=4142	
SES	M=0.20 [0.98]	M=-0.05 [0.98]	$t(18{,}006) = 16.64$*

Notes: SES = socioeconomic status; M = mean; numbers in [] = standard deviation; † % of total potential sample; * = $p<.0001$.

Table S5. Missing Outcome Data

	GCSEs			A-Levels			Transition to Higher Education		
	Present N=8,876 82.24%†	**Missing** N=1,195 17.76%†	**Test statistic**	**Present** N=6,309 58.47%†	**Missing** N=4,482 41.53%†	**Test statistic**	**Present** N=8,550 79.23%†	**Missing** N=2,241 20.77%†	**Test statistic**
Gender			χ^2 = 15.66*			χ^2=23.26*			χ^2 = 21.66*
Male	n=4047	n=969		n=2809	n=2207		n=3876	n=1140	
Female	n=4829	n=946		n=3500	n=2275		n=4674	n=1101	
SES	M=0.28	M=-0.16	*t*(10,250) = 16.99*	M=0.40	M=-0.08	*t*(10,250) = 25.16*	M=0.29	M=-0.14	*t*(10,250) = 17.82*
	[0.96]	[1.00]		[0.96]	[0.95]		[0.97]	[0.97]	
Negative Affect	0.23	0.31	*t*(10,789) = -16.63*	0.22	M=0.28	*t*(10,789) = -17.11*	M=0.23	M=0.31	*t*(10,789) = -17.50*
	[0.19]	[0.23]		[0.19]	[0.21]		[0.19]	[0.22]	
Negative Cognition	0.37	0.45	*t*(10,789) = -13.95*	M=0.36	M=0.41	*t*(10,789) = -11.78*	M=0.36	M=0.45	*t*(10,789) = -14.68*
	[0.24]	[0.26]		[0.24]	[0.25]		[0.24]	[0.25]	
Social Anxiety	0.27	0.31	*t*(10,789) = -8.60*	M=0.26	M=0.29	*t*(10,789) = -7.75*	M=0.27	M=0.31	*t*(10,789) = -7.94*
	[0.22]	[0.24]		[0.22]	[0.23]		[0.22]	[0.24]	
Fear	0.30	0.36	*t*(10,789) = -11.51*	M=0.30	M=0.33	*t*(10,789) = -8.72*	M=0.30	M=0.35	*t*(10,789) = -9.87*
	[0.21]	[0.24]		[0.21]	[0.22]		[0.21]	[0.23]	

Notes: SES = socioeconomic status; M = mean; numbers in [] = standard deviation; † % of study sample; * = p<.0001.

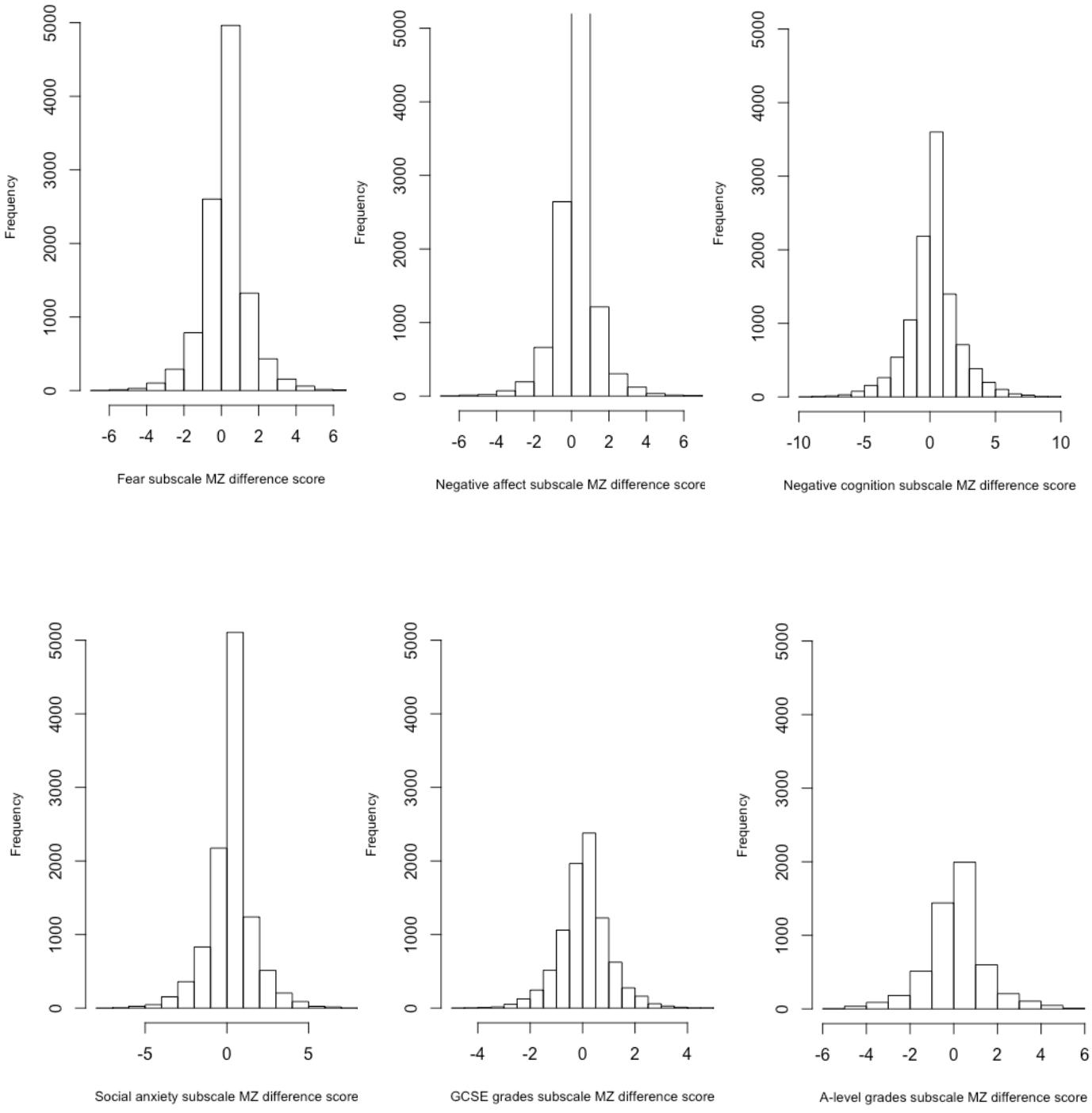

Figure S1. Histograms of distribution of MZ difference scores for Internalising Subscale Scores (ARBQ) and GCSE and A-level grades.