

A genome-wide test of the differential susceptibility hypothesis reveals a genetic predictor of differential response to psychological treatments for child anxiety disorders

Robert Keers PhD<sup>1,2</sup>, Jonathan R.I. Coleman MSc<sup>2</sup>, Kathryn J. Lester DPhil<sup>2,3</sup>, Susanna Roberts MSc<sup>2</sup>, Gerome Breen PhD<sup>2</sup>, Mikael Thastum PhD<sup>4</sup>, Susan Bögels PhD<sup>5</sup>, Silvia Schneider PhD<sup>6</sup>, Einar Heiervang MD PhD<sup>7,8</sup>, Richard Meiser-Stedman PhD<sup>9</sup>, Maaïke Nauta PhD<sup>10</sup>, Cathy Creswell DClInPsy PhD<sup>11</sup>, Kerstin Thirlwall DClInPsy<sup>11</sup>, Ronald M. Rapee PhD<sup>12</sup>, Jennifer L. Hudson PhD<sup>12</sup>, Cathryn Lewis PhD<sup>2</sup>, Robert Plomin PhD<sup>2</sup> and Thalia C. Eley PhD<sup>2</sup>

1. Department of Biological and Experimental Psychology, School of Biological and Chemical Sciences, Queen Mary University of London
2. King's College London, Institute of Psychiatry Psychology and Neuroscience, MRC Social, Genetic and Developmental Psychiatry (SGDP) Centre
3. University of Sussex, Brighton, UK
4. Department of Psychology, University of Aarhus, Aarhus, Denmark
5. Research Institute Child Development and Education, University of Amsterdam, Amsterdam
6. Ruhr-Universität Bochum, Bochum, Germany
7. Institute of Clinical Medicine University of Oslo, Norway and Anxiety Research Network
8. Haukeland University Hospital, Bergen, Norway
9. University of East Anglia, Norwich, UK
10. University Medical Center Groningen, University of Groningen, The Netherlands
11. School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK
12. Centre for Emotional Health, Department of Psychology, Macquarie University, Sydney, Australia

Corresponding Author: Thalia Eley (thalia.eley@kcl.ac.uk)

## **Disclosure**

Schneider is an author of the Diagnostisches Interview bei psychischen Störungen im Kindes- und Jugendalter, from which she receives royalties. Creswell is joint author of a book used in treatment within the Overcoming trial and receives royalties from sales of the book. Rapee is an author of the Cool Kids program but receives no direct payment from it. Hudson is an author of the Cool Kids program but receives no direct payment from it. Keers, Coleman, Lester, Roberts, Breen, Thastum, Bogels, Heiervang, Meiser-Stedman, Nauta, Thirlwall, Lewis, Plomin and Eley report no potential conflicts of interest.

## **Acknowledgements**

We gratefully acknowledge the ongoing contribution of the participants in the Twins Early Development Study (TEDS) and their families and participants in the GxT study.

TEDS is supported by a program grant to R.P. from the UK Medical Research Council (G0901245; previously G0500079), with additional support from National Institutes of Health Grants HD044454 and HD059215. The GxT study was supported by UK MRC grant G0901874/1 (T.C.E.) and individual trials supported by Australian Research Council grant DP0878609 (J.L.H., Jenny Donald, PhD, R.M.R., T.C.E.); Australian National Health and Medical Research Council grants (1027556: R.M.R., J.L.H., H.J.L., Cathy Mihalopolous, BSc[Hons], PhD), (488505: H.J.L., J.L.H., R.M.R.), and (382008: J.L.H. and R.M.R.); TrygFonden grant (7-10-1391: M.T. and Esben Hougaard, PhD); Edith og Godtfred Kirk Christiansens Fond grant (21-5675: M.T.); Swiss National Science Foundation grant (105314-116517: S.S.); Western Norway Regional Health Authority grants (911253: Odd E. Havik, PhD) and (911366: E.H.); UK MRC Clinical Fellowship (G0802821: R.M.-S.); National Institute of Mental Health R01MH079943 (W.K.S.); UK National Institute for Health Research (NIHR) grants (PB-PG-0110-21190: C.C.) and (PB-PG-0107-12042: P.C.); UK MRC grants (09-800-17: P.C. and C.C.; G0802326: K.T., P.C., and C.C.; G1002011: P.W., C.C., and P.C.; and G0601874: C.C.). Grant 09-800-17 was managed by NIHR on behalf of the MRC-NIHR partnership. Dr. Lester is supported by an MRC research grant (MR/J011762/1) and Jacobs Foundation Young Scholar Award. Dr. Keers is supported by an MRC Population Health Scientist Award (MR/K021281/1). This study presents independent research partly funded by the NIHR Biomedical Research Centre at South London and Maudsley, NHS Foundation Trust, and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health.

The authors thank Irma Knuistingh Neven, BEd, MSc, Kathryn Preece, BSc(Psych) PGDip, Annabel Price, BSc (Hons), and Julia White, BCom, BSc-Psych (Hons), Grad Dip Psych, of Macquarie University; Marianne Bjerregaard Madsen, administrator, Hjalti Jonsson, PhD, and Lisbeth Jørgensen, MSc, of Aarhus University; Carmen Adornetto, DPhil, Judith Blatter-Meunier, DPhil, Chantal Herren, DPhil, of University of Basel; Kristine Fonnes Griffin, administrator, Odd E. Havik, PhD, and Krister Fjermestad, PhD, of University of Bergen; Elisa Kulewski, MSc, Jonas Großekathöfer, MSc, Nora Dirks, BSc, Evelyn Kmelnitski, BSc, Sabrina Heuser, Dipl Psych, Karen Krause, Dipl Psych, of Bochum University; Adrian Boyle, MD, FCEM, Clare Dixon, BSc, MSc, Patrick Smith, MA, MSc, PhD, and Tim Dalgleish, MA, PhD, MSc, of University of Cambridge; Harma Moorlag, PhD, Nienke Boersma, MSc, and Sanne Hogendoorn, MSc, of University of Groningen; Sue Cruddace, PhD, Marie Weber, student, Zoe Hughes, BSc, Liz White, BSc, Françoise Hentges, PhD, DClínPsy, and Kiri Clarke, MSc, of Reading University; Nynke Wagenaar, MSc, of University of Amsterdam; Luci Motoca, PhD, Yesenia Rodriguez, MS, Klaudia Pereira, MSc, Cristina Del Busto, MS, Erin Hedemann, MS, and Yasmin Rey, PhD, of Florida International University; and Gillian O'Neill, BSc (Hons), Lucy Lindley, BSc (Hons), Rachel Banham, BSc (Hons), and Nora Delvendahl, BSc (Hons), of King's College London for their contributions.

## **Abstract**

*Background:* The differential susceptibility hypothesis suggests that certain genetic variants moderate the effects of both negative *and* positive environments on mental health and may therefore be important predictors of response to psychological treatments. Nevertheless, the identification of such variants has so far been limited to preselected candidate genes. In this study we extended the differential susceptibility hypothesis from a candidate gene to a genome-wide approach to test whether a polygenic score of environmental sensitivity predicted response to Cognitive Behavioural Therapy (CBT) in children with anxiety disorders.

*Methods:* We identified variants associated with environmental sensitivity using a novel method in which within-pair variability in emotional problems in 1026 monozygotic (MZ) twin pairs was examined as a function of the pairs' genotype. We created a polygenic score of environmental sensitivity based on the whole-genome findings and tested the score as a moderator of parenting on emotional problems in 1,406 children and response to individual, group and brief parent-led CBT in 973 children with anxiety disorders.

*Results:* The polygenic score significantly moderated the effects of parenting on emotional problems and the effects of treatment. Individuals with a high score responded significantly better to individual CBT than group CBT or brief parent-led CBT (remission rates: 70.9%, 55.5% and 41.6% respectively).

*Conclusions:* Pending successful replication, our results should be considered exploratory. Nevertheless, if replicated, they suggest that individuals with the greatest environmental sensitivity may be more likely to develop emotional problems in adverse environments, but also benefit more from the most intensive types of treatment.

## Introduction

Anxiety disorders are by far the most prevalent mental disorder and amongst the earliest to emerge, with the vast majority of adult cases beginning in childhood or adolescence [1].

While there is a substantial evidence base for the efficacy of psychological treatments for anxiety in children, response to treatment varies substantially between patients [2]. This means that identifying an effective treatment can be a long and costly process of trial and error that may both delay recovery and have a negative effect on long-term outcome.

Genetic predictors of treatment response may allow clinicians to select the most effective treatment for a given individual at the outset, enhancing outcomes and accelerating recovery times [3]. Such predictors could also offer valuable insights into the mechanisms underlying response to psychological treatments [4].

The “differential susceptibility hypothesis” suggests that genetic factors moderate the effects of both negative *and* positive environments on mental health: for better *and* for worse [5]. In line with this hypothesis, individuals with one or two copies of the short allele of the 5-HTTLPR have been shown to be at a greater risk of mood disorders following adversity than individuals homozygous for the long allele [6]. However, these same individuals also benefit more from *positive* environmental influences such as supportive parenting [7], positive life events [8], or social support [9]. Importantly, these associations have also been shown to extend to moderation of the positive effects of various interventions including psychosocial training on depression [10], high-quality foster care on disturbances of attachment [11] and externalizing behaviour [12], and the efficacy of Cognitive Behavioural Therapy (CBT) in children with anxiety disorders [13]. In addition to findings from the 5-HTTLPR, differential susceptibility has been reported for a small number of further markers [14] with results from intervention studies showing particular promise [15]. Nevertheless, findings have failed to replicate, even in high quality studies, with very similar methodologies [16] [17]. While the causes of non-replication are unclear, one

explanation is that environmental responsivity is a complex, polygenic trait, which is the result of multiple genetic variants of small effect, rather than a few select candidate genes.

Gene-environment interaction research therefore needs to move from a candidate gene to genome-wide methodology, which takes into account the aggregate effects of multiple variants [18].

Polygenic scoring allows the effects of multiple variants to be summarized in a single score. Specifically, alleles associated with a trait in a discovery sample at a given p value threshold are selected in an independent validation sample, and a score (the sum of these alleles weighted by their effect size) created for each individual [19]. Using this approach, a recent study reported that a polygenic score calculated using the results of a large case-control study of major depression moderated the effects of childhood maltreatment on depression in a further sample, with the interaction explaining a further 0.5% of the variance [20]. This approach to whole genome gene-environment interaction relies on the assumption that genetic variants have a main effect on outcome. This means that while this method may be suitable for detecting variants implicated in diathesis-stress interactions, it may not detect those involved in differential susceptibility, which are proposed to have no main effects [5]. One means of targeting these variants, is to explore genetic effects on intra-pair variability in outcomes in monozygotic (MZ) twin pairs. As they are genetically identical and share the same environment, discordance within MZ twin pairs on a measured outcome is considered to be the result of non-shared environmental effects. However, twin pairs with variants associated with increased sensitivity to the environment may have a greater intra-pair variability in outcome due to their increased responsivity to unmeasured non-shared environmental influences [21]. While this method has been previously used in a genome-wide study of metabolism [22], it is yet to be applied to analyses of mental health outcomes. Moreover, this approach is yet to incorporate polygenic scoring to consider of the aggregate effects of variants associated with environmental sensitivity.

In this study we aimed, for the first time, to test the differential susceptibility hypothesis using a genome-wide approach. First, we examined associations between genetic variants and intra-pair variability in emotional problems in MZ twins using genome-wide data. Next, in order to validate these findings, we calculated a polygenic score of sensitivity to the environment and tested whether this score moderated the effects of positive and negative parenting on emotional problems in a further sample of children. Finally, to test whether these same variants moderated response to psychological treatment, we tested the same polygenic environmental sensitivity score as a predictor of treatment response in a further clinical sample of children and adolescents with anxiety disorders treated with individual CBT, group CBT or brief parent-led CBT.

In addition to examining an effect of the polygenic environmental sensitivity score on overall treatment response, we also explored whether the polygenic score predicted differential response to the different types of treatment received. The effect of environmental sensitivity on response to psychological treatments with differing intensities remains unknown. It has been suggested that those with a low sensitivity to the environment may require a more intensive type of treatment to achieve the same results as those who are highly sensitive. In this case, individuals with a low sensitivity would respond better to individual CBT than brief parent-led CBT. Conversely, it has also been argued that individuals with a high sensitivity to the environment may benefit the most from more intensive forms of treatment. In this case individuals with a high sensitivity would respond more favourably to individual CBT, compared with lower intensity treatments such as brief parent-led CBT.

## **Methods**

### Samples

This study utilized three samples: a discovery sample, a validation sample and a treatment response sample.

#### *Discovery and validation samples*

The discovery and validation samples were both drawn from the Twins Early Development Study (TEDS). TEDS is an ongoing longitudinal study of more than 11,000 twin pairs born in England and Wales in 1994, 1995, and 1996, which has been shown to be representative of the UK population [23]. The discovery sample included 1026 monozygotic twin pairs from TEDS for whom genome-wide genotyping data were available, as well as data on emotional problems at age 12. The validation sample included a further 1,409 unrelated individuals from TEDS (a randomly selected individual from the remaining dizygotic twins pairs) with available data.

#### *Measures*

Emotional problems were measured in the discovery and validation samples at age 12 using the emotional symptoms subscale of the Strengths and Difficulties Questionnaire [24]. We created a composite score by summing the z scores from child and parent reports and dividing by two. Parenting was assessed at age 12 in the validation sample using two child-report measures: the Parental Feelings Questionnaire (PFQ) [25] and the Parental Strategies Questionnaire [26]. The PFQ includes 7 statements on the relationship with their parent on a 3-point scale (*very true, quite true, not true*). The measure included four negative items (e.g. "I make my parents angry") and three positive (e.g. "I feel happy about my relationship with my parents"). Positive items were reversed so that the total score reflected parental negativity. The Parental Strategies Questionnaire included four items in which children were

asked to rate on a 3 point scale (*rarely/never, sometimes, and often*) what their parent did if they misbehaved including two positive (e.g., “Explain or reason with me”) and two negative (e.g., “They give me a smack”) items. Positive items were reversed so that the total score reflected a more negative discipline strategy. An overall parenting score was created by summing the standardised scores from both scales. Separate positive and negative parenting scores were created by selecting the positive and negative items from each scale as reported previously in the TEDS data [27].

#### *Genetic data and quality control*

Both the discovery and validation samples were genotyped as part of the larger TEDS study. Full details of genotyping and quality control are provided elsewhere [28]. In brief, DNA was extracted from buccal cheek swabs samples and genotyped using Affymetrix GeneChip 6.0 SNP genotyping arrays. Individuals were removed for a low call rate or excessive heterozygosity, atypical population ancestry, relatedness or sample duplication and gender mismatches. SNPs were excluded if they had a call rate less than 98%, minor allele frequency less than 1% or a Hardy Weinberg  $p$ -value lower than  $1 \times 10^{-20}$ . Following quality control 679,050 SNPs remained for analysis.

#### *The treatment response sample*

The treatment response sample was drawn from the Genes for Treatment (GxT) study, a multi-site collaboration, including 1,519 individuals, which was designed to examine genetic and clinical predictors of response to psychological treatments in pediatric anxiety disorders. Full details of the sample are available elsewhere [29]. In brief, participants were included if they were aged 5-18 years (94% were 5-13 years old), met DSM-IV criteria for a primary diagnosis of an anxiety disorder and provided DNA. Parents provided written consent and children written or verbal assent. All sites administered the Anxiety Disorders Interview

Schedule for DSM-IV, Parent and Child Versions (ADIS-IV-C/P:[30]) except in two sites where the German equivalent, Kinder-DIPS, was used [31]. Participants were assessed before and immediately after treatment (post-treatment), with further assessments made 3, 6, or 12 months after treatment cessation where possible (follow-up). The severity of the primary anxiety disorder was measured at each time-point using the Clinicians Severity Rating (CSR) from the structured interview, which assigns a score of 0-8 (absent to very severe). A diagnosis was made when the child met the diagnostic criteria and received a CSR of 4 or more. Ten sites (n=1,396) also assessed comorbid mood (major depression or dysthymia) or externalizing disorders (oppositional defiant disorder, conduct disorder or attention-deficit/hyperactivity disorder [ADHD]) at baseline using the ADIS-C/P. All assessments were completed by graduate assistants or clinical staff (mainly psychologists) trained in the administration of the instruments. Sites have previously reported good inter-rater reliability for the diagnostic instruments using these samples [32-34]. In eight sites (n=1,289), parents also completed the Depression Anxiety Stress Scales (DASS) [35], assessing depression, anxiety, and stress symptoms experienced over the past week. For this study, the 3 subscales were summed to create an overall measure of parental psychopathology.

Of the 980 participants with available genome-wide genotyping data and at least one post baseline assessment, 269 (27.5%) were treated with individual CBT, 503 (51.3%) with group-based CBT, 201 (21.2%) with brief parent-led CBT and 7 (0.7%) with guided self-help CBT. In order to limit the heterogeneity of the sample and aid interpretation of treatment specific effects, individuals treated with guided self-help CBT were excluded from the analysis. For the remaining 973 participants, (female: 54.9%, mean age: 9.8, SD=2.2), primary diagnoses included Generalised Anxiety Disorder (GAD; n=362, 37.2%), Social Anxiety Disorder (SoAD; n=201, 20.7%), Specific Phobia (SP; n=106, 10.9%), Separation Anxiety Disorder (SAD; n=223; 22.9%). The remaining participants (n=81, 8.3) were grouped as “other” anxiety disorders which included panic disorder with and without agoraphobia

and agoraphobia without panic disorder (n = 26), obsessive-compulsive disorder (OCD; n = 34), posttraumatic stress disorder (PTSD; n = 13), selective mutism (in patients with primary selective mutism, a diagnosis of severe SoAD was also given; the selective mutism was considered by the clinician to be primary, the most interfering: n = 2) or anxiety disorder not otherwise specified (n = 6).

#### *Genetic data and quality control*

Genotyping and quality control procedures for the GxT study are documented elsewhere [36]. In brief, DNA was extracted from buccal swabs and saliva and genotyped on the Illumina Human Core Exome-12v1.0 microarrays. Individuals with a call rate <99% or excessive heterozygosity were removed, as well as those with gender mismatches or evidence for relatedness or sample duplication. SNPs were excluded if they had a call rate less than 99%, minor allele frequency of less than 5% or had a Hardy Weinberg  $p$ -value lower than  $1 \times 10^{-5}$ . Quality-controlled data was imputed to the December 2013 release of the 1000 using IMPUTE2. Only SNPs with an info metric >0.8, and with a minor allele frequency (MAF) >1% were retained for analysis.

#### Analyses

##### *Discovery sample*

Discordance in emotional symptom score was calculated as the absolute difference in scores between members of the pair. The effects of age, sex and the twin pair's mean score in emotional symptoms were regressed out to create a residual score, which was then included as an outcome variable in a linear regression in PLINK. In order to control for possible effects of population stratification, we included the first 10 principal components from previous analyses of the TEDS data [28] as covariates in all analyses.

### *Validation sample*

In the validation sample we aimed to test whether the environmental sensitivity polygenic score moderated the effects of parenting on emotional problems. Polygenic scores were calculated for each individual in the sample using the betas and p values from the discovery sample.

We used increasingly liberal significance thresholds to select 8 sets of SNPs from the discovery sample that reached  $P < 0.001$ , 0.01, 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5. Prior to inclusion, SNPs were pruned for linkage disequilibrium using p-value informed clumping in PLINK employing cut offs of LD ( $r^2 = 0.25$ ) and distance (a 200kb window).

As in the discovery sample, we created a standardised age and sex regressed residual score of emotional symptoms for individuals in the validation sample. We explored the main effects of the polygenic environmental sensitivity score and parenting on this outcome using linear regressions. Next, we tested whether the polygenic score moderated the effects of parenting on emotional problems by testing a polygenic score by parenting interaction term in these models. The presence of a gene-environment correlation (i.e. an effect of the polygenic score on parenting) could potentially bias any polygenic score by parenting interactions. We therefore also tested whether our measures of parenting were associated with the polygenic score using linear regressions. We included socio-economic status (SES) as a covariate, as well as the first 10 principal components previously derived from genome wide analyses of the TEDS data [28] in order to account for any population stratification effects.

### *Treatment response sample*

In the treatment response sample, we aimed to test whether the polygenic environmental sensitivity score predicted response to psychological treatments. We defined treatment response in the GxT sample as the change in severity (CSR score) of the primary anxiety

diagnosis from baseline to each time-point in the study including measurements from the post-treatment, 3, 6 and 12 month time points. In order to include all of the available outcome data simultaneously, and provide estimates in the presence of missing values, we used a linear mixed model fitted with full maximum likelihood.

We constructed a model including the fixed effects of baseline severity (CSR score of the primary diagnosis at baseline, centred at the mean) and the linear and quadratic effects of time to account for the curvilinear slope of treatment outcome. To account for correlations between repeated measures from the same subject all models included the random effects of individual. We also included a higher order random effect of trial to account for between trial differences. As in previous analyses, we covaried for clinical and demographic covariates including age, sex, primary diagnosis and treatment type by including these as fixed effects. We also included the first 10 principal components generated from previous genome wide analyses of the GxT data to account for confounding caused by population stratification.

A polygenic environmental sensitivity score was calculated for each individual in the GxT sample using the same approach as in the validation sample and entered into the above model as a fixed effect. First, we tested the effects of the polygenic score on overall treatment response. Next, we tested treatment specific effects by examining the effects of the polygenic score separately in participants treated with individual CBT, group CBT or brief parent-led CBT and by testing for treatment type by polygenic score interactions.

## Results

### *Discovery analyses*

In total 1026 monozygotic twin pairs (56.9% female, Mean age: 11.28, SD=0.02) with available genome-wide genotyping data and data on emotional symptoms were included in the discovery analyses. None of the included 679,050 SNPs reached genome-wide significance, nevertheless, several suggestively significant findings ( $P < 1 \times 10^{-5}$ ) were identified and are described in Table 1.

### *Validation analyses*

The validation sample included 1,406 unrelated individuals with available data. The sample was significantly younger than the discovery sample (mean age = 11.20, SD=0.70;  $t=2.53$ ,  $P=0.010$ ) and included significantly fewer females (52.1%;  $\chi^2=5.33$ ,  $P=0.021$ ). However, individuals did not differ in their mean emotional symptoms scores ( $t=-1.50$ ,  $P=0.133$ ).

All 679,050 SNPs from the discovery analysis passed quality control in the validation sample and following LD based pruning 155,019 SNPs remained to calculate the polygenic environmental sensitivity score in this sample. We generated 8 scores using increasingly liberal significance thresholds to select SNPs from the discovery sample ( $P < 0.001$ , 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5) which included 400; 3,161; 13,632; 25,384; 46,752; 66,205; 84,025 and 100,111 SNPs respectively.

Table 2 shows the results of a linear regression exploring the main effects of each polygenic environmental sensitivity score and the main effects of parenting on emotional problems. The polygenic score was not significantly associated with emotional symptom score, and findings were consistent across all significance thresholds. There was a significant main effect of parenting on emotional problems in the expected direction, with more negative parenting associated with increased emotional symptom scores. In order to

investigate whether the polygenic environmental sensitivity score moderated the effects of parenting on emotional problems we added an interaction term to the above models. Significant interactions were identified for polygenic scores calculated using 5 of the 8 P-value thresholds. Interaction effects began to emerge when using a threshold of  $P < 0.1$  in the discovery sample where they explained an additional 0.33% of the variance. The addition of further SNPs strengthened these effects, which were greatest for the polygenic scores based on a threshold of  $P < 0.5$  where the interaction term explained an additional 0.53% of the variance. The interaction from this model is illustrated in Figure 1 in which the polygenic score (based on a threshold of  $P < 0.5$ ) is divided into equal tertiles to represent low moderate and high scores, and parenting score is separated into equal tertiles to represent negative, moderate and positive parenting. Findings were in the expected direction. Specifically, for individuals with a low polygenic environmental sensitivity score, parenting had little effect on emotional problems. However, for those with a higher polygenic score, negative parenting was associated with an increased emotional symptom score, while positive parenting was associated decreased scores.

To explore these interaction effects further, we re-analysed the data considering the effects of positive and negative aspects of parenting separately (Tables S1-S2). Findings were consistent with those from the above analyses. Specifically, in individuals with a higher polygenic score, negative parenting was associated with increased emotional problems, while positive parenting was associated with decreased emotional symptom scores. However, in those with a polygenic score neither positive nor negative parenting had an effect on emotional problems. There was no evidence for gene-environment correlation. That is, there was no significant association between the polygenic environmental sensitivity score at any of the measured thresholds and our measures of parenting (Table S3). Finally, to test whether the same interaction effects were observed across raters, we reanalyzed the data, using child reported emotional problems and parent reported parenting. Findings were

similar to those from our initial analysis showing significant interaction effects, which emerged when using a threshold of  $P < 0.2$  (Table S4).

#### *Treatment response analyses*

We used a linear mixed model to identify predictors of response (change in the severity of the primary diagnosis). Initially we explored the effects of clinical and demographic factors and findings were similar to those reported for the full sample [29]. Specifically, individuals with Social Anxiety Disorder (SoAD) or Specific Phobia (SP) showed a significantly poorer response to treatment than those with Generalized Anxiety Disorder (GAD), ( $\beta = 0.43$ ,  $P < 0.001$  and  $\beta = 0.19$ ,  $P = 0.013$ ) respectively. However, treatment response did not differ according to any other factors including sex, age or treatment-type (all  $P$  values  $> 0.05$ ).

In total 277,893 SNPs from the discovery analysis were available in the treatment response sample and following LD based pruning 72,375 remained to calculate the polygenic environmental sensitivity score. We generated 8 scores using increasingly liberal significance thresholds to select SNPs from the discovery sample ( $P < 0.001$ , 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5) which included 159; 1,295; 5,905; 10,988; 20,423; 29,461; 37,668 and 45,371 SNPs respectively.

The polygenic score did not significantly predict overall response to treatment and results were consistent across the different thresholds used to calculate the score (See Table 3). However, the polygenic score did have treatment-specific effects on response.

Specifically, the score was positively associated with response to individual CBT and negatively associated with response to brief parent-led CBT. These effects only emerged when the polygenic score included SNPs reaching  $P < 0.05$  in the discovery sample. At this threshold, the polygenic score explained 1.55% of the variance of response to individual CBT and 4.80% of the variance of response to brief parent-led CBT. While the addition of further

SNPs improved the P value of these associations for brief parent-led CBT, they did not substantially improve the variance explained.

We further explored the treatment specific effects of the polygenic score on outcome by testing for treatment type by polygenic score interactions. These analyses showed that the polygenic score (based on SNPs reaching  $P < 0.05$  in the discovery sample) significantly moderated the effect of each treatment type on outcome (individual vs. group CBT x polygenic score interaction:  $\beta = -0.13$ , 95%CI =  $-0.24$  to  $-0.02$ ,  $P = 0.02$ ); individual vs. brief parent-led CBT x polygenic score interaction:  $\beta = -0.30$ , 95%CI =  $-0.42$  to  $-0.17$ ,  $P = 3.1 \times 10^{-6}$ ; group CBT vs. brief parent-led CBT x polygenic score interaction:  $\beta = -0.15$ , 95%CI =  $-0.26$  to  $-0.04$ ,  $P = 0.007$ ).

For those with a low polygenic environmental sensitivity score, treatment type had little effect on outcome. However, those with a high polygenic score responded well to individual CBT, moderately to group CBT and poorly to brief parent-led CBT. These effects are illustrated in Figure 2, which shows the mean change in anxiety severity score between baseline and the post-treatment time-point by tertiles of low, moderate and high polygenic score (using the threshold of  $P < 0.05$ ). Figure 3, shows the percentage of individuals in remission at the post-treatment time point by tertiles of low, moderate and polygenic score. 70.9% of individuals at the upper tertile of the score treated with individual CBT were in remission at the post-treatment time point. However, remission rates for those in the upper tertile of the score were only 55.5% in those treated with group CBT and 41.6% in those treated with brief parent-led CBT

As all analyses included baseline anxiety severity, diagnosis, age and gender as covariates, these factors are unlikely to confound the relationship between the polygenic score and treatment response. However, we previously showed that comorbid externalising and internalising disorders (measured in a subset of the sample ( $n = 935$ ), as well as parental psychopathology (measured in a smaller subset,  $n = 816$ ) were associated with treatment

response. Linear and logistic regressions showed that the polygenic score was not significantly related to parental psychopathology or the presence of comorbid externalising disorders, but was significantly associated with a lower likelihood of comorbid internalising disorders at the majority of the polygenic score thresholds tested (See Table S5)

In order to exclude the possibility that the presence of comorbid internalising disorders confounded the relationship between the polygenic scores and treatment response we therefore re-ran analyses controlling for this variable on the subsample in which they were available. Findings were similar to those from the main analysis (see Table S6).

Finally, the non-random allocation of treatments meant that individuals in each treatment group differed on several clinical and demographic factors including baseline severity, diagnosis, age, parental psychopathology and comorbid externalising and internalising disorders (See Table S7). To ensure that interactions between the polygenic score and treatment type on outcome were not biased by these differences we used propensity score matching to restrict analyses to individuals across treatment types who were matched for baseline severity, age, diagnosis, comorbid externalising and internalising disorders and parental psychopathology (see supplementary methods). Using this reduced sample, interaction effects were of a similar magnitude to those reported for the main analyses (individual vs. brief parent-led CBT x polygenic score interaction:  $\beta=-0.28$ , 95%CI=-0.46--0.09,  $P=0.003$ ; group CBT vs. brief parent-led CBT x polygenic score interaction:  $\beta=-0.15$ , 95%CI=-0.29--0.01,  $P=0.041$ ) suggesting that they were not the result of measured differences between treatment types at baseline.

## Discussion

The differential susceptibility hypothesis suggests that the same genetic variants moderate the effects of both positive and negative environments on mental health. While several candidate gene studies support this hypothesis, this was the first to find evidence for differential susceptibility using a genome-wide approach. We used an MZ differences design to detect variants that increase the effects of the environment on the development of emotional problems. Consistent with the differential susceptibility hypothesis, we found that a polygenic environmental sensitivity score based on these findings moderated the effects of both positive and negative parenting on emotional problems in a further sample of children. The same polygenic score also moderated response to different psychological treatments in children with anxiety disorders.

### *Main findings*

We examined within-pair variability in emotional symptoms in monozygotic twins to detect genetic variants associated with increased sensitivity to the environment. None of our findings reached genome-wide significance. Nevertheless, suggestively significant findings were identified in a region containing *UHMK1*, the gene, which encodes the brain-enriched protein kinase KIS. Animal models suggest that *UHMK1* is highly expressed in the brain, and knockdown of this gene the development of cortical neurons in culture [37]. In line with our findings, *UHMK1* knockout mice display a distinct deficit in fear conditioning which is accompanied by a down regulation of genes implicated in the aetiology of anxiety and fear including multiple components of GABA<sub>A</sub> receptors[38].

Consistent with our hypothesis, a polygenic environmental sensitivity score based on the whole genome results significantly moderated the effects of parenting on emotional problems in an unrelated sample. In line with the differential susceptibility hypothesis, this interaction applied to both the positive and negative aspects of this environmental measure.

That is, for individuals with low environmental sensitivity, parenting had little effect on emotional problems. In contrast, for those with high environmental sensitivity, negative parenting was associated with increased emotional problems, while positive parenting was associated with decreased emotional symptom scores. While statistically significant, the effects of the environmental sensitivity by environment interactions were very small, explaining at most an additional 0.53% of the variance in outcome. Nevertheless, these findings are comparable to those reported for the main effects of polygenic scores and polygenic score by environment interactions in a previous study of major depression [20]. The variance explained was also larger than that reported for the main effects of polygenic scores on depression symptoms in a population sample [39].

While the polygenic environmental sensitivity score did not predict overall response to treatment it did significantly predict differential response to individual CBT, group CBT and brief parent-led CBT. Importantly, these findings were not confounded by measured baseline clinical or demographic characteristics or biased by measured differences between treatment groups. The effects of environmental sensitivity appeared to increase linearly with the intensity of the treatment delivered, such that those with the highest environmental sensitivity responded best to individual CBT, moderately to group CBT and poorly to brief parent-led CBT. In contrast, those with a low environmental sensitivity responded equally well to each treatment type. The variance explained by the polygenic score was modest (1.62% in those treated with individual CBT and 5.77% in those treated with parent-led guided self help) but are nevertheless comparable to previous studies of treatment response using a polygenic approach [40].

Previous studies have created cumulative scores of environmental sensitivity based on small sets of hypothesised differential susceptibility alleles and tested them as moderators of the environment [41] and predictors of treatment response [3]. However, this was the first to use an MZ differences or genome-wide approach to detect and weight alleles

according to their effect on environmental sensitivity. Nevertheless, our findings are consistent with those of multiple candidate gene studies in which specific variants have been shown to enhance the effects of both negative and positive parenting on internalising and externalising phenotypes [14], response to CBT [13], and a range of interventions for internalizing and externalizing behaviour [15].

Studies are yet to examine the effects of sensitivity to the environment on response to CBT of varying intensity. It has been argued that individuals with a low sensitivity to the environment may require a more intensive type of treatment to achieve the same results as those who are highly sensitive. Our findings do not support this hypothesis. Outcomes for those with a low sensitivity to the environment were the same, regardless of the intensity of the treatment provided. In those with a high environmental sensitivity, the intensity of treatment was positively correlated with outcome such that they derived the most benefit from the most intensive forms of treatment. This finding is in line with the differential susceptibility hypothesis, which suggests that increasing exposure to an environment (positive or negative) has a greater effect on environmentally sensitive than environmentally insensitive individuals.

A more complete explanation may be that individuals with increased genetic sensitivity to the environment develop more of the cognitive biases underlying anxiety disorders (such as a bias towards threat [42]) and therefore require more intensive treatments to overcome these aberrant cognitions. A prospective longitudinal study with data both at the onset of illness and throughout treatment would be necessary to directly test this hypothesis. Such a design would allow for the investigation of aetiological factors [43], as well as the effects of the course of illness and disease progression [44].

### *Implications*

If replicated our findings may have several important implications for understanding the aetiology of emotional problems and treatment response.

We found our polygenic environmental sensitivity score was only a significant moderator of parenting or treatment response when it included variants reaching thresholds of  $P < 0.1$  and  $P < 0.05$  respectively in the discovery dataset. This suggests that sensitivity to the environment, rather than being the result of the effects of a handful of candidate genes, is a polygenic trait, which is due to the aggregate effects of tens of thousands of variants of small effect. These polygenic effects may explain why previous studies of gene-environment interaction, which focus on a single candidate gene, often fail to replicate.

A previous polygenic score study, which assumed a diathesis-stress model, suggested that around 0.5% of the variance in the liability of major depression is accounted for by gene-environment interaction [20]. We report that a similar amount of variance in the aetiology of emotional problems is accounted for by variants operating in a manner consistent with differential susceptibility. A significant role of such variants may explain why despite moderate estimates of heritability, attempts to identify the genetic variants responsible for child anxiety and depression have so far been unsuccessful [28]. They may also explain why SNP-level heritability estimates are also considerably lower than those predicted from twins, even within the same samples [45].

Polygenic predictors of environmental sensitivity may allow for a more accurate identification of those who are at risk of developing disorders in the face of adversity, but also who would be the most likely to benefit from which treatments. Previous findings from the current sample suggested that treatment type had little overall effect on outcome [29]. Indeed, a recent meta-analysis reported that individual CBT or group CBT were as effective as lower intensity self-help approaches [2]. However, our results suggest that the efficacy of different treatment types differs markedly according to environmental sensitivity. These effects are potentially clinically meaningful, with remission rates at the upper tertile of the

polygenic score of 70.9%, 55.1% and 40.6% for Individual CBT, group CBT and brief parent-led CBT respectively. If replicated, our findings suggest that for those with a relatively low genetic sensitivity to the environment more cost-effective, lower intensity approaches are equally as effective as face-to-face treatment. More importantly, they also suggest response rates may be substantially improved by targeting those with an increased genetic sensitivity to the environment with the most intensive psychological therapies.

### *Strengths and limitations*

It has been noted that there are two principal challenges facing gene-environment interaction research: the necessity to develop methods which include the whole genome and those which include and reliably measure the whole 'environome' of relevant environments [18]. By assessing the aggregate effects of genetic variants from across the genome on unmeasured non-shared environmental effects, the current study simultaneously addresses both of these challenges. Nevertheless, our findings should be interpreted in the light of several important limitations.

First, while our discovery analyses used a large, well-characterized sample of MZ twins, it was only adequately powered (80%) to detect individual variants with moderate effects on environmental sensitivity at genome-wide significance (explaining more than 1% of the variance of the variance) [21]. Although our polygenic approach did not rely solely on genome-wide significant findings, the discovery, validation and treatment samples were smaller than recommended for polygenic score analyses, particularly when testing treatment specific effects [19]. Our findings should therefore be considered exploratory, pending replication in further, larger samples.

Second, we aimed to identify genetic variants that moderated the effects of the non-shared environment on emotional problems. We chose to validate these findings by exploring the interaction between the score and child reported parenting as this is one of

the most robust environmental predictors of child anxiety [46] and has been shown to be moderated by genetic factors in manner consistent with differential susceptibility [7]. We identified similar interactions in cross rater analyses (using parent rated parenting and child rated emotional problems) and the same variants also moderated the arguably more objectively measured environment of psychological treatment. However, it remains unknown whether these findings extend to more objectively measured environments such as observed parenting.

Finally, our treatment response sample included children with anxiety disorders receiving a psychological treatment as part of a trial or treatment as usual in one of multiple studies [29]. The subsequent non-random allocations of treatments meant that treatment type was associated with several clinical and demographic characteristics at baseline. While additional analyses using propensity score matching allowed us to conclude that our findings were not biased by measured differences between treatment groups, we cannot exclude the possibility that individuals differed by unmeasured factors. Replication of our findings in a randomized trial comparing low and high intensity CBT is therefore necessary to fully exclude the effects of confounding by indication.

### *Conclusion*

The limited power provided by the cohorts used in each stage of our study means that our results should be considered exploratory until successfully replicated in larger samples. Nevertheless, if replicated, our findings suggest that responsivity to the environment is the result of multiple genetic variants of small effect, rather than a few select candidate genes. We show that these variants moderated the effects of parenting on the development of emotional problems in children. The same variants also predicted differential response to psychological treatments, such that those with the greatest sensitivity to the environment appeared to benefit the most from more intensive types of treatment. In line with previous

polygenic score studies, the gene-by-environment effects we identified explained a very small proportion of the variance (0.53%). However, the variance explained by gene-by-treatment effects was larger (1-5%). The potential clinical utility of these findings warrants further investigation of these effects in patients receiving low and high intensity CBT.

## References

- 1 Rapee RM, Schniering CA, Hudson JL: Anxiety disorders during childhood and adolescence: Origins and treatment. *Annual review of clinical psychology* 2009;5:311-341.
- 2 James AC, James G, Cowdrey FA, Soler A, Choke A: Cognitive behavioural therapy for anxiety disorders in children and adolescents. *Cochrane Database of Systematic Reviews* 2013;6:CD004690.
- 3 Hudson JL, Lester KJ, Lewis CM, Tropeano M, Creswell C, Collier DA, Cooper P, Lyneham HJ, Morris T, Rapee RM, Roberts S, Donald JA, Eley TC: Predicting outcomes following cognitive behaviour therapy in child anxiety disorders: The influence of genetic, demographic and clinical information. *Journal of Child Psychology and Psychiatry* 2013;54:1086-1094.
- 4 Morath J, Moreno-Villanueva M, Hamuni G, Kolassa S, Ruf-Leuschner M, Schauer M, Elbert T, Burkle A, Kolassa IT: Effects of psychotherapy on DNA strand break accumulation originating from traumatic stress. *Psychother Psychosom* 2014;83:289-297.
- 5 Belsky J, Pluess M: Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychol Bull* 2009;135:885-908.
- 6 Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, McClay J, Mill J, Martin J, Braithwaite A, Poulton R: Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science* 2003;301:386-389.
- 7 Hankin BL, Naderhof E, Oppenheimer CW, Jenness J, Young JF, Abela JR, Smolen A, Ormel J, Oldehinkel AJ: Differential susceptibility in youth: Evidence that 5-HTTLPR x positive parenting is associated with positive affect 'for better and worse'. *Translational psychiatry* 2011;1:e44.
- 8 Pluess M, Belsky J, Way BM, Taylor SE: 5-HTTLPR moderates effects of current life events on neuroticism: Differential susceptibility to environmental influences. *Prog Neuropsychopharmacol Biol Psychiatry* 2010;34:1070-1074.
- 9 Taylor SE, Way BM, Welch WT, Hilmert CJ, Lehman BJ, Eisenberger NI: Early family environment, current adversity, the serotonin transporter promoter polymorphism, and depressive symptomatology. *Biol Psychiatry* 2006;60:671-676.
- 10 Kohen R, Cain KC, Buzaitis A, Johnson V, Becker KJ, Teri L, Tirschwell DL, Veith RC, Mitchell PH: Response to psychosocial treatment in poststroke depression is associated with serotonin transporter polymorphisms. *Stroke* 2011;42:2068-2070.
- 11 Drury SS, Gleason MM, Theall KP, Smyke AT, Nelson CA, Fox NA, Zeanah CH: Genetic sensitivity to the caregiving context: The influence of 5-HTTLPR and BDNF Val66Met on indiscriminate social behavior. *Physiology & behavior* 2012;106:728-735.
- 12 Brett ZH, Humphreys KL, Smyke AT, Gleason MM, Nelson CA, Zeanah CH, Fox NA, Drury SS: Serotonin transporter linked polymorphic region (5-HTTLPR) genotype moderates the longitudinal impact of early caregiving on externalizing behavior. *Dev Psychopathol* 2015;27:7-18.
- 13 Eley TC, Hudson JL, Creswell C, Tropeano M, Lester KJ, Cooper P, Farmer A, Lewis CM, Lyneham HJ, Rapee RM, Uher R, Zavos HM, Collier DA: Therapygenetics: The 5-HTTLPR and response to psychological therapy. *Mol Psychiatry* 2012;17:236-237.

- 14 Belsky J, Pluess M: Beyond risk, resilience, and dysregulation: Phenotypic plasticity and human development. *Dev Psychopathol* 2013;25:1243-1261.
- 15 van Ijzendoorn MH, Bakermans-Kranenburg MJ: Genetic differential susceptibility on trial: Meta-analytic support from randomized controlled experiments. *Dev Psychopathol* 2015;27:151-162.
- 16 Fergusson DM, Horwood LJ, Miller AL, Kennedy MA: Life stress, 5-HTTLPR and mental disorder: Findings from a 30-year longitudinal study. *Br J Psychiatry* 2011;198:129-135.
- 17 Roberts S, Lester K, Keers R, Coleman J, Breen G, Wong C, Xu X, Arendt K, Blatter-Meunier J, Bögels S, Cooper P, Creswell C, Heiervang E, Herren C, Hogendoorn S, Hudson J, Krause K, Lyneham H, McKinnon A, Morris T, Nauta M, Rapee R, Rey Y, Schneider S, Schneider S, Silverman W, Smith P, Thastum M, Thirlwall K, Waite P, Wergeland GJ, Eley T, Roberts S, Lester K, Keers R, Coleman J, Breen G, Wong C, Xu X, Arendt K, Blatter-Meunier J, Bögels S, Cooper P, Creswell C, Heiervang E, Herren C, Hogendoorn S, Hudson J, Krause K, Lyneham H, McKinnon A, Morris T, Nauta M, Rapee R, Rey Y, Schneider S, Schneider S, Silverman W, Smith P, Thastum M, Thirlwall K, Waite P, Wergeland GJ, Eley T: Non-replication of the association between 5HTTLPR and response to psychological therapy for child anxiety disorders. *British Journal of Psychiatry* 2015
- 18 Uher R: Gene-environment interactions in severe mental illness. *Front Psychiatry* 2014;5:48.
- 19 Dudbridge F: Power and predictive accuracy of polygenic risk scores. *PLoS genetics* 2013;9:e1003348.
- 20 Peyrot WJ, Milaneschi Y, Abdellaoui A, Sullivan PF, Hottenga JJ, Boomsma DI, Penninx BW: Effect of polygenic risk scores on depression in childhood trauma. *Br J Psychiatry* 2014;205:113-119.
- 21 Visscher PM, Posthuma D: Statistical power to detect genetic loci affecting environmental sensitivity. *Behav Genet* 2010;40:728-733.
- 22 Surakka I, Whitfield JB, Perola M, Visscher PM, Montgomery GW, Falchi M, Willemsen G, de Geus EJ, Magnusson PK, Christensen K, Sorensen TI, Pietilainen KH, Rantanen T, Silander K, Widen E, Muiilu J, Rahman I, Liljedahl U, Syvanen AC, Palotie A, Kaprio J, Kyvik KO, Pedersen NL, Boomsma DI, Spector T, Martin NG, Ripatti S, Peltonen L, Genom EP: A genome-wide association study of monozygotic twin-pairs suggests a locus related to variability of serum high-density lipoprotein cholesterol. *Twin Res Hum Genet* 2012;15:691-699.
- 23 Haworth CM, Davis OS, 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* 2013;16:117-125.
- 24 Goodman R: The strengths and difficulties questionnaire: A research note. *Journal of child psychology and psychiatry, and allied disciplines* 1997;38:581-586.
- 25 Deater-Deckard K: Parenting and child behavioral adjustment in early childhood: A quantitative genetic approach to studying family processes. *Child Dev* 2000;71:468-484.
- 26 Deater-Deckard K, Dodge KA, Bates JE, Pettit GS: Multiple risk factors in the development of externalizing behavior problems: Group and individual differences. *Dev Psychopathol* 1998;10:469-493.
- 27 Oliver BR, Trzaskowski M, Plomin R: Genetics of parenting: The power of the dark side. *Dev Psychol* 2014;50:1233-1240.

- 28 Trzaskowski M, Eley TC, Davis OS, Doherty SJ, Hanscombe KB, Meaburn EL, Haworth CM, Price T, Plomin R: First genome-wide association study on anxiety-related behaviours in childhood. *PloS one* 2013;8:e58676.
- 29 Hudson J, Keers R, Roberts S, Coleman JRI, Breen G, Arendt K, Bogels S, Cooper P, Creswell C, Hartman CA, Heiervang E, Hotzel K, In-Albon T, Lavallee K, Lyneham HJ, Marin CE, McKinnon A, Meiser-Stedman R, Morris T, Nauta M, Rapee R, Schneider S, Schneider S, Silverman W, Thastum M, Thirlwall K, Waite P, Wergeland GJ, Lester K, Eley T, Hudson J, Keers R, Roberts S, Coleman JRI, Breen G, Arendt K, Bogels S, Cooper P, Creswell C, Hartman CA, Heiervang E, Hotzel K, In-Albon T, Lavallee K, Lyneham HJ, Marin CE, McKinnon A, Meiser-Stedman R, Morris T, Nauta M, Rapee R, Schneider S, Schneider S, Silverman W, Thastum M, Thirlwall K, Waite P, Wergeland GJ, Lester K, Eley T: Clinical predictors of response to cognitive-behavioral therapy in pediatric anxiety disorders. *The Genes for Treatment (GxT) Study* 2015
- 30 Silverman WK, Albano AM: *The anxiety disorders interview schedule for children-iv (child and parent versions)*. New York, Oxford University Press, 1996.
- 31 Schneider S, Unnewehr S, Margraf J: *Kinder-dips fur dsm-iv-tr. Diagnostisches interview bei psychischen strungen im Kindes- und jugendalter*. Heidelberg 2009.
- 32 Creswell C, Apetroaia A, Murray L, Cooper P: Cognitive, affective, and behavioral characteristics of mothers with anxiety disorders in the context of child anxiety disorder. *Journal of Abnormal Psychology* 2012;122:26-38.
- 33 Hudson JL, Newall C, Rapee RM, Lyneham HJ, Schniering CC, Wuthrich VM, Schneider S, Seeley-Wait E, Edwards S, Gar NS: The impact of brief parental anxiety management on child anxiety treatment outcomes: A controlled trial. *J Clin Child Adolesc Psychol* 2014;43:370-380.
- 34 Lyneham HJ, Abbott MJ, Rapee RM: Interrater reliability of the anxiety disorders interview schedule for dsm-iv: Child and parent version. *Journal of the American Academy of Child and Adolescent Psychiatry* 2007;46:731-736.
- 35 Lovibond PF, Lovibond SH: The structure of negative emotional states: Comparison of the depression anxiety stress scales (dass) with the beck depression and anxiety inventories. *Behav Res Ther* 1995;33:335-343.
- 36 Coleman JRI, Lester K, Roberts S, Curtis C, Arendt K, Bogels S, Cooper P, Creswell C, Dalgleish T, Hartman CA, Heiervang E, Hotzel K, Hudson J, In-Albon T, Lavallee K, Lyneham HJ, Marin CE, Meiser-Stedman R, Morris T, Nauta M, Rapee R, Schneider S, Schneider S, Silverman W, Thastum M, Thirlwall K, Waite P, Wergeland GJ, Breen G, Eley T: Genome-wide association study of response to cognitive behavioral therapy in children with anxiety disorders. Under review 2015
- 37 Cambray S, Pedraza N, Rafel M, Gari E, Aldea M, Gallego C: Protein kinase kis localizes to rna granules and enhances local translation. *Mol Cell Biol* 2009;29:726-735.
- 38 Manceau V, Kremmer E, Nabel EG, Maucuer A: The protein kinase kis impacts gene expression during development and fear conditioning in adult mice. *PloS one* 2012;7:e43946.
- 39 Chang SC, Glymour MM, Walter S, Liang L, Koenen KC, Tchetgen EJ, Cornelis MC, Kawachi I, Rimm E, Kubzansky LD: Genome-wide polygenic scoring for a 14-year long-term average depression phenotype. *Brain and behavior* 2014;4:298-311.
- 40 Investigators G, Investigators M, Investigators SD: Common genetic variation and antidepressant efficacy in major depressive disorder: A meta-analysis of three genome-wide pharmacogenetic studies. *The American journal of psychiatry* 2013;170:207-217.

- 41 Brett ZH, Sheridan M, Humphreys K, Smyke A, Gleason MM, Fox N, Zeanah C, Nelson C, Drury S: A neurogenetics approach to defining differential susceptibility to institutional care. *International journal of behavioral development* 2015;39:150-160.
- 42 Pergamin-Hight L, Bakermans-Kranenburg MJ, van Ijzendoorn MH, Bar-Haim Y: Variations in the promoter region of the serotonin transporter gene and biased attention for emotional information: A meta-analysis. *Biol Psychiatry* 2012;71:373-379.
- 43 Keers R, Uher R: Gene–environment interaction in major depression and antidepressant treatment response. *Curr Psychiatry Rep* 2011;DOI 10.1007/s11920-011-0251-x
- 44 Fava GA, Guidi J, Grandi S, Hasler G: The missing link between clinical states and biomarkers in mental disorders. *Psychother Psychosom* 2014;83:136-141.
- 45 Trzaskowski M, Dale PS, Plomin R: No genetic influence for childhood behavior problems from DNA analysis. *J Am Acad Child Adolesc Psychiatry* 2013;52:1048-1056 e1043.
- 46 McLeod BD, Wood JJ, Weisz JR: Examining the association between parenting and childhood anxiety: A meta-analysis. *Clin Psychol Rev* 2007;27:155-172.

Table 1. Associations with intra-pair differences in emotional problems in monozygotic twins reaching suggestive significance ( $P < 1 \times 10^{-5}$ )

| Chromosome | SNP ID     | Position  | Allele | Beta    | P value               | Nearest Gene |
|------------|------------|-----------|--------|---------|-----------------------|--------------|
| 1          | rs12131428 | 162426451 | C      | 0.3885  | $2.10 \times 10^{-7}$ | UHMK1        |
| 22         | rs5748871  | 17603477  | A      | -0.1915 | $1.63 \times 10^{-6}$ | CECR6        |
| 19         | rs7339483  | 24462409  | G      | 0.3683  | $6.20 \times 10^{-6}$ | ZNF254       |
| 5          | rs3864261  | 72358254  | A      | 0.2662  | $7.33 \times 10^{-6}$ | FCHO2        |
| 8          | rs10875469 | 142333425 | T      | -0.2144 | $9.29 \times 10^{-6}$ | GPR20        |
| 5          | rs1392412  | 72362289  | G      | 0.2631  | $9.41 \times 10^{-6}$ | FCHO2        |

Table 2. Validation analyses: Linear regression examining the main effects of polygenic environmental sensitivity score and parenting and their interaction on emotional problems

| P value threshold | Main effects of polygenic environmental sensitivity score <sup>a</sup> |            |       | Main effects of parenting <sup>a</sup> |           |                          | Polygenic environmental sensitivity score by parenting interaction <sup>b</sup> |            |       |                  |
|-------------------|--|------------|-------|--|-----------|--------------------------|---|------------|-------|------------------|
|                   | $\beta$  | 95% CI     | P     | $\beta$                                | 95% CI    | P                        | $\beta$   | 95% CI     | P     | % R <sup>2</sup> |
| 0.001             | -0.01  | -0.04-0.05 | 0.869 | 0.21                                   | 0.17-0.25 | 4.43 x 10 <sup>-22</sup> | -0.04   | -0.08-0.01 | 0.107 | 0.05             |
| 0.01              | -0.01  | -0.05-0.03 | 0.718 | 0.21                                   | 0.17-0.25 | 4.65 x 10 <sup>-22</sup> | 0.01  | -0.05-0.04 | 0.848 | 0.01             |
| 0.05              | 0.01   | -0.03-0.05 | 0.615 | 0.21                                   | 0.17-0.25 | 4.54 x 10 <sup>-22</sup> | 0.04  | -0.01-0.08 | 0.085 | 0.22             |
| 0.1               | 0.02   | -0.03-0.06 | 0.470 | 0.21                                   | 0.17-0.25 | 4.38 x 10 <sup>-22</sup> | 0.05  | 0.00-0.09  | 0.035 | 0.33             |
| 0.2               | 0.01   | -0.03-0.05 | 0.727 | 0.21                                   | 0.17-0.25 | 4.58 x 10 <sup>-22</sup> | 0.06  | 0.01-0.10  | 0.011 | 0.47             |
| 0.3               | -0.01  | -0.05-0.04 | 0.787 | 0.21                                   | 0.17-0.25 | 4.60 x 10 <sup>-22</sup> | 0.06  | 0.01-0.10  | 0.012 | 0.46             |
| 0.4               | -0.01  | -0.05-0.03 | 0.636 | 0.21                                   | 0.17-0.25 | 4.52 x 10 <sup>-22</sup> | 0.06  | 0.02-0.10  | 0.008 | 0.49             |
| 0.5               | -0.01  | -0.05-0.03 | 0.640 | 0.21                                   | 0.17-0.25 | 4.49 x 10 <sup>-22</sup> | 0.06  | 0.02-0.10  | 0.005 | 0.53             |

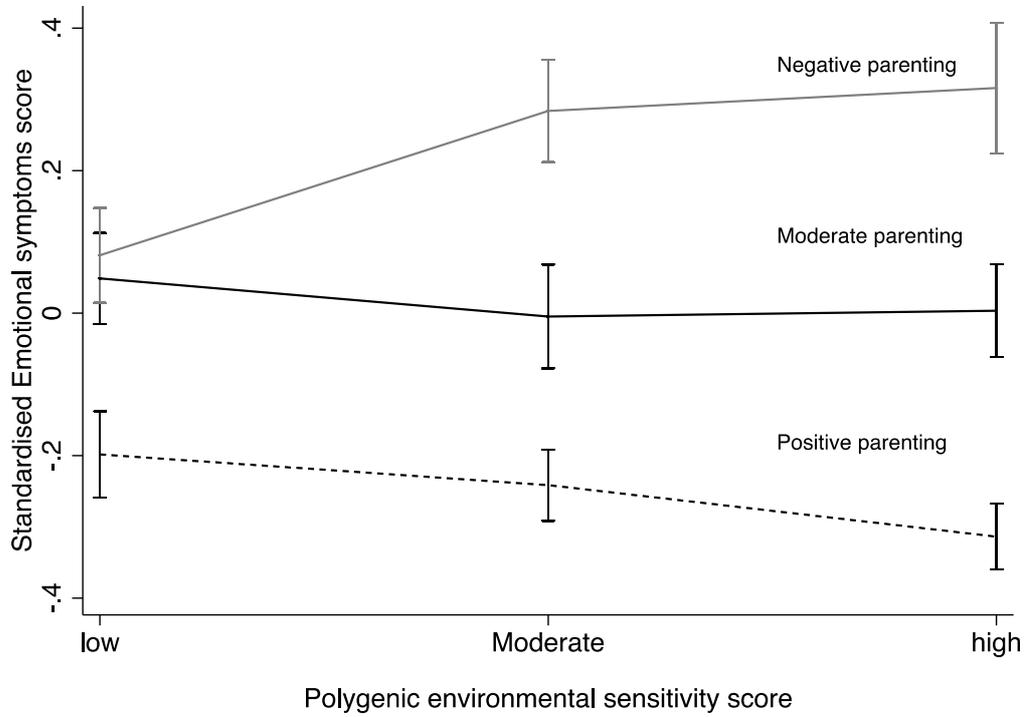
<sup>a</sup> models included the main effects of polygenic environmental sensitivity score and child-reported parenting on age and sex regressed combined child/adult rated emotional symptom score. <sup>b</sup> models included the main effects of polygenic environmental sensitivity score and child-reported parenting and their interaction on age and sex regressed combined child/adult rated emotional symptom score. To account for possible effects of population stratification all models also included the first 10 principal components previously derived from genome-wide analyses of the TEDS data.

Table 3. Treatment response analyses: Linear mixed model examining the effect of the polygenic environmental sensitivity score on treatment response (change in the severity of the primary anxiety disorder).

| P value threshold | Overall response |            |       |                  | Response to individual CBT |             |       |                  | Response to group based CBT |            |       |                  | Response to brief parent-led CBT |            |                         |                  |
|-------------------|------------------|------------|-------|------------------|----------------------------|-------------|-------|------------------|-----------------------------|------------|-------|------------------|----------------------------------|------------|-------------------------|------------------|
|                   | $\beta$          | 95% CI     | P     | % R <sup>2</sup> | $\beta$                    | 95% CI      | P     | % R <sup>2</sup> | $\beta$                     | 95% CI     | P     | % R <sup>2</sup> | $\beta$                          | 95% CI     | P                       | % R <sup>2</sup> |
| 0.001             | 0.01             | -0.03-0.05 | 0.699 | 0.03             | -0.01                      | -0.09-0.08  | 0.865 | 0.00             | 0.02                        | -0.04-0.08 | 0.482 | 0.08             | 0.04                             | -0.06-0.15 | 0.408                   | 0.45             |
| 0.01              | 0.02             | -0.02-0.06 | 0.357 | 0.04             | -0.08                      | -0.17-0.01  | 0.077 | 0.72             | 0.04                        | -0.02-0.10 | 0.182 | 0.23             | 0.07                             | -0.02-0.16 | 0.151                   | 0.80             |
| 0.05              | 0.03             | -0.02-0.07 | 0.267 | 0.05             | -0.12                      | -0.21--0.03 | 0.009 | 1.62             | 0.02                        | -0.04-0.08 | 0.456 | 0.10             | 0.18                             | 0.09-0.27  | 6.97 x 10 <sup>-5</sup> | 4.80             |
| 0.1               | 0.02             | -0.02-0.07 | 0.339 | 0.03             | -0.11                      | -0.19--0.02 | 0.014 | 1.50             | 0.01                        | -0.05-0.07 | 0.720 | 0.04             | 0.20                             | 0.11-0.29  | 1.92 x 10 <sup>-5</sup> | 5.21             |
| 0.2               | 0.02             | -0.02-0.07 | 0.277 | 0.05             | -0.09                      | -0.18--0.01 | 0.033 | 1.11             | 0.01                        | -0.05-0.07 | 0.841 | 0.03             | 0.21                             | 0.12-0.30  | 6.14 x 10 <sup>-5</sup> | 5.77             |
| 0.3               | 0.02             | -0.03-0.06 | 0.420 | 0.02             | -0.10                      | -0.18--0.01 | 0.022 | 1.23             | 0.01                        | -0.06-0.06 | 0.947 | 0.00             | 0.20                             | 0.11-0.29  | 1.99 x 10 <sup>-5</sup> | 5.20             |
| 0.4               | 0.02             | -0.03-0.06 | 0.485 | 0.01             | -0.10                      | -0.19--0.02 | 0.017 | 1.36             | 0.01                        | -0.06-0.06 | 0.971 | 0.01             | 0.19                             | 0.10-0.28  | 5.78 x 10 <sup>-5</sup> | 4.81             |
| 0.5               | 0.02             | -0.03-0.06 | 0.471 | 0.01             | -0.11                      | -0.19--0.02 | 0.014 | 1.44             | 0.01                        | -0.06-0.06 | 0.918 | 0.00             | 0.19                             | 0.10-0.29  | 3.47 x 10 <sup>-5</sup> | 5.14             |

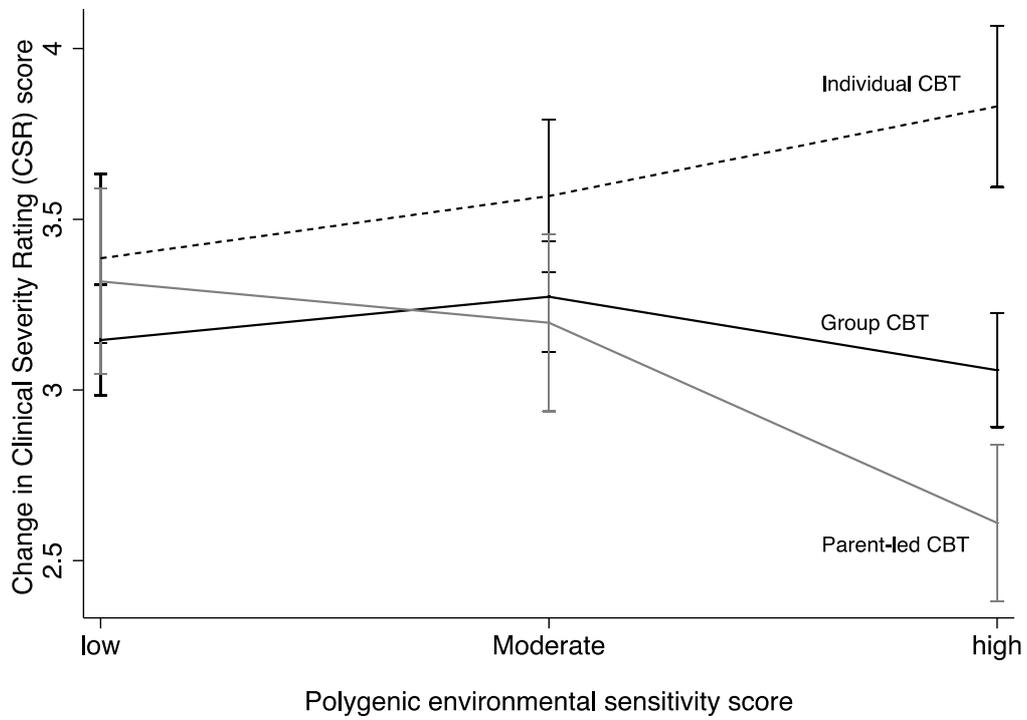
Note. To account for data collected longitudinally, all models included the random effects of participant and the linear and quadratic effects of time. All models also included sex, age (centred), primary diagnosis (Generalised Anxiety Disorder (GAD) Social Anxiety Disorder (SoAD) Specific Phobia (SP) Separation Anxiety Disorder (SAD) or "Other anxiety" disorder) and treatment type (individual based CBT group based CBT or brief parent-led CBT). All models included the random effects of trial. Regression weights ( $\beta$ ) significantly greater than zero indicate that this variable is associated with a poorer response following treatment.

**Figure 1.** Effects of the polygenic environmental sensitivity score, parenting and their interaction on emotional problems



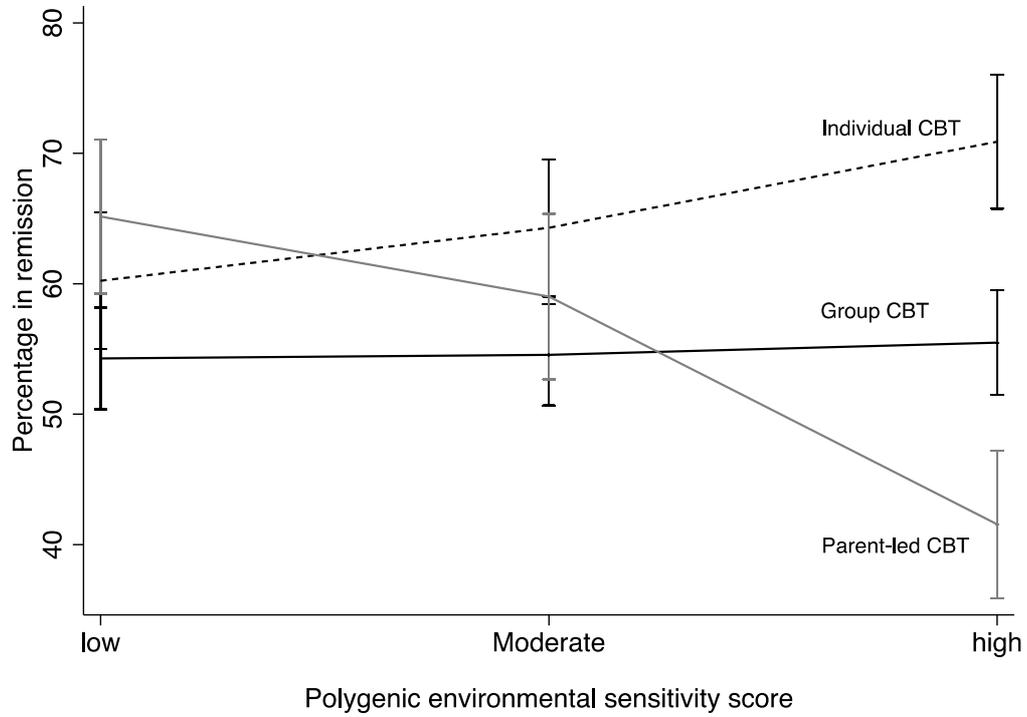
Mean standardised emotional symptom score by tertiles of parenting (representing negative, moderate and positive parenting) and tertiles of the polygenic environmental sensitivity score (low, moderate and high, threshold =  $P < 0.5$ ). Error bars represent 1 standard error.

**Figure 2.** Effects of the polygenic environmental sensitivity score on change in clinical severity rating score from baseline to the post-treatment time point



Mean change in clinical severity rating from baseline to post-treatment for individuals treated with individual CBT, group CBT and brief parent-led CBT by tertiles of the polygenic environmental sensitivity score (low, moderate and high). Error bars represent 1 standard error.

**Figure 3** Effects of the polygenic environmental sensitivity score on the percentage of individuals in remission at the post-treatment time point



Percentage of individuals in remission at the post treatment time point for individuals treated with individual CBT, group CBT and brief parent-led CBT by tertiles of the polygenic environmental sensitivity score (low moderate and high). Error bars represent 1 standard error.

## Supplementary methods

### *Propensity score matching*

Propensity score matching was used to match individuals across treatment types based on sex, age, baseline severity, primary diagnosis, parental psychopathology, comorbid internalizing and externalizing disorders. This approach allowed us to create a propensity score of each individual being treated with individual CBT vs group CBT, individual CBT vs brief parent-led CBT and group CBT vs Brief parent-led CBT. For each comparison individuals in one treatment group were matched, using the propensity score, to an individual receiving the alternative treatment using nearest neighbour matching in the psmatch2 package for STATA[1]. A threshold difference in propensity score of 0.02 within each pair was imposed to restrict analysis to the best-matched pairs[2]. This approach resulted in matched pairs for individual CBT vs group CBT (n=136, mean difference in propensity score =0.002, se=0.0003) individual CBT vs brief parent-led CBT (n=105, 0.006, se=0.0006) and group CBT vs Brief parent-led CBT (n=169; mean difference in propensity scores= 0.012, se=0.0004). Table S10 shows differences in sex, age, baseline severity, primary diagnosis, parental psychopathology, comorbid internalizing and externalizing disorders before and after matching.

Table S1. Validation analyses: Linear regression examining the main effects of the polygenic environmental sensitivity score, positive aspects of parenting and their interaction on emotional problems

| P value threshold | Main effects of the polygenic environmental sensitivity score <sup>a</sup> |            |       | Main effects of positive aspects of parenting <sup>a</sup> |             |                       | Polygenic environmental sensitivity score by positive aspects of parenting interaction <sup>b</sup> |             |       |                  |
|-------------------|--|------------|-------|--|-------------|-----------------------|---|-------------|-------|------------------|
|                   | $\beta$  | 95% CI     | P     | $\beta$  | 95% CI      | P                     | $\beta$   | 95% CI      | P     | % R <sup>2</sup> |
| 0.001             | 0.01   | -0.04-0.05 | 0.838 | -0.09  | -0.13--0.04 | $8.99 \times 10^{-5}$ | 0.05  | 0.01-0.10   | 0.020 | 0.38             |
| 0.01              | -0.01  | -0.05-0.04 | 0.755 | -0.09  | -0.13--0.04 | $8.98 \times 10^{-5}$ | 0.03  | -0.01-0.08  | 0.175 | 0.13             |
| 0.05              | 0.01   | -0.03-0.05 | 0.645 | -0.09  | -0.13--0.04 | $8.69 \times 10^{-5}$ | -0.01   | -0.06-0.03  | 0.563 | 0.02             |
| 0.1               | 0.02   | -0.03-0.06 | 0.458 | -0.09  | -0.13--0.04 | $8.77 \times 10^{-5}$ | -0.02   | -0.06-0.02  | 0.373 | 0.06             |
| 0.2               | 0.01   | -0.04-0.05 | 0.706 | -0.09  | -0.13--0.04 | $8.86 \times 10^{-5}$ | -0.04   | -0.08-0.01  | 0.105 | 0.18             |
| 0.3               | -0.01  | -0.05-0.04 | 0.789 | -0.09  | -0.13--0.04 | $8.78 \times 10^{-5}$ | -0.04   | -0.08-0.00  | 0.068 | 0.23             |
| 0.4               | -0.01  | -0.05-0.03 | 0.656 | -0.09  | -0.13--0.04 | $8.85 \times 10^{-5}$ | -0.04   | -0.09--0.00 | 0.040 | 0.29             |
| 0.5               | -0.01  | -0.05-0.03 | 0.677 | -0.09  | -0.13--0.04 | $8.86 \times 10^{-5}$ | -0.05   | -0.09--0.00 | 0.031 | 0.32             |

Table S2. Validation analyses: Linear regression examining the main effects of the polygenic environmental sensitivity score, negative aspects of parenting and their interaction on emotional problems

| P value threshold | Main effects of the polygenic environmental sensitivity score <sup>a</sup> |            |       | Main effects of negative aspects of parenting <sup>a</sup> |           |                        | Polygenic environmental sensitivity score by negative aspects of parenting interaction <sup>b</sup> |            |       |                  |
|-------------------|--|------------|-------|--|-----------|------------------------|---|------------|-------|------------------|
|                   | $\beta$  | 95% CI     | P     | $\beta$  | 95% CI    | P                      | $\beta$   | 95% CI     | P     | % R <sup>2</sup> |
| 0.001             | 0.01   | -0.04-0.04 | 0.908 | 0.23   | 0.19-0.27 | $1.66 \times 10^{-28}$ | -0.01   | -0.06-0.03 | 0.505 | 0.03             |
| 0.01              | -0.01  | -0.05-0.03 | 0.684 | 0.23   | 0.19-0.27 | $2.12 \times 10^{-28}$ | 0.01  | -0.03-0.05 | 0.630 | 0.02             |
| 0.05              | 0.01   | -0.03-0.05 | 0.581 | 0.23   | 0.19-0.27 | $2.14 \times 10^{-28}$ | 0.05  | 0.00-0.09  | 0.035 | 0.29             |
| 0.1               | 0.02   | -0.03-0.06 | 0.465 | 0.23   | 0.19-0.27 | $1.60 \times 10^{-28}$ | 0.05  | 0.01-0.10  | 0.014 | 0.40             |
| 0.2               | 0.01   | -0.03-0.05 | 0.706 | 0.23   | 0.19-0.27 | $2.06 \times 10^{-28}$ | 0.06  | 0.01-0.10  | 0.009 | 0.45             |
| 0.3               | 0.01   | -0.05-0.04 | 0.819 | 0.23   | 0.19-0.27 | $2.09 \times 10^{-28}$ | 0.05  | 0.01-0.10  | 0.020 | 0.35             |
| 0.4               | -0.01  | -0.05-0.03 | 0.631 | 0.23   | 0.19-0.27 | $2.02 \times 10^{-28}$ | 0.05  | 0.01-0.10  | 0.021 | 0.35             |
| 0.5               | -0.01  | -0.05-0.03 | 0.612 | 0.23   | 0.19-0.27 | $2.12 \times 10^{-28}$ | 0.05  | 0.01-0.10  | 0.020 | 0.36             |

Table S3. Validation analyses: Effects of the polygenic environmental sensitivity score on measures of parenting

| P value threshold | Overall parenting |            |       | Positive aspects of parenting |            |       | Negative aspects of parenting |            |       |
|-------------------|-------------------|------------|-------|-------------------------------|------------|-------|-------------------------------|------------|-------|
|                   | $\beta$           | 95% CI     | P     | $\beta$                       | 95% CI     | P     | $\beta$                       | 95% CI     | P     |
| 0.001             | 0.02              | -0.07-0.11 | 0.629 | -0.03                         | -0.11-0.06 | 0.507 | 0.01                          | -0.08-0.10 | 0.842 |
| 0.01              | 0.01              | -0.08-0.10 | 0.839 | 0.02                          | -0.07-0.10 | 0.717 | 0.02                          | -0.07-0.10 | 0.726 |
| 0.05              | 0.01              | -0.09-0.09 | 0.994 | 0.01                          | -0.08-0.09 | 0.849 | 0.01                          | -0.09-0.08 | 0.913 |
| 0.1               | 0.02              | -0.07-0.11 | 0.689 | 0.01                          | -0.09-0.08 | 0.958 | 0.02                          | -0.07-0.10 | 0.714 |
| 0.2               | 0.02              | -0.07-0.11 | 0.705 | -0.01                         | -0.09-0.08 | 0.858 | 0.01                          | -0.07-0.10 | 0.769 |
| 0.3               | 0.01              | -0.08-0.10 | 0.892 | -0.01                         | -0.09-0.08 | 0.884 | 0.01                          | -0.09-0.09 | 0.993 |
| 0.4               | 0.01              | -0.08-0.09 | 0.917 | 0.01                          | -0.08-0.09 | 0.950 | 0.01                          | -0.08-0.09 | 0.900 |
| 0.5               | 0.01              | -0.08-0.10 | 0.888 | 0.01                          | -0.08-0.09 | 0.910 | 0.01                          | -0.08-0.10 | 0.817 |

Table S4. Validation analyses: Linear regression examining the main effects of the polygenic environmental sensitivity score, parent-reported parenting and their interaction on child-reported emotional problems

| P value threshold | Main effects of the polygenic environmental sensitivity score <sup>a</sup> |            |       | Main effects of parenting <sup>a</sup> |           |                        | Polygenic environmental sensitivity score by parenting interaction <sup>b</sup> |            |       |                  |
|-------------------|--|------------|-------|--|-----------|------------------------|---|------------|-------|------------------|
|                   | $\beta$  | 95% CI     | P     | $\beta$                                | 95% CI    | P                      | $\beta$   | 95% CI     | P     | % R <sup>2</sup> |
| 0.001             | 0.01   | -0.04-0.06 | 0.783 | 0.17                                   | 0.12-0.22 | $5.30 \times 10^{-11}$ | 0.03  | -0.02-0.08 | 0.279 | 0.08             |
| 0.01              | 0.02   | -0.03-0.07 | 0.345 | 0.17                                   | 0.12-0.22 | $4.74 \times 10^{-11}$ | 0.03  | -0.02-0.08 | 0.265 | 0.09             |
| 0.05              | 0.05   | 0.00-0.10  | 0.038 | 0.17                                   | 0.12-0.22 | $3.62 \times 10^{-11}$ | 0.04  | -0.01-0.09 | 0.146 | 0.15             |
| 0.1               | 0.05   | 0.00-0.11  | 0.035 | 0.17                                   | 0.12-0.22 | $4.22 \times 10^{-11}$ | 0.05  | -0.00-0.10 | 0.065 | 0.24             |
| 0.2               | 0.05   | -0.01-0.10 | 0.081 | 0.17                                   | 0.12-0.22 | $4.99 \times 10^{-11}$ | 0.06  | 0.00-0.11  | 0.034 | 0.31             |
| 0.3               | 0.03   | -0.02-0.08 | 0.242 | 0.17                                   | 0.12-0.22 | $4.84 \times 10^{-11}$ | 0.06  | 0.00-0.11  | 0.036 | 0.30             |
| 0.4               | 0.03   | -0.02-0.08 | 0.303 | 0.17                                   | 0.12-0.22 | $4.79 \times 10^{-11}$ | 0.07  | 0.01-0.12  | 0.016 | 0.40             |
| 0.5               | 0.03   | -0.02-0.08 | 0.284 | 0.17                                   | 0.12-0.22 | $4.84 \times 10^{-11}$ | 0.07  | 0.02-0.12  | 0.011 | 0.44             |

Table S5. Associations between the polygenic environmental sensitivity score and parental psychopathology and comorbid internalising and externalising disorders

| P value threshold | Parental psychopathology |            |       | Comorbid externalising disorders |            |       | Comorbid internalising disorders |             |                         |
|-------------------|--------------------------|------------|-------|----------------------------------|------------|-------|----------------------------------|-------------|-------------------------|
|                   | $\beta$                  | 95% CI     | P     | $\beta$                          | 95% CI     | P     | $\beta$                          | 95% CI      | P                       |
| 0.001             | -0.53                    | -1.81-0.75 | 0.417 | 0.05                             | -0.12-0.22 | 0.557 | 0.01                             | -0.19-0.22  | 0.940                   |
| 0.01              | -0.86                    | -2.13-0.41 | 0.182 | -0.10                            | -0.27-0.07 | 0.241 | -0.31                            | -0.52--0.10 | 0.003                   |
| 0.05              | -0.69                    | -1.96-0.58 | 0.288 | -0.07                            | -0.24-0.09 | 0.382 | -0.38                            | -0.59--0.17 | 4.51 x 10 <sup>-4</sup> |
| 0.1               | -0.32                    | -1.61-0.97 | 0.627 | -0.05                            | -0.22-0.11 | 0.539 | -0.32                            | -0.52--0.11 | 0.004                   |
| 0.2               | -0.05                    | -1.35-1.25 | 0.941 | -0.04                            | -0.21-0.12 | 0.627 | -0.24                            | -0.44--0.03 | 0.023                   |
| 0.3               | -0.01                    | -1.30-1.28 | 0.988 | -0.03                            | -0.20-0.13 | 0.682 | -0.21                            | -0.41--0.00 | 0.044                   |
| 0.4               | 0.13                     | -1.17-1.42 | 0.848 | -0.04                            | -0.20-0.13 | 0.647 | -0.20                            | -0.40-0.00  | 0.053                   |
| 0.5               | 0.06                     | -1.24-1.36 | 0.926 | -0.03                            | -0.20-0.14 | 0.727 | -0.21                            | -0.41--0.01 | 0.044                   |

Table S6. Treatment response analyses: Linear mixed model examining the effects of the polygenic environmental sensitivity score on treatment response (change in the severity of the primary anxiety disorder) adjusting for comorbid internalising disorders

| P value threshold | Overall response |            |       |                  | Response to individual CBT |             |       |                  | Response to group based CBT |            |       |                  | Response to brief parent-led CBT |            |                         |                  |
|-------------------|------------------|------------|-------|------------------|----------------------------|-------------|-------|------------------|-----------------------------|------------|-------|------------------|----------------------------------|------------|-------------------------|------------------|
|                   | $\beta$          | 95% CI     | P     | % R <sup>2</sup> | $\beta$                    | 95% CI      | P     | % R <sup>2</sup> | $\beta$                     | 95% CI     | P     | % R <sup>2</sup> | $\beta$                          | 95% CI     | P                       | % R <sup>2</sup> |
| 0.001             | 0.01             | -0.04-0.05 | 0.698 | 0.02             | -0.05                      | -0.14-0.04  | 0.265 | 0.22             | 0.03                        | -0.03-0.09 | 0.283 | 0.17             | 0.05                             | -0.06-0.15 | 0.390                   | 0.46             |
| 0.01              | 0.03             | -0.02-0.07 | 0.226 | 0.08             | -0.09                      | -0.19--0.00 | 0.049 | 0.97             | 0.05                        | -0.01-0.11 | 0.089 | 0.38             | 0.08                             | -0.02-0.17 | 0.107                   | 0.97             |
| 0.05              | 0.04             | -0.01-0.08 | 0.091 | 0.13             | -0.11                      | -0.21--0.02 | 0.015 | 1.74             | 0.04                        | -0.02-0.10 | 0.234 | 0.24             | 0.19                             | 0.10-0.28  | 2.72 x 10 <sup>-5</sup> | 5.21             |
| 0.1               | 0.03             | -0.01-0.08 | 0.157 | 0.08             | -0.11                      | -0.20--0.02 | 0.016 | 1.65             | 0.02                        | -0.04-0.08 | 0.458 | 0.11             | 0.20                             | 0.12-0.29  | 7.30 x 10 <sup>-6</sup> | 5.61             |
| 0.2               | 0.03             | -0.01-0.08 | 0.126 | 0.11             | -0.08                      | -0.17-0.01  | 0.076 | 1.01             | 0.01                        | -0.05-0.08 | 0.633 | 0.07             | 0.21                             | 0.12-0.31  | 3.61 x 10 <sup>-6</sup> | 5.95             |
| 0.3               | 0.03             | -0.02-0.07 | 0.233 | 0.06             | -0.09                      | -0.17-0.00  | 0.052 | 1.18             | 0.01                        | -0.06-0.07 | 0.866 | 0.02             | 0.20                             | 0.11-0.29  | 1.21 x 10 <sup>-5</sup> | 5.36             |
| 0.4               | 0.03             | -0.02-0.07 | 0.262 | 0.05             | -0.09                      | -0.18--0.00 | 0.049 | 1.24             | 0.01                        | -0.05-0.07 | 0.824 | 0.03             | 0.19                             | 0.10-0.29  | 3.63 x 10 <sup>-5</sup> | 4.96             |
| 0.5               | 0.03             | -0.02-0.07 | 0.245 | 0.05             | -0.09                      | -0.18--0.00 | 0.047 | 1.26             | 0.01                        | -0.06-0.06 | 0.892 | 0.02             | 0.20                             | 0.11-0.29  | 2.15 x 10 <sup>-5</sup> | 5.29             |

Note. To account for data collected longitudinally, all models included the random effects of participant and the linear and quadratic effects of time. All models also included sex, age (centred), primary diagnosis (Generalised Anxiety Disorder (GAD) Social Anxiety Disorder (SoAD) Specific Phobia (SP) Separation Anxiety Disorder (SAD) or "Other anxiety" disorder) and treatment type (individual based CBT group based CBT or brief parent-led CBT) and the presence of comorbid internalising disorders. All models included the random effects of trial. Regression weights ( $\beta$ ) significantly greater than zero indicate that this variable is associated with a poorer response following treatment.

Table S7. Differences between treatment groups before and after propensity score matching

|   | Individual vs. Group |        |                    |       | Individual CBT vs. brief parent-led CBT |        |                    |       | Group CBT vs. Brief parent-led CBT |        |                    |       |
|---|----------------------|--------|--------------------|-------|---|--------|--------------------|-------|------------------------------------|--------|--------------------|-------|
|   | Unmatched            |        | Matched            |       | Unmatched                               |        | Matched            |       | Unmatched                          |        | Matched            |       |
|   | B (95% CI)           | P      | B (95% CI)         | P     | B (95% CI)                              | P      | B (95% CI)         | P     | B (95% CI)                         | P      | B (95% CI)         | P     |
| Severity of primary diagnosis at baseline | 0.39 (0.13-0.66)     | 0.004  | -0.20 (-0.55-0.14) | 0.249 | -0.37 (-0.57--0.16)                     | 0.001  | -0.08 (-0.32-0.17) | 0.544 | 0.82 (0.60-1.03)                   | <0.001 | -0.10 (-0.37-0.18) | 0.499 |
| Sex                                       | 0.01 (-0.47-0.50)    | 0.959  | 0.54 (-0.05-1.12)  | 0.072 | 0.21 (-0.20-0.62)                       | 0.322  | 0.16 (-0.33-0.66)  | 0.514 | 0.02 (-0.36-0.39)                  | 0.933  | 0.17 (-0.27-0.61)  | 0.458 |
| Age                                       | 0.29 (0.14-0.44)     | <0.001 | 0.02 (-0.16-0.19)  | 0.858 | 0.11 (0.01-0.21)                        | 0.032  | -0.07 (-0.19-0.05) | 0.272 | 0.13 (0.04-0.23)                   | 0.008  | -0.01 (-0.14-0.11) | 0.829 |
| Primary diagnosis                         |                      |        |                    |       |   |        |                    |       |                                    |        |                    |       |
| GAD                                       |                      |        |                    |       |   |        |                    |       |                                    |        |                    |       |
| SoAD                                      | 0.34 (-0.33-1.01)    | 0.316  | 0.03 (-0.76-0.82)  | 0.937 | 1.03 (0.46-1.60)                        | <0.001 | 0.14 (-0.58-0.86)  | 0.701 | -0.59 (-1.09--0.09)                | 0.021  | 0.25 (-0.34-0.84)  | 0.411 |
| SP  | 0.99 (0.24-1.73)     | 0.010  | -0.15 (-1.05-0.75) | 0.738 | 1.95 (1.30-2.60)                        | <0.001 | -0.13 (-0.92-0.67) | 0.757 | -0.87 (-1.51--0.23)                | 0.008  | 0.11 (-0.63-0.85)  | 0.769 |
| SAD                                       | 0.76 (0.10-1.42)     | 0.024  | -0.13 (-0.89-0.64) | 0.747 | 1.53 (0.96-2.09)                        | <0.001 | -0.20 (-0.91-0.50) | 0.574 | -0.70 (-1.21--0.20)                | 0.007  | 0.10 (-0.49-0.69)  | 0.748 |
| Other                                     | 0.61 (-0.40-1.63)    | 0.236  | -0.32 (-1.57-0.94) | 0.620 | 0.80 (0.03-1.57)                        | 0.043  | -0.03 (-0.99-0.92) | 0.945 | -0.01 (-0.80-0.78)                 | 0.980  | 0.40 (-0.51-1.30)  | 0.388 |
| Comorbidity                               |                      |        |                    |       |   |        |                    |       |                                    |        |                    |       |
| Mood disorder                             | -0.03 (-0.86-0.79)   | 0.942  | -0.18 (-1.14-0.77) | 0.709 | 0.28 (-0.34-0.90)                       | 0.374  | 0.37 (-0.45-1.19)  | 0.378 | -0.23 (-0.88-0.42)                 | 0.491  | -0.23 (-1.06-0.60) | 0.581 |
| Externalising disorder                    | 0.44 (-0.15-1.04)    | 0.144  | 0.11 (-0.58-0.81)  | 0.747 | 0.90 (0.41-1.40)                        | <0.001 | 0.02 (-0.57-0.62)  | 0.936 | -0.32 (-0.82-0.18)                 | 0.205  | 0.35 (-0.22-0.93)  | 0.225 |
| Parental psychopathology                  | 0.56 (0.29-0.84)     | <0.001 | -0.05 (-0.37-0.28) | 0.777 | 0.25 (0.03-0.47)                        | 0.023  | -0.04 (-0.31-0.23) | 0.769 | 0.35 (0.13-0.58)                   | 0.002  | -0.04 (-0.31-0.23) | 0.763 |

## References

1. PSMATC H2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing [program]. version 4.0.11 22oct2014 version: <http://ideas.repec.org/c/boc/bocode/s432001.html>, 2003.
2. Austin PC. Some methods of propensity-score matching had superior performance to others: results of an empirical investigation and Monte Carlo simulations. *Biometrical journal. Biometrische Zeitschrift* 2009;**51**(1):171-84 doi: 10.1002/bimj.200810488[published Online First: Epub Date].