

Genetic and Environmental Links between  
General Factors of Psychopathology and Cognitive Ability in Early Childhood

*In press at Clinical Psychological Science*

Andrew D. Grotzinger<sup>a,\*</sup>

Amanda K. Cheung<sup>a,\*</sup>

Megan W. Patterson<sup>a</sup>

K. Paige Harden<sup>a,b</sup>

Elliot M. Tucker-Drob<sup>a,b</sup>

<sup>a</sup>Department of Psychology, The University of Texas at Austin, 108 E Dean Keeton, Stop A8000, Austin, TX 78712-0187.

<sup>b</sup>Population Research Center, University of Texas at Austin, Austin, TX

\*These authors contributed equally to this work.

Correspondence concerning this article should be addressed to Elliot M. Tucker-Drob, Department of Psychology, The University of Texas at Austin, 108 E Dean Keeton, Stop A8000, Austin, TX 78712-0187. Phone: 512-471-1157, E-mail: [tuckerdrob@utexas.edu](mailto:tuckerdrob@utexas.edu)

*Keywords:* psychopathology, intelligence, behavioral genetics, early childhood

**Abstract**

In adults, psychiatric disorders are highly comorbid, and are negatively associated with cognitive abilities. Individual cognitive measures have been linked with domains of child psychopathology, but the specificity of these associations and the extent to which they reflect shared genetic influences are unknown. This study examines the relation between general factors of cognitive ability (*g*) and psychopathology (*p*) in early development using two genetically-informative samples: the Texas “Tiny” Twin project (TXtT;  $N = 626$  individuals, age range = 0.16 – 6.31 years) and the Early Childhood Longitudinal Study-Birth Cohort (ECLS-B;  $N \approx 1,300$  individual twins, age range = 3.7 – 7.1 years). The total *p-g* correlation (-.21 in ECLS-B; -.34 in TxtT) was primarily attributable to genetic and shared environmental factors. The early age range of participants indicates that the *p-g* association is a reflection of overlapping genetic and shared environmental factors that operate in the first years of life.

**Genetic and Environmental Links between  
General Factors of Psychopathology and Cognitive Ability in Early Childhood**

Psychopathology is a leading source of health burden in both children and adults worldwide, with behavioral and mental health problems affecting up to 20% of children (Egger & Angold, 2006; Whiteford et al., 2013). The three domains of psychopathology of primary concern during early childhood are externalizing (aggressive and oppositional behaviors), internalizing (fearfulness and sadness), and attentional difficulties (Egger & Angold, 2006; Koot, van den Oord, Verhulst, & Boomsma, 1997). These dimensions can be reliably identified at very young ages (Achenbach & Rescorla, 2000a) and forecast mental health problems across the lifespan (Caspi, Moffitt, Newman, & Silva, 1996; Keenan, Shaw, Delliquadri, Giovannelli & Walsh, 1998; Mesman & Koot, 2001).

Recent research in late childhood, adolescence and adulthood supports a dimensional model of psychopathology, in which clinical diagnoses represent extreme ends of a continuous quantitative trait spanning normal-range functioning, subclinical symptoms, and clinically-defined disorders (Kotov et al., 2017; Plomin, Haworth, & Davis, 2009). Moreover, these continuous dimensions of liability overlap considerably across symptom domains. Elevated symptoms in one domain of psychopathology are associated with elevated symptoms in other domains of psychopathology, and clinically-severe levels of psychiatric disease are highly comorbid (Caspi & Moffitt, 2018; Smoller et al., 2018). Factor analytic work has identified a transdiagnostic dimension, *p*, representing a general pattern of cross-domain symptomatology (Caspi et al., 2014; Lahey et al., 2012), which is partially the result of non-specific genetic liabilities that confer risk across different psychiatric disorders (Antilla et al., 2018; Bulik-Sullivan et al., 2015; Grotzinger et al., 2018; Lahey, Krueger, Rathouz, Waldman, & Zald, 2016; Tackett et al., 2013).

A key outstanding question is when in development transdiagnostic vulnerabilities to psychopathology are apparent. The *p*-factor might arise from dynamic processes that unfold across development: A temporally primary disorder causes the emergence of secondary disorders. For example, genetic risk for attention difficulties may come to be correlated with depressive symptoms when a child's behavioral problems in school elicit harsh interactions with his parents and teachers. Alternatively, risks for a variety of different mental health problems might operate through shared biological processes, such that co-occurring elevations in mental health symptoms are evident even early in childhood. To date, the earliest ages that a *p*-factor has been estimated is late childhood (Martel et al., 2017 [ages 6-12]; Tackett et al., 2013 [ages 6-17]). We are not aware of any investigations of *p* during the infancy and preschool years. Identifying the extent to which psychopathology symptoms co-occur across domains in early childhood informs the potential viability of developing transdiagnostic therapies to target a broad set of psychiatric symptoms (Barlow, Allen, & Choate, 2004; Caspi & Moffitt, 2018). In addition, the ubiquitous phenotypic comorbidity and genetic correlations observed for adult psychiatric disorders have reinvigorated calls to refine psychiatric nosology (Antilla et al., 2018), and genetic research on the overlap among childhood disorders also has the potential to inform similar classification questions.

In addition to estimating the existence of general genetic liability to psychopathology in early childhood, this paper also examines the genetic association between *p* and general cognitive ability. Early deficits in cognitive ability have been proposed as a key contributor to general vulnerability to psychopathology across the life course. For instance, Caspi et al. (2014) reported negative associations between a "brain integrity" factor, formed from cognitive, psychomotor, and neurological indices measured at age three years, and a general factor of psychopathology, formed from measures taken in early adulthood. In addition, a negative

association between global executive function and the *p*-factor was found at age six years ( $\beta = -.24$ ; Martel et al., 2017), and a similar negative association was found between the *p*-factor and both IQ and teacher-reported school functioning in females at ages 5-11 years (Lahey et al., 2015).

The mechanisms underlying negative associations between cognitive ability early in life and psychopathology in later development are ambiguous. Such associations might arise because of the reciprocal effects of cognitive ability on mental health and vice versa. For example, cognitive deficits might impede social interactions, and behavioral problems might impede learning. Consistent with this latter account, internalizing and externalizing problems at age 24 months were reported to prospectively predict lower cognitive ability (as measured using the Woodcock-Johnson) in the first grade (Bub, McCartney, & Willett, 2007). Similarly, internalizing and externalizing at age 7 years were found to prospectively predict lower GPA, math, and reading ability at age 10 years (Bodovski and Youn, 2011; Zhou, Main, & Wang, 2010). If these dynamic mechanisms are the exclusive basis for the association between psychopathology and cognitive development, we would expect such associations to be weaker or entirely absent in early childhood and emerge and strengthen with development. In contrast, an early link between cognitive ability and psychopathology in the preschool years would lend support to the hypothesis that both are influenced by a common set of risk factors that are already present in early life. These could include genetic risks, neurodevelopmental problems, and/or early environmental deprivation.

Prior findings indicate that the *p*-factor is moderately heritable in school age samples, and that individual symptom domains are negatively, genetically correlated with intelligence. For example, using molecular genetic methods, Neumann et al. (2016) found that a general psychopathology factor was 38% heritable in children ages 6-8 years. In twin samples, the *g*-

factor has been estimated to be approximately 23% heritable in early childhood (ages 2-4 years), with increases to approximately 62% heritability by age 7 years (Davis, Haworth, & Plomin, 2009). Negative genetic associations have been found between IQ and antisocial behavior in children at age 5 years ( $r_g = -.41$ ; Koenen et al., 2006), ADHD in children at age 5 years ( $r_g = -.45$ ; Kuntsi et al., 2004), and hyperactivity in children ages 8-11 years ( $r_g = -.07$ ; Paloyelis, Rijdsdijk, Wood, Asherson, & Kuntsi, 2010). Moreover, genetic associations were found to mediate over 65% of the phenotypic association between IQ and behavior problems in all three studies. In addition, negative genetic correlations have also been reported for specific pre-academic skills and behavior problems in early childhood (e.g., reading and inattention,  $r_g = -.26$ ; Ebejer et al., 2010). However, despite a large literature base on grade-school samples, no study of preschool-aged children has sought to identify a *p*-factor or decompose this general liability, and its association with the *g*-factor, into genetic and environmental factors. Whether genetic influences on *g* overlap with non-specific genetic influences on psychopathology in early childhood remains an open question.

The current study evaluates the association between a general factor of psychopathology (*p*) and a general cognitive ability factor (*g*), and decomposes this association into genetic and environmental influences using quantitative genetic modeling. Data are drawn from two American twin studies of early child development, each of which provided five measures of cognitive and psychomotor development, and parent-report measures of internalizing, externalizing, and attentional/self-regulatory problems. We fit confirmatory factor models to the eight phenotypes and, using an integrative data analysis approach, estimate multivariate biometric models of the associations between abilities and psychopathology.

## Method

### Participants

**Early Childhood Longitudinal Study – Birth Cohort (ECLS-B) Sample**

Twins were drawn from the third (age 4 years), fourth (age 5 years), and fifth (age 6 years) wave of the ECLS-B study, a nationally representative sample of children born in the U.S. in 2001 (Snow et al., 2009). Data were available from 650 pairs of twins (1,300 individuals) in the third wave and 550 pairs of twins (1,100 individuals) in the fourth wave. The subset of participants who had not yet entered kindergarten by wave 4 were invited to participate in a fifth wave, yielding 150 pairs of twins (300 individuals) in the fifth wave.<sup>1</sup> Participants ranged in age from 3.71 to 7.07 years old across all waves ( $M_{\text{age at wave3}} = 4.40$  years,  $SD_{\text{age at wave3}} = 0.33$  years). The sample was racially/ethnically diverse (62% Caucasian, 16% African-American, 16% Hispanic, 2% Asian, 4% multi-racial), and 50% of participants were female. Tucker-Drob et al. (2011) have previously reported that the twin subsample is similarly representative of family SES compared to the full ECLS-B sample. SES is computed in ECLS-B as the composite of five variables: paternal education, maternal education, paternal occupation, maternal occupation, and family income (Hollingshead, 1975). These individual variables were standardized against the full ECLS-B sample to have a mean of 0 and standard deviation of 1 and then averaged to create an unstandardized SES composite. The composite SES measure had a mean of -0.05 and a standard deviation of 0.86 in the full sample, and a mean of 0.13 and standard deviation of 0.87 in the twin subsample. The SES composite score ranged from -2.13 to 2.12 in the twin subsample. Informed parental consent was obtained from all study participants.

Opposite-sex twin pairs were coded as dizygotic by default. Trained researchers rated the similarity of same-sex twin pairs (1 = “no difference”, 2 = “slight difference”, and 3 = “clear difference”) on eye color, hair texture, hair color, complexion, facial appearance, and ear lobe shape. In line with the procedure reported in Tucker-Drob et al. (2011), scores were summed for each twin pair resulting in a composite ranging from 6 to 18. Based on the bimodal shape of this

distribution, twins whose scores fell in the 6-8 range were classified as monozygotic (MZ), and twins scoring above 9 were classified as dizygotic (DZ).<sup>2</sup> In the final sample, 30% of twins were classified as MZ, 30% as same-sex DZ, and 40% as opposite-sex DZ.

### **Texas “Tiny” Twin Project Sample**

A second sample of twin participants was drawn from the downward extension of the Texas Twin Project (Harden, Tucker-Drob, & Tackett, 2013). The Texas “Tiny” Twin Project (TXtT) recruited families with twins or multiples of ages 0-6 years on an on-going basis. Potential families were identified from birth records provided by the Texas Department of State Health Services and from community outreach. Data were collected via paper or online surveys for a maximum of 20 follow-up waves until the twins or multiples reached age 6 years (see Cheung, Harden, & Tucker-Drob, 2015, 2016 for a detailed schedule of the follow-up surveys). Data were available on 626 individual twins or multiples ( $M_{\text{age at baseline}} = 2.55$  years,  $SD_{\text{age at baseline}} = 1.28$  years). This sample was racially/ethnically diverse (67.41% Caucasian, 5.27% Latino, 5.75% African American, 2.24% Asian, and 16.77% racially/ethnically mixed). Among their primary caregivers (92.33% biological mother), 4.15% completed no more than high school, 8.47% completed no more than some college, 39.62% completed no more than 2-year or 4-year college, and 47.12% completed education beyond college. Informed parental consent was obtained from all study participants.

To diagnose zygosity of same-sex pairs, we analyzed parental ratings on the pair-wise physical similarity of their twins or multiples using two-class Latent Class Analysis (see Harden, Kretsch, Tackett, & Tucker-Drob, 2014). Zygosity assignment based on physical similarity ratings is highly reliable (Forget-Dubois et al., 2003; Heath et al., 2003; Price et al., 2000; Rietveld et al., 2000). This resulted in 142 MZ twins (74 male and 68 female), 234 same-sex DZ twins (124 male and 110 female), 190 opposite-sex DZ twins (95 male and 95 female), and 60

triplets (8 male MZ, 12 female MZ, 19 male DZ, and 21 female DZ). Among the 626 individuals, 268 of them provided up to 11 sets of follow-up data. This resulted in a final effective sample of 1,398 observations from 626 individuals.

### **Ethical Standards**

For ECLS-B, the authors received a license from the National Center for Education Statistics (NCES) to access the deidentified and anonymized restricted-use ECLS-B dataset. The ECLS-B study was approved by the NCES Institutional Review Board for human subject research. The TXtT study was approved by the University of Texas at Austin Institutional Review Board (2009-12-0040: A Sibling and Twin Study of Healthy Development in Children and Adolescents). Procedures for both studies complied with the Helsinki Declaration of 1975, as revised in 2008.

### **Measures**

#### **Cognitive and Psychomotor Abilities**

**ECLS-B:** Participants completed 85 items designed by ECLS-B to capture pre-reading skills in the content areas of recognizing simple words, phonological awareness, knowledge of print conventions, and matching words (Najarian, Snow, Lennon, & Kinsey, 2010). Early mathematics skills were assessed using 45 items at wave 3 and 58 items at waves 4 and 5 designed by ECLS-B in the content areas of number sense, geometry, counting, operations, and patterns. Pre-reading and math ability estimates were obtained for each child using Item Response Theory. Children's receptive and expressive verbal abilities (i.e., oral language) were assessed using the 'Let's Tell Stories' task, adapted from the pre-LAS subtest. An experimenter started by telling two scripted stories while pointing to a series of pictures; the child was then asked to re-tell the story using the pictures as prompts. Responses were scored by trained coders on a 5-point scale using standardized procedures.

Trained researchers also assessed children's gross motor skills in hopping, balancing on one foot, skipping, walking backwards, and catching a bean bag. Performance was scored according to standardized procedures, and scores were summed to create a composite ranging from 0 to 13. Fine motor skills were assessed using a building task and a copying task. For the first task, a child watched an experimenter build a gate out of a set of blocks and then was asked to build the gate using a second set of blocks. For the second task, the child was asked to copy the shapes of a square, a triangle, and an asterisk using a pencil and paper. Scores from the two tasks were summed together to create a composite score ranging from 0 to 4.

**TXtT:** As reported in our previous work (Cheung, Harden, & Tucker-Drob, 2015), children in the TXtT sample were assessed on five domains of cognitive and psychomotor functioning using the Ages and Stages Questionnaire, Third Edition (ASQ; Squires & Bricker, 2009). For each of the five ASQ domains (see Table S1), primary caregivers completed a set of 5-10 age-appropriate items that varied across age and measurement waves. Item-sets for adjacent age groups contained overlapping items to allow for vertical scaling of scores, which facilitates cross-age comparisons. Most items described a specific task and provided concrete guidelines for primary caregivers to rate their children's abilities on a 3-point Likert scale (0 = *no*, 1 = *sometimes*, and 2 = *yes*). ASQ has been reported to demonstrate, on average, sensitivity and specificity of 86% in identifying children with developmental concerns (Squires et al., 2009) and shows concurrent validity with other cognitive and psychomotor assessments (Gollenberg et al., 2010; Schonhaut, Armijo, Schönstedt, Alvarez, & Cordero, 2013; Simard et al., 2012; Yu et al., 2007). Rasch Item Response Theory (1PL IRT) analyses were used to obtain each domain score, with higher scores indicating more advanced development.

### **Psychopathology**

**ECLS-B:** Primary caregivers rated their twins' behaviors on a number of five-point Likert items (1 = *Never*, 5 = *Very Often*). In line with Tucker-Drob and Harden (2013), five items were summed to create an externalizing composite that indexed how often a child has temper outbursts, gets angry, engages in physical aggression, destroys others' things, and bothers and annoys other children. The internalizing composite consisted of parent report on how often the child appeared unhappy or worried. An attention-deficit/hyperactivity composite was created using parent report on how often the child is overly active, keeps working until finished, and pays attention well. The latter two items were reverse coded for consistent direction in scoring. All psychopathology items were drawn from the Preschool and Kindergarten Behavioral Scale—2<sup>nd</sup> edition (Riccio, 1995) and the Social Skills Rating System (Van Horn et al., 2007). Item-level confirmatory factor analyses (CFAs) revealed that all factor loadings were positive and highly significant.

**TXtT:** Internalizing and externalizing were measured with the ASEBA Child Behavior Checklist for Ages 1.5-5 (CBCL). CBCL is a parent report of young children's emotional and behavioral problems, which can be broadly categorized as internalizing and externalizing (Achenbach & Rescorla, 2000b; see Table S2). It has been shown to reliably measure children's problem behavior and is commonly used in developmental research (Achenbach & Rescorla, 2000a). Primary caregivers rated how well various problem behaviors apply to their children on a 3-point Likert scale (0 = *not true*, 1 = *somewhat or sometimes true*, and 2 = *very true or often true*). Raw scores were converted to standardized scores as outlined in the ASEBA manual (Achenbach & Rescorla, 2000a), with higher scores indicating greater extent of problem behaviors. CBCL internalizing is composed of four subscales: emotionally reactive, anxious/depressed, somatic complaints, and withdrawn. CBCL externalizing is composed of two subscales: aggressive behavior and attention problems. Excluding attention problems from the

CBCL externalizing scores in order to more directly parallel the externalizing scores used in ECSL-B produced a very similar pattern of results as those reported in this paper.

Self-regulation problems were assessed with the Ages and Stages Questionnaire: Social-Emotional (ASQ:SE), a measure of psychosocial adjustment during early childhood (Squires et al., 2003). Each primary caregiver completed a set of 18-32 age-appropriate items that varied across measurement waves (see Table S2). Item-sets for adjacent age groups contained overlapping items to allow valid comparison of scores across children of different ages. Most items described a specific task and provided concrete guidelines for primary caregivers to rate their children's social and emotional competence on a 3-point Likert scale (0 = *rarely or never*, 1 = *sometimes*, and 2 = *most of the time*). ASQ:SE demonstrated, on average, sensitivity of 78% and specificity of 95% in identifying children with developmental concerns (Squires et al., 2003). Twenty-one items were reverse-coded and all 77 items were analyzed using 1PL Item Response Theory to obtain an overall self-regulation score, with higher scores indicating lower competence.

### **Analyses**

The combined dataset across both studies and all waves was > 2,000 twin pairs. We included all waves of data for both studies using a sandwich estimator implemented by the complex survey option in *Mplus*. This statistical approach accounts for the non-independence among data on the same individual from different waves, and between individuals within a twin pair. Before testing the hypothesized models, multiple regression analyses were used to residualize for study-specific sex differences, linear and quadratic effects of age, and linear and quadratic age  $\times$  sex interactions on all indicators. Before synthesizing data across studies for the biometric twin models, we first examined whether the hypothesized factor configuration produced acceptable model fit in the individual studies. The model divided the indicators into

broad clusters of *p* and *g*. For these phenotypic models, we used absolute measures of model fit—model  $\chi^2$ , Comparative Fit Index (CFI), Tucker Lewis Index (TLI), and the Root Mean Square Error of Approximation (RMSEA)—that can be used to determine whether the model provides appropriate fit given the observed data.

We next fit a series of biometric models where variance in the higher-order latent factors (*g* and *p*), as well as the residual variance in the indicators, was decomposed into additive genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) factors. By definition, the *A* factors are fixed to correlate at 1.00 in MZ twins and .50 in DZ twins to reflect the fact that MZ and DZ twins share approximately 100% and 50% of their segregating genetic variants, respectively. The *C* factors, by definition, were fixed to correlate at 1.0 within all twins as these reflect environmental factors that are shared across twins and serve to make them more similar. As *E* reflects the environmental influences that are unique to each twin, the *E* factor for twin 1 and twin 2 are, by definition, uncorrelated. Biometric models were also used to examine the overlap between genetic and environmental estimates of *p* and *g*. These  $r_A$ ,  $r_C$ , and  $r_E$  estimates reflect the correlations between the additive genetic (*A*), shared environmental (*C*), and non-shared environmental (*E*) components of *p* and *g*.

Using integrative data analysis (Curran & Hussong, 2009), which is a form of meta-analysis that capitalizes on individual-level data, the samples were combined in a single model with four data groups – TXtT MZ, TXtT DZ, ECLS-B MZ, and ECLS-B DZ. As the manifest content of the measures used differed across studies (e.g., personal-social skills in TXtT only and reading in ECLS-B only), we allowed for study-specific loadings of the indicators on the *p*- and *g*-factors, study-specific *ACE* loadings on residual variance in the indicators, and study-specific intercepts. We note that within the individual studies, the integrative approach in this context still involves fitting standard bivariate twin models. By using a broad sample of indicators from the construct

space, we expected the indicators to triangulate on the same *p*- and *g*-factors in both studies (Little, Lindenberger, & Nesselroade, 1999), a principal that Spearman (1927) originally termed “the indifference of the indicator.”

We used a series of nested model comparisons to examine whether the biometric structure of these higher-order factors and their correlations could be constrained to be invariant across studies. In addition to using absolute measures of model fit, as in the phenotypic models, we also used relative measures of model fit—Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), and Satorra-Bentler  $\chi^2$  comparisons—to determine whether imposing the constraints across studies produced a significant decrement in fit to the data.

We also examined whether there was moderation by age of the genetic and environmental contributions to *p*, *g*, and their correlation. Age was included as a predictor of *p* and *g* in age moderation models. As longitudinal data were combined across children who varied in age at baseline, this approach can be considered an accelerated longitudinal analysis (Bell, 1953). These parametric moderation models (Purcell, 2002) allow for the paths from *ACE* factors to *p* and *g* and their correlations to be specified as consisting of a main effect and an interaction with age. Age moderation models were run twice: once with age centered at 3 years and again with age centered at 6 years of age. This choice of centering allowed us to estimate genetic and environmental contributions at the lower and upper regions of the observed age range. Absolute model fit indices (e.g.,  $\chi^2$ , RMSEA) are not provided for the age moderation models, as these indices are based on comparing a model-implied covariance matrix to an observed covariance matrix, but moderation models imply that the covariance structure of the data is age dependent (such that no single covariance matrix is implied).

As a final sensitivity analysis, we examined the consistency of the results from the main models reported with those in which we allowed for specific associations across specific

indicators of *p* and *g* on a data-driven basis. This aided in understanding whether (a) the *p-g* association was biased by not including associations between indicators, and (b) if there were associations between indicators of general psychopathology and cognitive ability that existed above and beyond what was explained by the general factors.

All variables, including *p* and *g*, were standardized in all models. All models were run using full information maximum likelihood (FIML) estimation in *Mplus* version 7.4 (Muthén & Muthén, 2012).

## Results

### Confirmatory Factor Models

The *g*-factor was defined by oral language/communication, gross motor, fine motor, math/problem-solving, reading (ECLS-B only) and personal-social abilities (TXtT only). The *p*-factor was defined by internalizing and externalizing scores. Attention problems (ECLS-B only) and self-regulation problems (TXtT only) were allowed to load on both *p* and *g*. Model fit statistics suggested that this model fit our data well in both ECLS-B ( $\chi^2 [18] = 88.89, p < .001$ , MLR scaling = 1.442, RMSEA = .04, CFI = .98, TLI = .97) and TXtT ( $\chi^2 [18] = 20.04, p = .33$ , MLR scaling = 1.573, RMSEA = .01, CFI = .99, TLI = .99). In ECLS-B, the *p*- and *g*-factors correlated at -.21 (95% Confidence Interval [-.27, -.16]). In TXtT, the *p*- and *g*-factors correlated at -.34 (95% CI [-.52, -.16]).

### *g* – *p* Associations Within and Across Twins

Our next goal was to depict the shape of the relationship between *g* and *p*. To accomplish this, we first output standardized factor score estimates for *g* and *p* from the above confirmatory models and plotted the continuous relationship between these two outcomes (Figure 1); these results indicated that there was an inverse linear relationship between *g* and *p*. Next, we sought to visualize the extent to which the *g-p* association is driven by shared environmental and/or genetic

risks that are shared within families. Beginning with one randomly selected twin per pair, we selected those scoring low (at least 1 *SD* below the mean) and high (at least 1 *SD* above the mean) on the *g*-factor. We calculated the average scores on the *p*-factor for each of these two groups of participants, as well as for their co-twins. These results are depicted in Figure 2. In both samples, children with high scores on *g* were characterized by lower average levels of psychopathology than those with low scores on *g*. Moreover, compared to co-twins of the low *g* group, the co-twins of the high *g* group also had lower average scores on *p*, indicating that the *g*-*p* association is due, at least in part, to influences shared by families. That this cross-twin cross-trait association appeared to be somewhat larger in MZ compared to DZ twins suggested that the familial component is, in part, genetically mediated. We then went on to examine cross-twin within-trait and cross-twin cross-trait correlations to gain an understanding of genetic and environmental components of *g* and *p* across the full range of data.

For cross-twin, within-trait correlations (e.g., twin 1's *p*-factor correlated with twin 2's *p*-factor) larger differences between MZ and DZ twins indicate genetic effects. The pattern of correlations was suggestive of genetic effects on *p* for both ECLS-B ( $r_{MZ} = .76$ ,  $SE = .06$ ;  $r_{DZ} = .44$ ,  $SE = .05$ ) and TXtT ( $r_{MZ} = .93$ ,  $SE = .04$ ;  $r_{DZ} = .70$ ,  $SE = .06$ ). There was also evidence for additive genetic effects on *g* for ECLS-B ( $r_{MZ} = .96$ ,  $SE = .01$ ;  $r_{DZ} = .76$ ,  $SE = .03$ ) and TXtT ( $r_{MZ} = .95$ ,  $SE = .02$ ;  $r_{DZ} = .79$ ,  $SE = .06$ ). Cross-twin, cross-trait correlations were then examined to understand the causes of phenotypic correlations. Larger differences between MZ and DZ cross-twin, cross-trait, correlations (e.g., twin 1's *p*-factor correlated with twin 2's *g*-factor) indicate that *p*-*g* associations are driven by genetic effects. The pattern again was suggestive of genetic effects for both ECLS-B ( $r_{MZ} = -.20$ ,  $SE = .03$ ;  $r_{DZ} = -.15$ ,  $SE = .03$ ) and TXtT ( $r_{MZ} = -.29$ ,  $SE = .08$ ;  $r_{DZ} = -.25$ ,  $SE = .09$ ). Biometric models were next used to formally confirm the genetic and environmental components suggested by these patterns.

### Biometric Models

We tested a quantitative genetic model with the same hierarchical structure used in the phenotypic CFAs. Model 1 estimated study-specific *ACE* correlations and *ACE* loadings on *p* and *g*. Model 2 differed from Model 1 only by constraining the higher-order *p*- and *g*-factor *ACE* correlations to be equal across studies. Building on the equality constraints on the *ACE* correlations specified under Model 2, Model 3a constrained the loadings on *p* and *ACE* correlations to be equal across studies and Model 3b constrained the loadings on *g* and *ACE* correlations to be equal across studies. Standardized path estimates for all four models are presented in Table 1 and estimates of model fit in Table S3. Model 1 provided good fit to the data (AIC = 72124.7, BIC = 72663.7, RMSEA = .029, CFI = .978). Model 2, in which the *ACE* correlations were constrained to be equal across studies, did not produce a significant decrement in model fit (AIC = 72122.3, BIC = 72644.4, RMSEA = .029, CFI = .979,  $\Delta\chi^2[3] = 1.86$ ,  $p = .602$ ). Models that additionally constrained *ACE* loadings for *p* (Model 3a: AIC = 72149.2, BIC = 72654.5, RMSEA = .030, CFI = .977) and *g* (Model 3b: AIC = 72162.1, BIC = 72677.4, RMSEA = .031, CFI = .976) provided good fit to the data, but fit significantly worse than Model 2, which only constrained the *ACE* correlations (Model 2 vs 3a:  $\Delta\chi^2[3] = 10.75$ ,  $p = .013$ ; Model 2 vs 3b:  $\Delta\chi^2[3] = 14.78$ ,  $p = .002$ ). Satorra-Bentler chi-square comparisons, along with AIC and BIC comparisons, indicated that the *ACE* correlations, but not the loadings on *p* or *g*, could be constrained to be equal without a significant drop in model fit. We therefore consider results from Model 2, which are displayed in Figure 3.

For both studies, variation in *p* was moderately heritable (TXtT: 46%, 95% CI [.18, .74]; ECLS-B: 63%, 95% CI [.37, .90]). The shared environment played a larger role for *p* in TxtT (47%, 95% CI [.22, .72]) than in ECLS-B (13%, 95% CI [-.09, .34]), while the reverse pattern was observed for non-shared environment (TXtT: 7%, 95% CI [-.01, .15]; ECLS-B: 24%, 95%

CI [.13, .35]). The *g*-factor was also moderately heritable (TXtT: 33%, 95% CI [.10, .56]; ECLS-B: 38%, 95% CI [.27, .49]), but was primarily influenced by the shared environment (TXtT: 62%, 95% CI [.40, .85]; ECLS-B: 57%, 95% CI [.47, .68]) in both datasets. Non-shared environmental effects on *g* were small, explaining only 5% of the variation in both studies (TXtT: 95% CI [.01, .08]; ECLS-B: 95% CI [.02, .07]). The correlation between shared environmental components of *p* and *g* was the largest ( $r_C = -.39$ , 95% CI [-.72, -.07],  $p = .017$ ), followed by the non-shared environmental correlation ( $r_C = -.33$ , 95% CI [-.58, -.09],  $p = .007$ ), and the lowest correlation between additive genetic factors ( $r_A = -.19$ , 95% CI [-.35, -.03],  $p = .019$ ).

Contributions of genetic and environmental factors to the phenotypic correlation between *p* and *g* were derived by multiplying the respective standardized *ACE* factor loadings and *ACE* correlation (e.g.,  $a_{p\text{-factor}} \times r_A \times a_{g\text{-factor}}$ ). These contributions were expressed as proportions by dividing them by the model-implied phenotypic correlation. Genetic (TXtT: 25%, 95% CI [-.08, .57]; ECLS-B: 40%, 95% CI [.07, .74]) and shared environmental factors (TXtT: 69%, 95% CI [.37, 1.00]; ECLS-B: 45%, 95% CI [.15, .75]) accounted for the majority of the phenotypic association between *p* and *g*. Despite evincing the largest *ACE* correlation, non-shared environment contributed to only 6% (95% CI [-.01, .13]) of the *p-g* association in TXtT and 15% in ECLS-B (95% CI [.05, .25]) due to relatively small *E* factor loadings. Sensitivity analyses that only included same-sex twins produced the same pattern of results for all models reported.

Overall, using data from two independent samples of very young twins, we found evidence that (a) a general factor of cognitive ability is negative associated with a general vulnerability to emotional and behavioral problems in early childhood, and (b) this association is primarily attributable to shared risk factors that are stratified between families, including both genetic liabilities and early environments shared by twins raised together.

### Age Moderation

Next, we were interested in examining whether or not the detected association between *p* and *g* was driven by older ages. To test this, we fit a model in which we allowed genetic and environmental factor loadings and correlations from Model 2 to be moderated by age (in years). We highlighted interaction parameters from the model where age was centered at age 3 years (results were similar from a model in which age was centered at age 6 years; see Table S4 for full results). The interaction terms for genetic ( $r_A' = .09, SE = .09, p = .305$ ), and shared environmental correlations were non-significant ( $r_C' = -.27, SE = .18, p = .139$ ), while the moderating effect of the non-shared environment correlation was just significant ( $r_E' = .22, SE = .11, p = .042$ ). The main effect estimate of the genetic correlation was higher when age was centered at age 3 years than when it was centered at age 6 years (Age 3:  $r_A = -.41, SE = .19, p = .028$ ; Age 6:  $r_A = -.31, SE = .14, p = .028$ ). This result is inconsistent with the hypothesis that the genetic correlation between *p* and *g* is driven by the upper range of the age distribution of the samples.

The moderating effect of age on *ACE* for *p* was non-significant for the genetic (ECLS-B:  $a_p' = .01, SE = .04, p = .846$ ; TXtT:  $a_p' = .03, SE = .06, p = .651$ ), shared environmental (ECLS-B:  $c_p' = .05, SE = .08, p = .509$ ; TXtT:  $c_p' = .08, SE = .07, p = .214$ ), and non-shared environmental interactions (ECLS-B:  $e_p' = -.04, SE = .05, p = .406$ ; TXtT:  $e_p' = .03, SE = .04, p = .450$ ). Age was a significant moderator of genetic effects on *g* for both ECLS-B ( $a_g' = .13, SE = .04, p = .001$ ) and TXtT ( $a_g' = .23, SE = .08, p = .002$ ), in the direction of increasing heritability of *g* with age (cf. Briley & Tucker-Drob, 2013).

### Sensitivity Analysis: Domain-Specific Associations

Our extraction of common *p*- and *g*-factors from the psychopathology and cognitive function data should not be taken to indicate that each of these two constellations of variables is

unidimensional. Rather, each of the individual indicators is likely to contain systematic and meaningful variation specific to the domain tapped by that indicator. For example, although there is shared variance between reading and other *g*-factor variables, there is certainly still variance unique to reading above and beyond what is explained by the *g*-factor (i.e., reading is not *only g*). Had we used multiple indicators for each narrow domain within psychopathology and cognitive function, we would expect that hierarchical structures would emerge with *p* and *g* at the apexes of the respective hierarchies. To determine whether the *p-g* association was driven by more domain-specific associations, we conducted separate sensitivity analyses using the phenotypic models for ECLS-B and TXtT.

We used modification indices with a cut-off of 3.84  $\chi^2$  units to expand our structural equation models to allow for pairwise associations between specific domains of psychopathology and cognitive function while simultaneously estimating a higher order *p-g* association. These modification indices index whether there are associations between indicators (e.g., externalizing and math) that were not included in the model, but would significantly improve model fit were they to be added to the model. This allows us to address the question of whether the association between *p* and *g* is, in fact, driven by associations between specific indicators of each of these factors. For ECLS-B, a domain-specific association was added between reading and internalizing ( $r = .04$ , 95% CI [.01, .07],  $p = .017$ ). In this model, the higher order *p-g* association was estimated at  $-.21$  (95% CI  $[-.28, -.16]$ ,  $p < .001$ ), indicating that this association was not simply driven by associations between specific pairs of psychopathology and cognitive function domains. No domain-specific associations were identified for TXtT.

### **Discussion**

Research using adolescent and adult samples has found that genetic liabilities for psychopathology are largely nonspecific and that a general factor of psychopathology is

negatively associated with intelligence. Although previous work in early childhood has examined pairwise associations between measures of specific cognitive abilities and specific dimensions of behavioral and emotional problems, no previous work has examined whether the more general relation between cognitive ability and psychopathology is already apparent in early childhood, and the extent to which this association is attributable to overlapping genetic risks. Using an integrative data analysis approach, we combined data from two American twin samples of early child development to estimate genetic and environmental influences on a general factor of psychopathology (*p*) and their links with a general ability factor (*g*).

We found that internalizing symptomology, externalizing symptomology, attention-deficit/hyperactivity, and self-regulation problems all loaded positively on a general *p*-factor of psychopathology. The *p*- and *g*-factors were themselves negatively correlated in both studies ( $r = -.34$  for TXtT and  $r = -.21$  for ECLS-B). For ECLS-B, associations between *p* and *g* remained unchanged when allowing for a domain-specific association between internalizing and reading, while no domain-specific associations were identified for TXtT. Although model comparisons indicated that genetic and environmental loadings on *p* and *g* could not be constrained across studies, the substantive conclusions were consistent. Behavioral genetic decomposition indicated that the *p*-factor was 63%/46% heritable (in the order: ECLS-B/TXtT), 13%/47% shared environmental, and 24%/7% nonshared environmental. The *g*-factor was 38%/33% heritable, 57%/62% shared environmental, and 5% nonshared environmental in both studies. Genetic variants and shared environmental factors influencing both phenotypes accounted for the majority of the phenotypic association between *p* and *g* (40%/25% and 45%/69%, respectively). Age only significantly moderated non-environmental correlations. The current findings did not suggest that the predominately genetic and shared-environmental basis for the *p*-*g* association

was driven by the upper end of the age distribution under examination. In fact, the point estimate of the genetic correlation was slightly larger at age 3 years than at age 6 years.

### **Strengths and Limitations**

The studies included in the integrative data analysis had complementary strengths and weaknesses. The TXtT sample might have included parents who were more highly educated than the average in the surrounding area. The TXtT study also employed home-based measures of cognitive and psychomotor abilities that were completed by the twins' primary caregivers. Although this does increase the potential biases in cognitive scores due to social desirability or caregivers' subjective impressions, we were careful to employ instruments that asked caregivers to report on performance on concrete tasks, rather than making unanchored subjective judgements. Moreover, the cognitive assessment used in TXtT has been reported to have interrater reliability of .86 (Squires et al., 2009) and to correlate with scores from professionally administered measures of cognitive development at .51-.75 (Schonhaut et al., 2013). On the contrary, the ECLS-B study employed cognitive and psychomotor measures administered by trained examiners during home-visits. A weakness of the ECLS-B study, however, is that the measures of internalizing and externalizing psychopathology were derived from short general-purpose rating systems. In contrast, the TXtT study employed a comprehensive highly validated clinical measure of internalizing and externalizing psychopathology (Ivanova et al., 2010). Despite the differences in measurement and sample ascertainment, overall results were consistent across both samples, thus increasing confidence in results.

### **Conclusions**

Our results suggest that the link between cognitive function and psychopathology emerges much earlier than is implied by several prominent theories in cognitive epidemiology. According to Koenen et al. (2009), the "cognitive reserve" hypothesis holds that early life mental ability

buffers against the effects of neuropathology that emerges across the life course, such that “lower premorbid IQ increases the risk of subsequent disorder.” Similarly, Deary (2008) has suggested several possible mechanisms for the link between early cognitive ability and life course physical health, among which is what he refers to as the “system integrity” hypothesis (Deary, 2012). This hypothesis posits that “mental test scores obtained in youth might be an indicator of a well-put-together system” (Deary, 2008). In reference to the system integrity hypothesis, Deary (2008) speculated that “a well-wired body is more able to respond effectively to environmental insults.” Under these perspectives, the link between early IQ and psychopathology should emerge with age, as the incidence of psychopathology becomes stratified by levels of intelligence over time.

In contrast, the observed association between *p* and *g* in these early childhood samples, prior to the beginning of formal schooling, indicates that the association is not exclusively a consequence of early ability buffering against the incidence of adolescent- and adult-onset psychiatric disease. Nor is the association only an emergent consequence of prolonged academic problems stemming from or leading to behavioral problems in the context of educational institutions. Rather, given the age range of both samples, the link between cognitive deficits and psychiatric symptoms stems, at least in part, from factors that operate in the first years of life. These include genetic variants that affect both phenotypes, and environmental factors that are stratified between families.

**Author Contributions**

K. P. Harden, E. M. Tucker-Drob, and A. K. Cheung developed the study concept and design. Testing and data collection was performed by A. K. Cheung. A. D. Grotzinger, A. K. Cheung, and E. M. Tucker-Drob performed the data analysis and interpretation. All authors helped in drafting the manuscript and providing critical revisions. All authors approved the final version of the manuscript for submission.

**Acknowledgements**

Data from the Texas “Tiny” Twin Project were collected and managed using Research Electronic Data Capture (REDCap) hosted at The University of Texas at Austin (UT-Austin). During the time that this article was prepared, EMTD, KPH, and ADG were supported by National Institutes of Health (NIH) grants R01HD083613, R21HD081437, and R21AA023322. KPH and EMTD are supported by Research Fellowships from the Jacobs Foundation. The Population Research Center at the University of Texas is supported by NIH grant R24HD042849.

### References

- Achenbach, T. M., & Rescorla, L. A. (2000a). *Child Behavior Checklist for Ages 1.5-5*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- Achenbach, T. M., & Rescorla, L. A. (2000b). *Manual for the ASEBA Preschool Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., Duncan, L., ... & Patsopoulos, N. A. (2018). Analysis of shared heritability in common disorders of the brain. *Science*, *360*(6395), eaap8757.
- Barlow, D. H., Allen, L. B., & Choate, M. L. (2004). Toward a unified treatment for emotional disorders. *Behavior therapy*, *35*(2), 205-230.
- Bell, R. Q. (1953). Convergence: An accelerated longitudinal approach. *Child Development*, *24*(2), 145-152.
- Bodovski, K., & Youn, M. (2011). The long term effects of early acquired skills and behaviors on young children's achievement in literacy and mathematics. *Journal of Early Childhood Research*, *9*(1), 4-19.
- Bub, K. L., McCartney, K., & Willett, J. B. (2007). Behavior problem trajectories and first-grade cognitive ability and achievement skills: A latent growth curve analysis. *Journal of Educational Psychology*, *99*(3), 653-670.
- Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P. R., ... & Daly, M. J. (2015). An atlas of genetic correlations across human diseases and traits. *Nature Genetics*, *47*(11), 1236-1241.
- Briley, D. A., & Tucker-Drob, E. M. (2013). Explaining the increasing heritability of cognition over development: A meta-analysis of longitudinal twin and adoption studies. *Psychological Science*, *24*(9), 1704-1713.
- Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., ... & Moffitt, T. E. (2014). The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science*, *2*(2), 119-137.

- Caspi, A., Moffitt, T. E., Newman, D. L., & Silva, P. A. (1996). Behavioral observations at age 3 years predict adult psychiatric disorders: Longitudinal evidence from a birth cohort. *Archives of General Psychiatry*, *53*(11), 1033-1039.
- Caspi, A., & Moffitt, T. E. (2018). All for one and one for all: Mental disorders in one dimension. *American Journal of Psychiatry*.
- Cheung, A. K., Harden, K. P., & Tucker-Drob, E. M. (2015). From specialist to generalist: Developmental transformations in the genetic structure of early child abilities. *Developmental Psychobiology*, *57*(5), 566-583.
- Cheung, A. K., Harden, K. P., & Tucker-Drob, E. M. (2016). Multivariate behavioral genetic analysis of parenting in early childhood. *Parenting*, *16*(4), 257-283.
- Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: the simultaneous analysis of multiple data sets. *Psychological Methods*, *14*(2), 81-100.
- Davis, O. S., Haworth, C. M., & Plomin, R. (2009). Dramatic increase in heritability of cognitive development from early to middle childhood: An 8-year longitudinal study of 8,700 pairs of twins. *Psychological Science*, *20*(10), 1301-1308.
- Deary, I. J. (2008). Why do intelligent people live longer? *Nature*, *456*(7219), 175-176.
- Deary, I. J. (2012). Looking for 'system integrity' in cognitive epidemiology. *Gerontology*, *58*(6), 545-553.
- Ebejer, J. L., Coventry, W. L., Byrne, B., Willcutt, E. G., Olson, R. K., Corley, R., & Samuelsson, S. (2010). Genetic and environmental influences on inattention, hyperactivity-impulsivity, and reading: Kindergarten to grade 2. *Scientific Studies of Reading*, *14*(4), 293-316.
- Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry*, *47*(3), 313-337.
- Forget-Dubois, N., Pérusse, D., Turecki, G., Girard, A., Billette, J., Rouleau, G., ... & Tremblay, R. E. (2003). Diagnosing zygosity in infant twins: Physical similarity, genotyping, and chorionicity. *Twin Research*, *6*(6), 479-485.
- Gollenberg, A., Lynch, C. D., Jackson, L. W., McGuinness, B. M., & Msall, M. E. (2010). Current validity of the parent-completed Ages and Stages Questionnaires, with the

- Bayley Scales of Infant Development II in a low-risk sample. *Child: Care, Health and Development*, 36(4), 485-490.
- Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., ... & Harden, K. P. (2018). Genomic SEM provides insights into the multivariate genetic architecture of complex traits. *bioRxiv*, 305029.
- Harden, K. P., Kretsch, N., Tackett, J. L., & Tucker-Drob, E. M. (2014). Genetic and environmental influences on testosterone levels in adolescents: Evidence for sex differences. *Developmental Psychobiology*, 56(6), 1278-1289.
- Harden, K. P., Tucker-Drob, E. M., & Tackett, J. L. (2013). The Texas Twin Project. *Twin Research and Human Genetics*, 16(1), 385-390.
- Heath, A. C., Nyholt, D. R., Neuman, R., Madden, P. A. F., Bucholz, K. K., Todd, R. D., ... & Martin, N. G. (2003). Zygosity diagnosis in the absence of genotypic data: An approach using latent class analysis. *Twin Research*, 6(1), 22-26.
- Hollingshead, A.B. (1975). *Four factor index of social status*. New Haven, CT: Yale University Press.
- Ivanova, M. Y., Achenbach, T. M., Rescorla, L. A., Harder, V. S., Ang, R. P., Bilenberg, N., ... & Dobrean, A. (2010). Preschool psychopathology reported by parents in 23 societies: testing the seven-syndrome model of the child behavior checklist for ages 1.5–5. *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(12), 1215-1224.
- Keenan, K., Shaw, D., Delliquadri, E., Giovannelli, J., & Walsh, B. (1998). Evidence for the continuity of early problem behaviors: Application of a developmental model. *Journal of Abnormal Child Psychology*, 26(6), 441-452.
- Koenen, K. C., Caspi, A., Moffitt, T. E., Rijdsdijk, F., & Taylor, A. (2006). Genetic influences on the overlap between low IQ and antisocial behavior in young children. *Journal of Abnormal Psychology*, 115(4), 787-797.
- Koenen, K. C., Moffitt, T. E., Roberts, A. L., Martin, L. T., Kubzansky, L., Harrington, H., ... & Caspi, A. (2009). Childhood IQ and adult mental disorders: A test of the cognitive reserve hypothesis. *American Journal of Psychiatry*, 166(1), 50-57.
- Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., ... & Eaton, N. R. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A

- dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, 126(4), 454.
- Koot, H. M., van den Oord, E. J., Verhulst, F. C., & Boomsma, D. I. (1997). Behavioral and emotional problems in young preschoolers: Cross-cultural testing of the validity of the Child Behavior Checklist/2-3. *Journal of Abnormal Child Psychology*, 25(3), 183-196.
- Kuntsi, J., Eley, T. C., Taylor, A., Hughes, C., Asherson, P., Caspi, A., & Moffitt, T. E. (2004). Co-occurrence of ADHD and low IQ has genetic origins. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 124(1), 41-47.
- Lahey, B. B., Applegate, B., Hakes, J. K., Zald, D. H., Hariri, A. R., & Rathouz, P. J. (2012). Is there a general factor of prevalent psychopathology during adulthood? *Journal of Abnormal Psychology*, 121(4), 971-977.
- Lahey, B. B., Krueger, R. F., Rathouz, P. J., Waldman, I. D., & Zald, D. H. (2016). A hierarchical causal taxonomy of psychopathology across the life span. *Psychological Bulletin*, 143(2), 142-186.
- Lahey, B. B., Rathouz, P. J., Keenan, K., Stepp, S. D., Loeber, R., & Hipwell, A. E. (2015). Criterion validity of the general factor of psychopathology in a prospective study of girls. *Journal of Child Psychology and Psychiatry*, 56(4), 415-422.
- Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods*, 4(2), 192-211.
- Martel, M. M., Pan, P. M., Hoffmann, M. S., Gadelha, A., do Rosário, M. C., Mari, J. J., ... & Rohde, L. A. (2017). A general psychopathology factor (P factor) in children: Structural model analysis and external validation through familial risk and child global executive function. *Journal of Abnormal Psychology*, 126(1), 137-148.
- Mesman, J., & Koot, H. M. (2001). Early preschool predictors of preadolescent internalizing and externalizing DSM-IV diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(9), 1029-1036.

- Muthén, L. K., & Muthén, B. O. (2010). *Mplus User's Guide, 6th edn.* Muthén & Muthén, Los Angeles, CA.
- Najarian, M., Snow, K., Lennon, J., & Kinsey, S. (2010). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), preschool– kindergarten 2007 psychometric report (NCES 2010-009)*. Washington, DC: U.S. Department of Education, National Center for Education Statistics, Institute of Education Sciences.
- Neumann, A., Pappa, I., Lahey, B. B., Verhulst, F. C., Medina-Gomez, C., Jaddoe, V. W., ... & Tiemeier, H. (2016). Single nucleotide polymorphism heritability of a general psychopathology factor in children. *Journal of the American Academy of Child & Adolescent Psychiatry, 55*(12), 1038-1045.
- Paloyelis, Y., Rijdsdijk, F., Wood, A. C., Asherson, P., & Kuntsi, J. (2010). The genetic association between ADHD symptoms and reading difficulties: The role of inattentiveness and IQ. *Journal of Abnormal Child Psychology, 38*(8), 1083-1095.
- Plomin, R., Haworth, C. M., & Davis, O. S. (2009). Common disorders are quantitative traits. *Nature Reviews Genetics, 10*(12), 872-878.
- Price, T. S., Freeman, B., Craig, I., Petrill, S. A., Ebersole, L., & Plomin, R. (2000). Infant zygosity can be assigned by parental report questionnaire data. *Twin Research, 3*(3), 129-133.
- Purcell, S. (2002). Variance components models for gene–environment interaction in twin analysis. *Twin Research and Human Genetics, 5*(6), 554-571.
- Riccio, C. (1995). Book review: preschool and kindergarten behavior scales. *Journal of Psychoeducational Assessment, 13*(2), 194-196.
- Rietveld, M. J. H., van der Valk, J. C., Bongers, I. L., Stroet, T. M., Slagboom, P. E., & Boomsma, D. I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Research, 3*(3), 134-141.
- Schonhaut, L., Armijo, I., Schönstedt, M., Alvarez, J., & Cordero, M. (2013). Validity of the Ages and Stages Questionnaires in term and preterm infants. *Pediatrics, 131*, e1468-e1474.

- Simard, M., Luu, T. M., & Gosselin, J. (2012). Concurrent validity of Ages and Stages Questionnaires in preterm infants. *Pediatrics*, *130*, e108-e114.
- Smoller, J. W., Andreassen, O. A., Edenberg, H. J., Faraone, S. V., Glatt, S. J., & Kendler, K. S. (2018). Psychiatric genetics and the structure of psychopathology. *Molecular Psychiatry*, *1-12*.
- Snow, K., Derecho, A., Wheelless, S., Lennon, J., Rosen, J., Rogers, J., ... & Einaudi, P. (2009). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), kindergarten 2006 and 2007 data file user's manual (2010-010)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, US Department of Education.
- Spearman, C. (1927). *The abilities of man*. London: Macmillan
- Squires, J., & Bricker, D. (2009). *Ages & Stages Questionnaires, Third Edition (ASQ-3)*. Baltimore, MD: Paul H. Brookes Publishing Co.
- Squires, J., Bricker, D., & Twombly, E. (2003). *The ASQ:SE User's Guide for the Ages & Stages Questionnaires: Social-Emotional. A Parent-Completed, Child-Monitoring System for Social-Emotional Behaviors*. Baltimore, MD: Paul H. Brookes Publishing Co.
- Tackett, J. L., Lahey, B. B., Van Hulle, C., Waldman, I., Krueger, R. F., & Rathouz, P. J. (2013). Common genetic influences on negative emotionality and a general psychopathology factor in childhood and adolescence. *Journal of Abnormal Psychology*, *122*(4), 1142-1153.
- Tucker-Drob, E. M., Rhemtulla, M., Harden, K. P., Turkheimer, E., & Fask, D. (2011). Emergence of a gene  $\times$  socioeconomic status interaction on infant mental ability between 10 months and 2 years. *Psychological Science*, *22*(1), 125-133.
- Van Horn, M. L., Atkins-Burnett, S., Karlin, E., Ramey, S. L., & Snyder, S. (2007). Parent ratings of children's social skills: Longitudinal psychometric analyses of the Social Skills Rating System. *School Psychology Quarterly*, *22*(2), 162-199.
- Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., ... & Burstein, R. (2013). Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *The Lancet*, *382*(9904), 1575-1586.

Yu, L., Hey, E., Doyle, L. W., Farrell, B., Spark, P., Altman, D. G., & Duley, L. (2007).

Evaluation of the Ages and Stages Questionnaires in identifying children with neurosensory disability in the Magpie Trial follow-up study. *Acta Paediatrica*, *96*(12), 1803-1808.

Zhou, Q., Main, A., & Wang, Y. (2010). The relations of temperamental effortful control and anger/frustration to Chinese children's academic achievement and social adjustment: A longitudinal study. *Journal of Educational Psychology*, *102*(1), 180-196.

**Footnotes**

<sup>1</sup>The sample size is rounded to the nearest 50 in accordance with ECLS-B regulations.

<sup>2</sup>Same-sex twins that were classified as DZ, but had parents who indicated there was a medical reason for their dissimilarity, were excluded from analyses. Less than 50 twin pairs were excluded using this criterion.

Table 1

*Estimates for Quantitative Genetic Models that Estimated Pooled or Study-Specific Parameters*

<i>Parameter</i>	<b>Model 1</b>		<b>Model 2</b>		<b>Model 3a</b>		<b>Model 3b</b>	
	<i>ECLS-B</i>	<i>TXtT</i>	<i>ECLS-B</i>	<i>TXtT</i>	<i>ECLS-B</i>	<i>TXtT</i>	<i>ECLS-B</i>	<i>TXtT</i>
<b><i>p</i>-Factor</b>								
Genetic Effect ( $A_p$ )	.77 (.08)***	.70 (.11)***	.80 (.09)***	.68 (.11)***	.77 (.07)***		.81 (.09)***	.68 (.10)***
Shared Environment ( $C_p$ )	.39 (.12)***	.68 (.10)***	.35 (.15)*	.69 (.09)***	.51 (.09)***		.33 (.18)	.69 (.09)***
Non-shared Environment ( $E_p$ )	.51 (.06)***	.23 (.10)**	.49 (.06)***	.26 (.08)**	.37 (.05)***		.49 (.06)***	.26 (.09)***
<b><i>g</i>-Factor</b>								
Genetic Effect ( $A_g$ )	.62 (.05)***	.58 (.10)***	.62 (.05)***	.57 (.10)***	.61 (.05)***	.58 (.10)***		.60 (.05)***
Shared Environment ( $C_g$ )	.76 (.04)***	.79 (.08)***	.76 (.04)***	.79 (.07)***	.76 (.04)***	.79 (.08)***		.78 (.03)***
Non-shared Environment ( $E_g$ )	.22 (.03)***	.21 (.04)***	.21 (.03)***	.22 (.04)***	.21 (.03)***	.22 (.04)***		.21 (.02)***
<b><i>ACE</i> Correlations</b>								
Genetic ( $r_A$ )	-.16 (.10)	-.29 (.19)		-.19 (.08)*		-.21 (.09)*		-.18 (.08)*
Shared Environment ( $r_C$ )	-.37 (.18)*	-.38 (.19)*		-.39 (.17)*		-.31 (.12)*		-.42 (.20)*
Non-shared Environment ( $r_E$ )	-.30 (.11)**	-.53 (.36)		-.33 (.12)**		-.42 (.14)**		-.33 (.12)**

*Note.* Parameters reported are standardized path estimates and standard errors are given in parentheses. Model 1 estimated study-specific *p* and *g* ACE estimates and correlations. Model 2 estimated study-specific ACE estimates and study-averaged ACE correlations. Model 3a estimated study-averaged ACE estimates for *p* and cross-trait ACE correlations. Model 3b estimated study-averaged ACE estimates for *g* and cross-trait ACE correlations. \*\*\*significantly different than zero at  $p < .001$ ; \*\*  $p < .01$ ; \*  $p < .05$ .

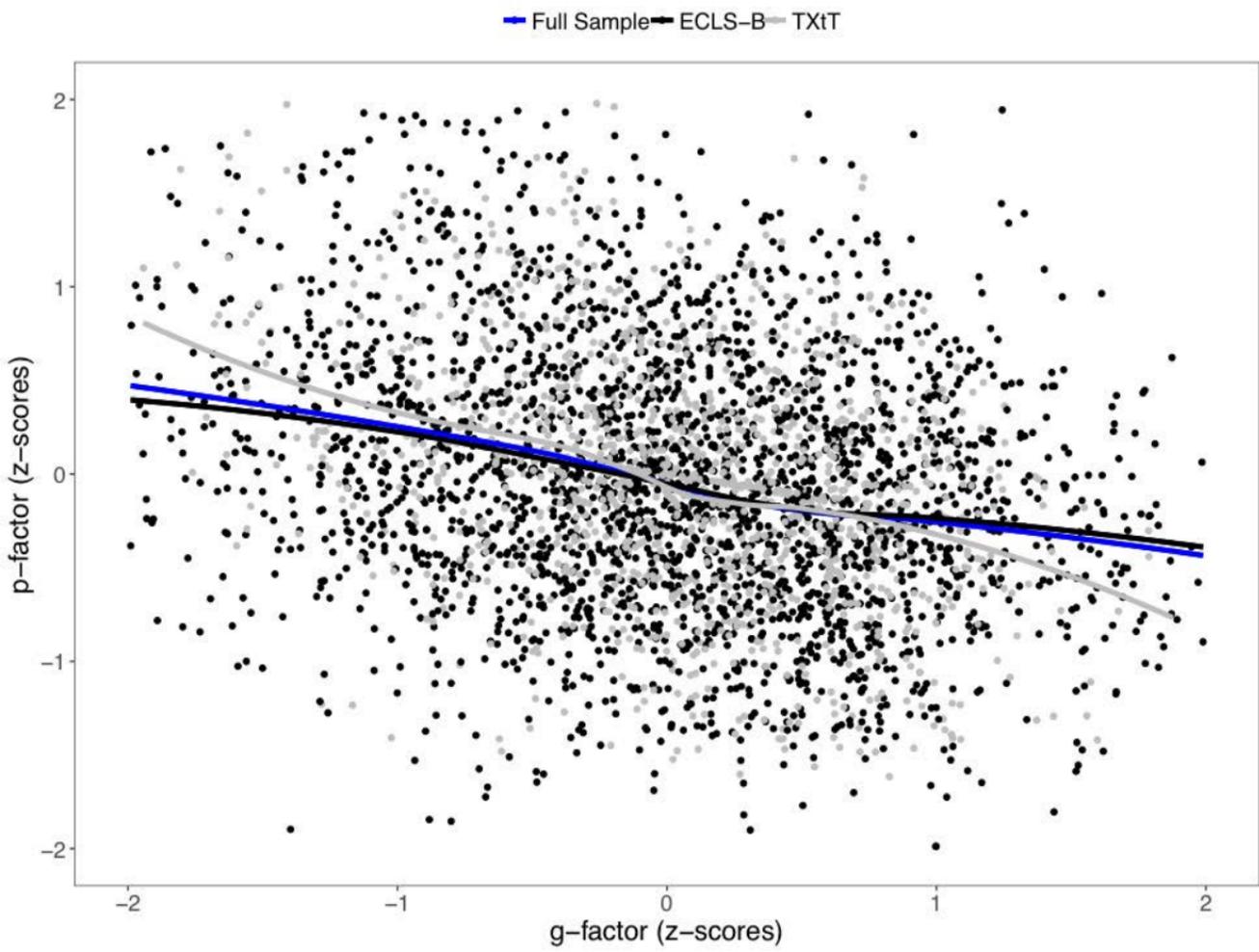
**Figure Captions**

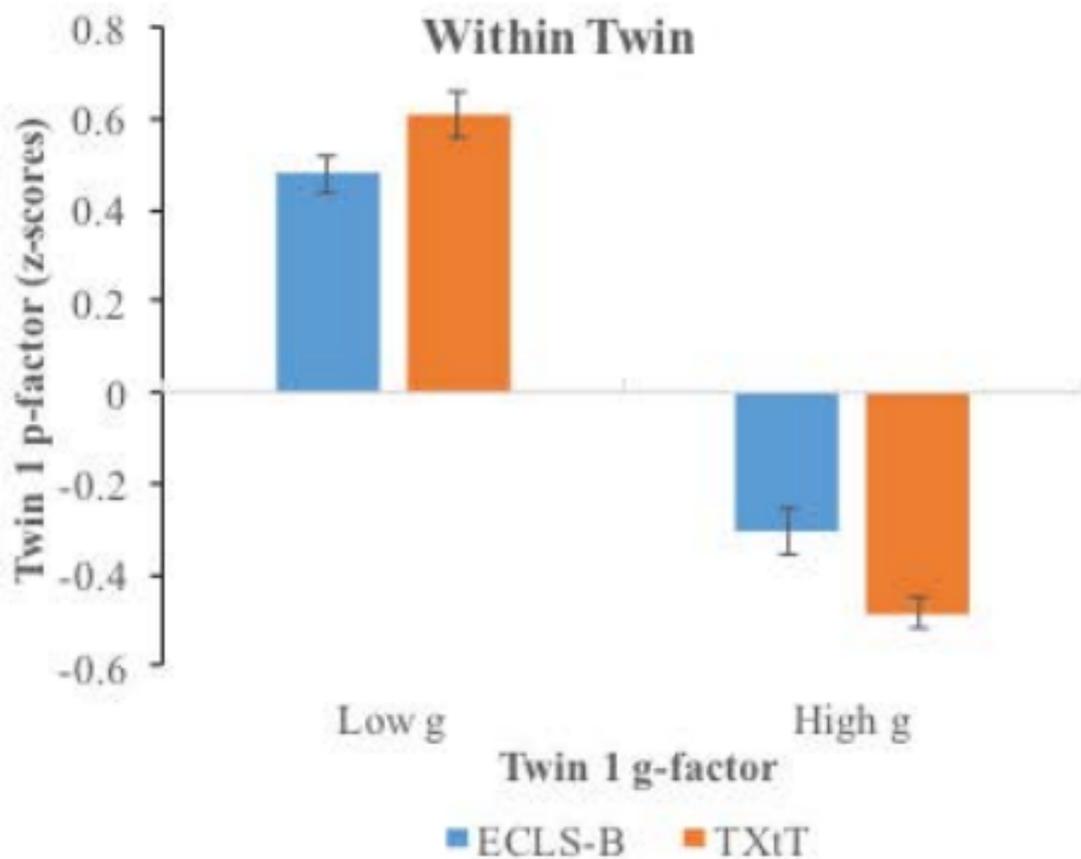
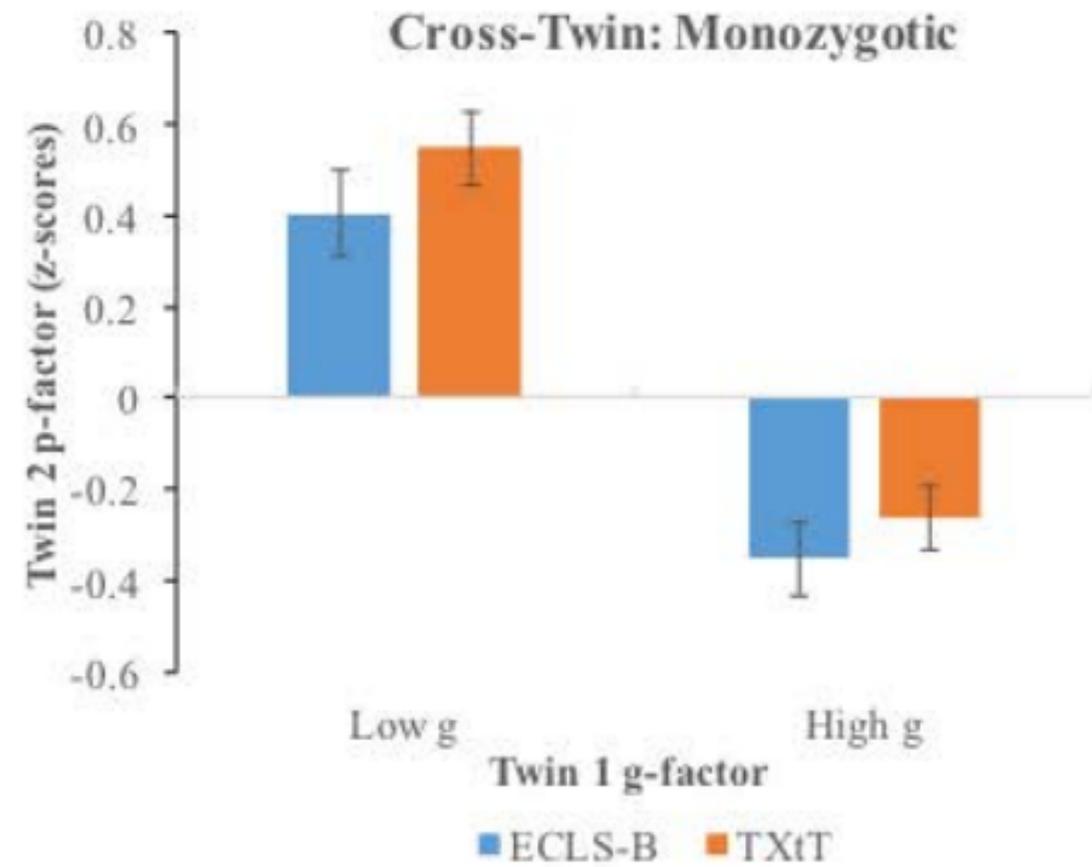
*Figure 1.* Figure depicts standardized *p*-factor scores across standardized *g*-factor scores for ECLS-B and TXtT participants. Plotted lines (ECLS-B = black, TXtT = gray, full sample = blue) were fit using a locally weighted scatterplot (LOESS) regression function in *R*.

*Figure 2.* Figure depicts mean *p*-factor differences between participants with low ( $\leq -1 SD$ ) and high ( $\geq +1 SD$ ) *g*-factor scores. Panel A depicts differences within the individual. Panels B and C depict the average of Twin 2 *p*-factor scores as a function of Twin 1 *g*-factor scores for MZ and DZ twins, respectively. Both factors were standardized prior to calculating averages. Error bars represent  $\pm 1 SE$  of the sample mean.

*Figure 3.* Standardized parameter estimates from Model 2 that constrained *ACE* correlations, but not *ACE* loadings, of the *p*- and *g*-factor to be equal across samples. For study-specific parameters, estimates are presented in the order: ECLS-B/Texas Tiny Twins. All indicator intercepts, unique *ACE* loadings, and indicator residuals were freely estimated within each sample. Indicator level estimates are standardized with respect to both *ACE* parameters and *p/g*-factor loadings. Personal-social and Reading indicators were present only in Texas Tiny Twins and ECLS-B, respectively, and NAs are reported for non-existent paths. All indicators were residualized for the effects of age, age<sup>2</sup>, sex, age  $\times$  sex, and age<sup>2</sup>  $\times$  sex and then standardized prior to

being entered into the model. Only one twin per pair is depicted for ease of presentation. *A*, additive genetic; *C*, shared environment; *E*, non-shared environment. Parameters depicted in bold are significant at  $p < .05$



**A****B****C**