



Regular Article

Using an adoption design to test genetically based differences in risk for child behavior problems in response to home environmental influences

Robyn A. Cree¹ , Chang Liu², Ralitza Gueorguieva¹, Jenae M. Neiderhiser³, Leslie D. Leve⁴, Christian M. Connell³, Daniel S. Shaw⁵, Misaki N. Natsuaki⁶ , Jody M. Ganiban², Charles Beekman⁷, Megan V. Smith^{1,8} and David Reiss⁹

¹Chronic Disease Epidemiology, Yale School of Public Health, New Haven, CT, USA; ²Department of Psychology, The George Washington University, Washington, DC, USA; ³Department of Psychology, Pennsylvania State University, State College, PA, USA; ⁴Prevention Science Institute, University of Oregon, Eugene, OR, USA; ⁵Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA; ⁶Department of Psychology, University of California, Riverside, CA, USA; ⁷Department of Psychology, Arizona State University, Phoenix, AZ, USA; ⁸Department of Psychiatry, Yale University, New Haven, CT, USA and ⁹ Child Study Center, Yale University, New Haven, CT, USA

Abstract

Differential susceptibility theory (DST) posits that individuals differ in their developmental plasticity: some children are highly responsive to both environmental adversity and support, while others are less affected. According to this theory, “plasticity” genes that confer risk for psychopathology in adverse environments may promote superior functioning in supportive environments. We tested DST using a broad measure of child genetic liability (based on birth parent psychopathology), adoptive home environmental variables (e.g., marital warmth, parenting stress, and internalizing symptoms), and measures of child externalizing problems ($n = 337$) and social competence ($n = 330$) in 54-month-old adopted children from the Early Growth and Development Study. This adoption design is useful for examining DST because children are placed at birth or shortly thereafter with nongenetically related adoptive parents, naturally disentangling heritable and postnatal environmental effects. We conducted a series of multivariable regression analyses that included Gene \times Environment interaction terms and found little evidence of DST; rather, interactions varied depending on the environmental factor of interest, in both significance and shape. Our mixed findings suggest further investigation of DST is warranted before tailoring screening and intervention recommendations to children based on their genetic liability or “sensitivity.”

Keywords: adoption design, differential susceptibility, externalizing, psychopathology factor, social competence

(Received 6 September 2018; revised 8 March 2020; accepted 15 March 2020)

Individual differences in response to environmental exposures are well recognized in the study of behavioral problems in young children. Diathesis–stress, or the idea that some individuals are more negatively affected by exposure to stress than others, has been useful in explaining why certain children are resilient while others are susceptible to environmental stressors (e.g., Monroe & Simons, 1991; Rosenthal et al., 1968). Differential susceptibility theory (DST) (Belsky, 2005) builds on diathesis–stress by positing that individuals with “plasticity” genes will do poorly under stressful circumstances, yet thrive under promotive environmental situations, as compared to those without these genes who will function relatively consistently regardless of environmental influence (e.g., Belsky, Bakermans-Kranenburg, & van IJzendoorn, 2007; Belsky

& Pluess, 2009; Boyce & Ellis, 2005; Reiss, Leve, & Neiderhiser, 2013).

Kendler and Eaves (1986) first conceptualized “genetic control of sensitivity of the environment,” noting that individuals, once assumed to carry genotypes that place them at risk for psychopathology, may also be more sensitive to their environments, both positively and negatively. Wahlberg et al. (1997) first tested the theory using data from the Finnish Adoptive Family Study of Schizophrenia. They found that children with increased genetic risk, as measured by whether the child’s biological mother had schizophrenia, were more likely to have symptoms of thought disorder in response to their adoptive mothers’ communication deviance, but were less likely to have symptoms of thought disorder in response to favorable rearing conditions, as compared to children with biological mothers who did not have schizophrenia.

Since the reformulation of DST by Belsky et al., correlational studies support DST in the context of parenting and child outcomes (for a review, see Belsky & Pluess, 2013a). Similarly, a meta-analysis of randomized controlled trials (RCTs) comparing the behavior of children with and without sensitive variants of

Author for correspondence: Robyn Cree, 60 College Street, New Haven, CT 06511, USA. Email: robyncree@gmail.com.

Cite this article: Cree RA *et al.* (2021). Using an adoption design to test genetically based differences in risk for child behavior problems in response to home environmental influences. *Development and Psychopathology* 33, 1229–1247. <https://doi.org/10.1017/S0954579420000450>

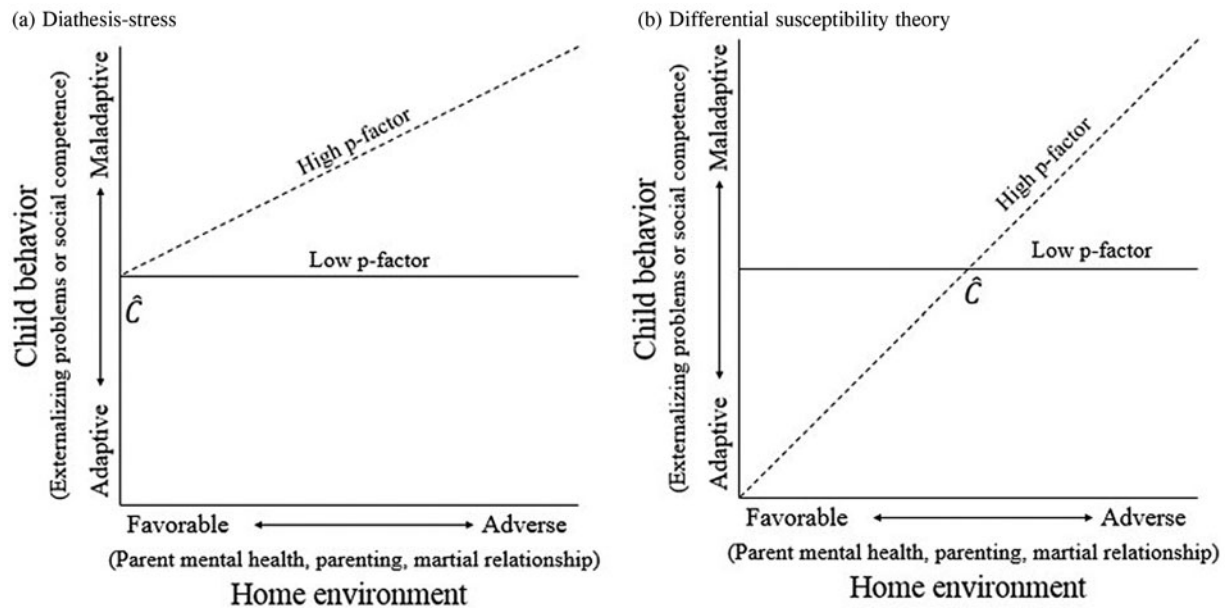


Figure 1. Graphical representation of hypothesized results under differential susceptibility theory (DST) versus diathesis-stress. *Note:* Expected results under (a) diathesis-stress and (b) differential susceptibility with cross-over point C. Figures modeled after Reiss *et al.* 2013. p-factor = general psychopathology factor.

stress-related polymorphisms after parenting interventions found an overall moderate effect size with results consistent with DST (van Ijzendoorn & Bakermans-Kranenburg, 2015); however, it is important to note that the participants in the studies included in this review were genotyped after the trial completion with consequent compromises in subject retention and benefits of randomization and intention-to-treat analyses. Further, a recent study using a genome-wide approach to calculate polygenic scores in a large sample of twins also found exploratory evidence of DST in relation to a parenting intervention and child anxiety (Eley *et al.*, 2017). These results are in accord with an earlier review including phenotypic (*i.e.*, child behavior) and genetic data from longitudinal cohort studies, cross-sectional data, and RCTs (Pluess & Belsky, 2010).

However, not all forms of gene \times environment interactions ($G \times E$ s) implicated in psychological outcomes are consistent with DST (*e.g.*, Leve *et al.*, 2009), nor are they expected to be (Belsky *et al.*, 2007; Reiss *et al.*, 2013; Rutter, Moffitt, & Caspi, 2006). Further, concerns with replicability, power, and publication bias surrounding $G \times E$ findings, in general, have been noted (see Dick *et al.*, 2015; Duncan & Keller, 2011). Therefore, replication with longitudinal samples that are specifically designed to explore $G \times E$ s, such as adoption studies (*e.g.*, where parents are genetically unrelated to their children, enabling us to disentangle heritable and prenatal influences, on the one hand, from postnatal rearing environmental influences, on the other), augment the existing molecular genetic work using statistical tests to examine DST.

Testing differential susceptibility: Necessary design elements

Several threats to the validity of results must be considered when designing studies to test for DST.

First, studies testing for $G \times E$ s should first assess whether the genetic factor of interest is correlated with the environmental variable in question to rule out confounding by gene-environment correlation (rGE) (Belsky *et al.*, 2007). For example, passive rGE

occurs when associations between parental behavior and child functioning are due to shared genetic influences on both phenotypes, while evocative rGE occurs when a child who is genetically predisposed to difficult behavior might evoke a negative response from the parent (Ge *et al.*, 1996). Both can masquerade as a significant $G \times E$ and should either be accounted for or ruled out. A major benefit of adoption design is that it naturally eliminates confounding by passive rGE and, as a consequence, observed associations between parenting and child outcomes are due to environmental effects, either child-to-parent or parent-to-child (Plomin, DeFries, & Loehlin, 1977; Scarr & McCartney, 1983). However, the validity of results from adoption studies rests on the assumption that children are not selectively placed with adoptive families based on heritable shared characteristics (Reiss *et al.*, 2013), which would inflate both environmental and genetic estimates through a process similar to passive rGE (Scarr & Weinberg, 1977, 1983).

Second, restricted ranges in variation in environmental exposures and child outcomes can threaten the validity of DST results (Belsky *et al.*, 2007). Many prior studies examining DST have defined promotive environments as a lack of detrimental environmental stimuli (*i.e.*, lack of environmental stress) and positive child outcomes as the absence of psychopathology rather than operationalizing these dimensions to reflect promoting contexts and enhanced child outcomes (Hankin *et al.*, 2011). Sufficient variation is needed so that promotive and detrimental environmental exposures, as well as adaptive and maladaptive outcomes are represented, as shown in the four quadrants of Figure 1b. For example, if a promotive environment is measured by the absence of a risky condition, the interaction could be misinterpreted as consistent with diathesis-stress simply because a truly supportive environmental condition (*i.e.*, the area represented in the left half of Figure 1b) was not assessed.

Third, adequate statistical testing of significant interactions is necessary (Widaman *et al.*, 2012). Significant $G \times E$ s must be disordinal and similar in shape to Figure 1b, rather than ordinal, which is in line with the diathesis-stress model depicted in Figure 1a. Further, this shape should be statistically “confirmed”

with tests of the cross-over point of the interaction and its confidence interval (CI) (Widaman et al., 2012). To support DST, both must fall within the observable range of the environmental variable. Although a cross-over point can be calculated using the exploratory tests mentioned above (Aiken & West, 1991), re-parameterized equations are necessary to obtain the associated CI. Researchers should test four statistical models—"strong" and "weak" differential susceptibility, and "strong" and "weak" diathesis–stress—and then select the best fitting model (Belsky, Pluess, & Widaman, 2013). For a full discussion and example of this technique, see Belsky et al. (2013). Finally, results should be consistent across a variety of different promotive and detrimental environmental exposures and behavioral outcomes (Belsky et al., 2007).

Relevant early environmental exposures

Home environmental exposures are experienced as early as prenatally and during infancy onward, shaping how a child will react to social situations throughout the lifespan. Children consistently exposed to negative environments in early development are more likely to view benign social situations as hostile and may react in socially unacceptable ways; conversely, children in healthy family emotional climates are more likely to develop healthier social information processing and, as a result, have fewer externalizing problems and greater social competence (Dodge & Crick, 1990; Schultz & Shaw, 2003). Self-reported parental emotional states during early childhood, such as over-reactivity (e.g., Lipscomb et al., 2012) and daily hassles (e.g., Crnic & Greenberg, 1990; Deater-Deckard, 1998; Respler-Herman, Mowder, Yasik, & Renee Shamah, 2012), are often associated with poorer child psychological functioning, while parental self-efficacy is generally considered a promotive environmental exposure (for a review, see Jones & Prinz, 2005). For example, parenting daily hassles are associated with higher observational ratings of behavior problems in preschool children (Crnic, Gaze, & Hoffman, 2005) and teacher-rated emotion regulation in toddlers (Mathis & Bierman, 2015). Marital functioning has also been associated with child psychological functioning that may be moderated by a child's genetic risk (Fearon et al., 2015; O'Connor, Caspi, Defries, & Plomin, 2003; Rhoades et al., 2011). Further, parents who report greater conflict in their marriages engage in more negative parenting practices with their infants (e.g., Owen & Cox, 1997) and toddlers (as shown in the current sample; e.g., Stover et al., 2012). Finally, children exposed to parent internalizing symptoms are more likely to experience poorer parenting practices (Goodman, 2007; Goodman et al., 2011; Lovejoy, Graczyk, O'Hare, & Neuman, 2000; Natsuaki et al., 2014; Wilson & Durbin, 2010), dysregulated parent–child interactions (Roben et al., 2015), and difficulties with emotion regulation (Feng et al., 2008), contributing to problems with their own psychological functioning (Bagner, Pettit, Lewinsohn, Seeley, & Jaccard, 2013; Goodman et al., 2011) and the development of externalizing problems (Gartstein & Fagot, 2003; Goodman & Gotlib, 1999; Luoma et al., 2001). However, it should be noted that parental influences on child outcomes found in previous studies using biologically related families cannot be fully attributed to the environmental influences of rearing because, in biologically related families, the associations between parenting behaviors and child outcomes could be explained by shared heritable influences between parents and children, prenatal and postnatal environmental influences, or passive *r*GE influences (Ge et al., 1996; Scarr & McCartney, 1983).

Evidence for a general genetic liability for psychopathology

Caspi et al. (2014) and Lahey et al. (2012) proposed a method of estimating an indicator of a general genetic risk for psychopathology that takes into account genetic influences common to a broad range of mental disorders. Both groups of researchers noted that clinical psychological disorders have high comorbidities, with roughly half of all individuals with at least one disorder also having a second disorder, and half of those with two disorders having a third disorder (Newman, Moffitt, Caspi, & Silva, 1998). Two major latent liabilities help explain these high comorbidities: underlying tendencies to develop internalizing (e.g., major depressive disorder) or externalizing (e.g., antisocial personality disorders) behaviors (Kendler, Prescott, Myers, & Neale, 2003; Krueger & Markon, 2006). Although there is wide consensus on the existence of these two dimensions, work has failed to explain the high correlation between latent internalizing and externalizing dimensions, estimated to be about 0.5 (Wright et al., 2013). Caspi et al. (2014) and Lahey et al. (2012) proposed that this high correlation suggests the possibility of a latent general psychopathology factor or vulnerability.

Similar to the commonly known "g-factor" for intelligence, Caspi et al. (2014) and Lahey et al. (2012) suggest that the psychopathology factor (p-factor) represents a general, underlying genetic propensity to develop psychopathology, and is more parsimonious than the widely used two-factor structure models. Recently, evidence for the p-factor has been replicated in a number of studies (Laceulle, Vollebergh, & Ormel, 2015; Patalay et al., 2015; Tackett et al., 2013), and the structure of the p-factor remains consistent across childhood and adolescence (Murray, Eisner, & Ribeaud, 2016). In the bifactor measurement model, multiple specific factors (e.g., "externalizing" or "internalizing" factors) are identified. Each specific factor influences a subset of diagnostic symptoms of psychopathologies above and the beyond influences from the p-factor (Caspi et al., 2014). Previous work indicates that after including the p-factor in the model, the associations between the specific factors (externalizing or internalizing factors) and early environmental risks (e.g., childhood maltreatment), family psychiatric history, compromised cognitive functioning (e.g., low IQ), and life impairments significantly decreased (Caspi et al., 2014). However, the associations between externalizing (or internalizing) factors and gendered personality styles remained the same or grew stronger (Caspi et al., 2014). Taken together, this evidence suggests that the externalizing and internalizing factors generated in the bifactor model may represent gendered personality styles/traits rather than psychopathology. Thus, in the current work, we used the p-factor as our measure of genetic liability and controlled for specific externalizing and internalizing factors in our models.

Neumann et al. (2016) showed that the p-factor has a significant single nucleotide polymorphism (SNP) heritability of 38% ($SE = 16\%$), although it should be noted that these analyses were based on a large number of measures by multiple raters assessed in 6- to 8-year-old European children, and the results would not necessarily extend to samples assessed using fewer measures or at different ages. Nevertheless, significant SNP heritability may explain why many psychological disorders appear to share many of the same genetic influences (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Kendler, Neale, Kessler, Heath, & Eaves, 1992). Considering that children share 50% of their genes with each biological parent and that links between children and the p-factor score for birth parents are

genetically based (e.g., Lahey *et al.*, 2015; Patalay *et al.*, 2015; Selzam, Coleman, Caspi, Moffitt, & Plomin, 2018; Waldman, Poore, van Hulle, Rathouz, & Lahey, 2016), children with biological parents who score high on the p-factor (i.e., an index of risk alleles) would presumably be more likely to demonstrate psychopathology themselves. However, using the p-factor as a measure of genetic risk could dilute tests of DST if a subset of children categorized in the low-risk group had birth parents with genetic susceptibility who were raised in environments that protected against expression of their risk. Nevertheless, testing G×Es using the p-factor as a measure of genetic liability—a method that takes into account known and unknown genetic influences related to psychopathology—affords a new opportunity to test DST.

Current study

We used data from the Early Growth and Development Study (EGDS), a prospective adoption study (Leve *et al.*, 2019), to perform a stringent test of DST in relation to two important public health outcomes at 54 months of age: childhood externalizing problems (including aggression, oppositional behavior, and hyperactivity (Liu, 2004)) and the child's social competence (including interpersonal skills, peer relations, and social functioning (Semrud-Clikeman, 2007)). This early developmental period is particularly salient for examining differential susceptibility as evidence for this theory has already been found at this age for both of our outcomes of interest using single genes (e.g., dopamine receptor D4 gene (DRD4) 7-repeat; Belsky *et al.*, 2013). We used a general latent psychopathology factor (p-factor; Caspi *et al.*, 2014; Lahey *et al.*, 2012) in birth mothers as a measure of genetic susceptibility to provide a new opportunity to test DST using the most stringent statistical tests available (Widaman *et al.*, 2012).

We considered three distinct rearing environment variables, each independently associated with child externalizing problems or social competence: (a) parenting stress (Crnic *et al.*, 2005; Fearon *et al.*, 2015; Jones & Prinz, 2005; Mathis & Bierman, 2015), (b) parent internalizing symptoms (Bagner *et al.*, 2013; Goodman *et al.*, 2011), and (c) parent marital warmth or hostility (McCoy, Cummings, & Davies, 2009). We included both promotive (i.e., high self-efficacy as a parent and high marital warmth) and adverse (over-reactivity, parenting hassles, parental internalizing symptoms, and marital hostility) environmental factors and positive (e.g., social competence) and negative (e.g., externalizing problems) child outcomes to allow for better exploration of DST. Finally, we measured child outcomes using father ratings of behavior with mother ratings of environmental exposures to control for shared method variance and to ensure the observed correlations were conservative estimates of true relationships rather than an artifact of variables being measured using identical sources. We chose to use maternal home environment variables because previous work using the EGDS sample has reported stronger environmental effects on children at the age we studied for mothers than fathers during infancy and toddlerhood (e.g., Natsuaki *et al.*, 2014).

We hypothesized that the relations between the three home environmental variables (parenting behaviors and emotions, parent internalizing symptoms, and marital warmth/hostility) and later child externalizing behavior and social competence would be moderated by a genetic propensity for mental health problems. Consistent with DST, we further hypothesized that children with a genetic propensity for psychopathology would have the highest

levels of externalizing problems and the lowest levels of social competence when exposed to negative home environmental factors and the highest level of functioning when exposed to promotive home environmental factors, as compared to children without the genetic propensity.

Method

Participants

We used data from a sample of 561 adoptive families, including children, birth mothers, adoptive mothers, and adoptive fathers from the EGDS (Leve *et al.*, 2012). The birth mothers were primarily white (71.1%), in their mid-twenties (24.1 ± 5.9 years, range = 13.63–43.39), with an average annual income of less than US\$15,000 (US\$ 2003–2010) at the time of adoption. Birth fathers shared similar demographic characteristics but were not included in this analysis because the available sample with non-missing psychopathology data was too small to fit the bifactor measurement model ($n = 179$). The adoptive mothers and fathers were predominantly white (mothers = 91.4%; fathers = 90.4%), older than the birth parents (mothers = 37.4 ± 5.6 years, range = 23.7–55.1; fathers = $38 \pm .59$ years, range = 24.4–59.8) and had an average income of over US\$100,000 per year (US\$ 2003–2010). The children had a median adoption placement age of 2 days ($M = 6.2 \pm 12.45$ days; range = 0–91), were 57.2% male and 55.6% white, followed by multiracial (19.3%), African American (13%), Latino (10.9%), and Asian, American Indian, and unknown ethnicity (each <1%). There was little evidence of selective placement (i.e., children were not placed based on demographic or behavioral characteristics of the birth and adoptive parents), as confirmed by nonsignificant correlations between birth and adoptive parent on characteristics such as personality and intelligence (Leve *et al.*, 2007, 2012).

Procedures

Participants were recruited from March 2003 to January 2010 from 45 adoption agencies. The birth parents completed assessment questionnaires between 3 and 18 months postpartum and adoptive parents at child ages of 9, 18, and 27 months. All participants provided written consent prior to entering the study and were paid for their time. The project was approved by the institutional review boards at the sponsoring institutions. Additional details regarding the EGDS sample, recruitment, and procedures can be found elsewhere (Leve *et al.*, 2019).

Measures

Indices of heritable risk

Birth mother general psychopathology factor (p-factor) was used to index heritable risk for adopted child problem behaviors and lack of heritable risk for child social competence. Following the work of Caspi *et al.* (2014) and Lahey *et al.* (2012), we tested the general factor model for psychopathology using birth mother psychopathology data to create a continuous p-factor variable to represent heritable influences in the adoptee. We used data from the Composite International Diagnostic Interview (CIDI; Kessler, Chiu, Demler, Merikangas, & Walters, 2005) and the Diagnostic Interview Schedule (DIS; Robins, Helzer, Croughan, & Ratcliff, 1981) measured at child age 18 months. Across the eight mental disorders assessed in birth mothers, prevalence

Table 1. Factor loadings for the bifactor measurement model of birth mother psychopathology

	Hierarchical bifactor		
	Externalizing factor	Internalizing factor	General psychopathology factor (p-factor)
Alcohol dependence	.43		.55
Drug dependence	.52		.75
Tobacco dependence	.08		.65
Conduct disorder	.27		.54
Major depressive disorder		.41	.61
Generalized anxiety disorder		.33	.74
Social phobia		.46	.49
Specific phobia		.59	.15
Externalizing factor with internalizing factor		$r = -.58$	

estimates ranged from 7 to 21%. Disorders included alcohol dependence, tobacco dependence, illicit drug dependence, antisocial personality disorder, generalized anxiety disorder, major depressive disorder, specific phobia, and social phobia. Absence of a disorder diagnosis was coded as 0 and the presence of a disorder diagnosis was coded as 1.

We fitted a hierarchical bifactor measurement model using structural equation modeling in the statistical package Mplus (Muthén & Muthén, 2019). The best fitting model was chosen using a combination of the Tucker–Lewis index (TLI) (Tucker & Lewis, 1973), Bentler’s comparative fit index (CFI) (Bentler, 1990), and the root mean square error of approximation (RMSEA) (Browne & Cudeck, 1992). Values above .90 for the CFI and TLI and values below .08 for the RMSEA are considered to reflect an adequate fit. The general factor model had an adequate fit ($\chi^2 = 8.3$, $df = 10$, $p = .6$, CFI > .99, TLI > .99, RMSEA < .01); for this reason, the general psychopathology, specific internalizing problems, and specific externalizing problems factors from this model were extracted and used in subsequent regression analyses. As shown in Table 1, the factor loadings of the general psychopathology factor (p-factor) across birth mother mental disorders were all above .49 (ranging from .49 to .75) except for specific phobia (.15). The p-factor score was correlated with child externalizing problems ($r = .10$, $p = .05$) and marginally correlated with child social competence ($r = -.09$, $p = .06$) at child age 27 months, supporting the validity of this measure for testing a G×E from 27 months to 54 months.

We dichotomized the p-factor variable into low and high risk—low risk ($n = 157$) was set to the minimum score (−2.48) and high risk ($n = 353$) was set to all scores above the minimum (−2.47 to 1.30). Importantly, the minimum score (−2.48) represents birth mothers without any diagnosed disorder, whereas the other scores (−2.47 to 1.30) represent birth mothers with one or more diagnosed disorders. We chose this categorization rather than the median split because the variable was highly skewed, with 31.40% of scores falling at the minimum and the remaining scattered relatively evenly (skew = .96). Furthermore, in the current study, birth mothers in the high-risk p-factor group showed higher antisocial personality symptoms obtained from the Elliott Youth Questionnaire (Elliott & Huizinga, 1983), more negative life events obtained from the Negative Life Events Scale (Dohrenwend, Askenasy, Krasnoff, & Dohrenwend, 1978), lower general life satisfaction, lower financial satisfaction

(higher financial needs), lower self-perceived social competence and self-esteem obtained from the Harter Adult Self-Perception Profile (Messer & Harter, 1986), higher current anxiety and depressive symptoms obtained from the Beck Anxiety Inventory (BAI) (Beck & Steer, 1993) and the Beck Depression Inventory (BDI) (Beck, Steer, & Brown, 1996), and higher self-reported physical health problems obtained from the SF-36 Health Survey (Ware, Kosinski, & Dewey, 2000) as compared to birth mothers in the low-risk p-factor group (each at $p < .05$). Using this cutoff, children categorized in the low-risk group should have a lower likelihood of having genetic risk (or “plasticity”) genes and psychopathology if DST is true. The approach we used here is similar to the threshold model, where we set −2.48 as the threshold (cutoff) score, classifying children into two groups: (a) no/low heritable risk for psychopathology group and (b) moderate to high heritable risk for psychopathology group.

Child home environment at 9, 18, and 27 months

Adoptive mother self-reported parenting stress

Confirmatory factor analysis yielded a self-reported “parenting stress” variable representing the internal mental state of the parent surrounding issues of parenting (versus a direct measurement of parenting behaviors), comprising three self-reported parenting behaviors that have all been validated against observational ratings in other samples. These behaviors included over-reactivity measured using 14 items from the Parenting Scale applicable to infants (Arnold, O’Leary, Wolf, & Acker, 1993), self-efficacy in caring for the child on the 10-item Parental Efficacy scale (Teti & Gelfand, 1991), and perceived hassle intensity using the 20 intensity items from the Parenting Daily Hassles scale (Crnic & Greenberg, 1990). Higher scores on this scale represent promotive parenting behaviors and emotions, with low over-reactivity, low reported hassle intensity, and high self-efficacy, while lower scores represent adverse parenting behaviors and emotions with high over-reactivity, high reported hassle intensity, and low self-efficacy. Inter-item alphas were acceptable for over-reactivity (9-month $\alpha = .70$, 18-month $\alpha = .72$, 27-month $\alpha = .79$), self-efficacy (9-month $\alpha = .73$, 18-month $\alpha = .72$, 27-month $\alpha = .74$), and hassle intensity (9-month $\alpha = .91$, 18-month $\alpha = .92$, 27-month $\alpha = .86$). Factor scores were moderate (.65–.78, .43–.51, and .49–.60, respectively) and significant (all p values < .001). Factor scores were exported for all three waves and averaged

together in a summary score ($\alpha = .85$). Additional details are available from the first author upon request.

Adoptive mother internalizing symptoms

We used 20 items from the BDI (Beck et al., 1996) and the 21-item BAI (Beck & Steer, 1993) to measure parent internalizing symptoms. A question referring to suicidal ideation on the BDI was not administered to the full sample and was excluded. We converted both variables to a scale of 0–1 (subtracted the minimum value and divided by the range) before averaging across 9, 18, and 27 months to create a summary score. The summary score was centered at the mean to improve interpretation of parameter estimates. Inter-item alphas were acceptable for both the BDI (9-month $\alpha = .73$, 18-month $\alpha = .79$, 27-month $\alpha = .83$) and the BAI (9-month $\alpha = .76$, 18-month $\alpha = .78$, 27-month $\alpha = .82$). The alpha coefficient for the summary score was high ($\alpha = .86$).

Adoptive mother marital quality

Marital quality was created based on marital warmth/hostility measured using 22 items from the Behavior Affect Rating Scale (BARS) (Melby, Conger, Ge, & Warner, 1995) on a scale of 1 (always) to 7 (never). Participants rated partners on nine items related to warmth (range 9–63) and 13 related to hostility (range 13–16), with higher scores indicating higher levels of the behavior. We transformed both marital warmth and hostility to scales ranging from 0 to 1, reverse coding marital warmth so that higher scores corresponded to lower warmth before computing an average for each domain across 9-, 18-, and 27-month time points. We then created a summary score by averaging together the converted reverse-coded marital warmth and marital hostility. The summary score was centered at the mean to improve interpretation of parameter estimates. Inter-item alphas were high for both marital warmth (9-month $\alpha = .92$, 18-month $\alpha = .92$, 27-month $\alpha = .94$) and marital hostility (9-month $\alpha = .90$, 18-month $\alpha = .91$, 27-month $\alpha = .90$), and the alpha coefficient for the summary score was high ($\alpha = .92$).

Child outcomes at 54 months

Consistent with previous work testing DST (e.g., Belsky et al., 2013), we included child externalizing and child social competence to test both positive and negative outcomes. Children can be high on both externalizing problems and social competence; therefore, we examined each construct separately rather than creating a single composite construct representing child psychological functioning. If DST holds, patterns should be consistent across multiple outcomes, both positive and negative.

Adoptive-father-reported child externalizing problems

We measured child externalizing problems using *T*-scores on the 1½–5 year Child Behavior Checklist (CBCL) (Achenbach & Rescorla, 2001), which is a parental self-report measure of child noncompliant, hyperactive, or aggressive behaviors that has been validated with observational reports (Achenbach, 2009). We used father-reported rather than mother-reported outcomes to account for correlations due to using the same source for both exposure and outcome. CBCL *T*-scores were used to account for child sex and age norms in externalizing problems. The alpha coefficient was high ($\alpha = .89$). Of note, maternal and paternal ratings of child externalizing problems were moderately correlated ($r = .51$, $p < .01$).

Adoptive-father-reported child social competence

We measured child social competence using 39 items from the preschool version age 3–5 year Social Skills Rating System (SSRS) (Gresham & Elliott, 1990), which is a well-validated parental self-report measure of prosocial behavior including agreeableness, ability to get along with peers, and cooperation (Demaray et al., 1995). The alpha coefficient for this scale was high ($\alpha = .87$). Of note, maternal and paternal ratings of child social competence were moderately correlated ($r = .46$, $p < .01$).

Covariates

To account for potential confounders that may be responsible for observed similarities between birth parents and children, we controlled for several birth parent environmental factors that might impact the child and confound estimates of genetic effects, including adoption openness (contact between birth and adoptive families) and prenatal obstetric complications including exposure to drugs, toxins, and birth complications. We measured adoption openness on a five-point scale (from very closed to very open) between birth and adoptive parents at 5–9, 18, and 27 months (Ge et al., 2008). Previous analyses of 323 adoptive and birth families included in the EGDS assessed at 9 months showed high levels of inter-rater agreement of adoption openness by the three informants (adoptive fathers, adoptive mothers, and birth mothers) ($r = .66$ – $.81$, all p values $< .001$; Ge et al., 2008). In the current study, birth-parent-reported and adoptive-parent-reported adoption openness across the three time points was highly correlated, ranging from .82 to .91, indicating that contacts between birth parents and adoptive parents were consistent. Thus, data were averaged across three waves between birth mothers and adoptive parents to create a composite score ($\alpha = .95$). Obstetric complications were calculated from both self-report and medical records and added together in a summary score (Marceau et al., 2013). In cases when both self-report and medical record data assessed the same type of risk, we computed rates of agreement. After evaluating the agreement across medical record data and self-report, we created “best” scores that incorporate the most reliable and comprehensive score for each construct (Neiderhiser et al., 2016). For more detailed information about creating the “best” scores, please refer to Neiderhiser et al. (2016). Mothers had a mean number of two complications ($SD = 1.24$, range 1–5). We also controlled for child sex (coded 1 = male, 2 = female) because gender differences in genetic susceptibility to environmental factors have been found (Sjöberg et al., 2006) and, for internalizing and externalizing factor scores created in the bifactor model, to obtain the p-factor to account for variance in birth mother risk that is not explained by the psychopathology factor. In a final set of analyses, we controlled for earlier (27-month) child externalizing problems using the CBCL ($\alpha = .90$) and social competence using the Preschool Socio-Affective Profile ($\alpha = .81$) because the SSRS was only collected at 54 months to better understand the timing of the interaction. The descriptive statistics of the main study variables are presented in Table 2.

Analytic strategy

First, we conducted a series of preliminary analyses to explore other explanatory mechanisms or potential confounders that could influence our tests of DST. We used correlational analyses to examine the associations between the birth mother p-factor score and the home environment variables (parenting,

Table 2. Descriptive statistics of main study variables

	Mean	SD	Min.	Max.
9 months				
Adoptive mother over-reactivity	1.72	0.58	1.00	4.57
Adoptive mother self-efficacy	34.70	2.81	25	40
Adoptive mother daily hassles	30.46	12.53	0	73
Adoptive mother anxiety symptoms	3.74	3.58	0	21
Adoptive mother depressive symptoms	3.63	3.30	0	17
Adoptive mother marital warmth	51.54	8.08	24	63
Adoptive mother marital hostility	23.67	7.54	13	66
18 months				
Adoptive mother over-reactivity	2.02	0.68	1.00	4.71
Adoptive mother self-efficacy	33.92	2.80	24	40
Adoptive mother daily hassles	35.90	10.46	11	70
Adoptive mother anxiety symptoms	3.43	3.59	0	28
Adoptive mother depressive symptoms	3.86	3.92	0	25
Adoptive mother marital warmth	50	8.63	18	63
Adoptive mother marital hostility	25.23	8.33	13	62
27 months				
Adoptive mother over-reactivity	2.21	0.68	1	4.43
Adoptive mother self-efficacy	33.89	2.89	23	40
Adoptive mother daily hassles	39.11	10.29	16	78
Adoptive mother anxiety symptoms	3.46	4.02	0	31
Adoptive mother depressive symptoms	3.89	4.19	0	30
Adoptive mother marital warmth	49.75	9.40	12	63
Adoptive mother marital hostility	24.92	8.26	13	68.55
54 months				
Adoptive father reported child externalizing behavior (<i>T</i> -score)	48.66	9	28	83
Adoptive father reported child social competence (standard score)	98.16	13.58	63	129
Birth mother psychopathology factor	0.69	0.47	0	1

internalizing symptoms, and marital quality) to exclude the possibility of measured evocative *rGE*. We tested whether the *p*-factor variable was associated with either child outcome, which would have suggested a diathesis–stress relationship rather than DST (Belsky et al., 2007). We also verified that our three environmental variables of interest were three distinct environmental constructs through additional correlational analyses. We then centered the three environmental variables and created product terms with the *p*-factor variable.

Second, because of a substantial amount of missing data on both child outcomes (33–34%) because of attrition, as well as some minimal level of missing data on the *p*-factor (11%) and our environmental variables of interest (0–2%), we compared children included in our analysis to those excluded because they were missing study variables by select demographic characteristics. We then created imputation datasets ($n = 100$) using the fully conditional specification (FCS) method as recommended for datasets with categorical data and arbitrary missing data patterns (van Buuren, 2007). We included all interaction terms in the

imputation model to avoid biasing our results to the null (Von Hippel, 2009). Imputation was performed using PROC MI and model estimates were combined using PROC MIANALYZE in SAS software version 9.4 (SAS Institute Inc., 2013). We compared differences in parameter estimates between the reduced datasets (externalizing problems $n = 337$, social competence $n = 330$) and the combined estimates from the 100 imputed datasets ($N = 561$) before continuing with our testing of DST on the reduced datasets to allow for full functionality of later procedures using SAS software.

Third, we tested one main effects model and three $G \times E$ models for each child outcome, in line with the work of Widaman et al. (2012). The main effects model for both outcomes contained 10 parameters, including the intercept, one genetic effect, three environmental main effects, and five control variables. The $G \times E$ models contained one additional product term. We used the analysis of variance *F*-test to examine overall model significance for each model and to compare interaction models with the main effects model for each outcome. We considered the $G \times E$ significant if

the model with the interaction term fitted significantly better than the main effects model. We also compared models using the Akaike information criterion (AIC) and Bayesian information criterion (BIC), where lower values suggest better model fit. We then added 27-month child externalizing problems or social competence as additional control variables to test whether the G×E effects remained significant in an additional set of models. For previously significant interactions that were nonsignificant following the addition of earlier behavior problems to the model, we conducted a series of post hoc analyses to better elucidate when in development significant G×Es occurred. We ran models including (a) an average of 9 and 18 months of our three environmental variables predicting 27-month child outcomes with and without controlling for 18-month child behavior and (b) environmental data at child age 9 months predicting 18-month child externalizing behavior, with and without controlling for infant temperament at 9 months (i.e., our best approximation for child behavior in this developmental period). Finally, we ran a cross-lagged correlation analysis using Mplus to simultaneously estimate effects from both parent to child and child to parent in order to assess possible child evocative effects on parental behavior.

Fourth, we probed the interactions for differential susceptibility. We plotted significant interactions for visual inspection using the SAS procedure PROC PLM. We then formally distinguished the interaction between “strong” and “weak” forms of DST and diathesis–stress using re-parameterized equations, by placing a number of constraints on the original regression models using PROC NLIN as outlined by Widaman et al. (2012) and Belsky et al. (2013). For example, we tested DST models by subtracting the cross-over point of the two p-factor group slopes from the observed environmental variable of interest (see Figure 1b). We tested diathesis–stress models by subtracting the most extreme possible promotive environment as the cross-over point from the observed environmental variable of interest (see Figure 1a). We distinguished strong and weak forms by setting the slope of the low p-factor group to zero, representing a more extreme form of DST or diathesis–stress where the low p-factor group is not reactive at all to their home environment. The step-by-step procedure is outlined in Supplementary Material S.1. Our full SAS code adapted from supplementary material provided by Widaman et al. (2012) is available from the first author upon request. We needed a sample size of $N = 221$ to detect a G×E with 11 predictors and 90% power using an alpha level of .05 based on results from a previous EGDS study examining the interaction of parenting and genetic liability calculated by birth parent psychopathology with a total $R^2 = .21$ and a change in R^2 of .04 when adding interaction terms (Leve et al., 2009) using the SAS procedure PROC POWER, well below our available sample non-missing on our variables of interest ($n = 337$ for externalizing problems and $n = 330$ for social competence). We performed all analyses using SAS software version 9.4 (SAS Institute Inc., 2013).

Results

All associations between the birth mother p-factor score and the home environmental variables (parenting, psychopathology, and marital quality) were nonsignificant ($r = -.002$ – $-.08$; *ns*), reducing the concern of the effects of measured evocative rGE and allowing for an uncomplicated test of all forms of G×E. Additionally, the p-factor was not correlated with either child externalizing problems ($r = .08$; *ns*) or social competence ($r = .03$; *ns*) at 54 months.

The three environmental variables were related, but appeared to be distinct constructs with correlations ranging from .24 to .26 (all p values $< .001$). Child social competence and externalizing problems were also inversely correlated ($r = -.39$, $p < .001$).

There were no significant demographic differences between participants with and without missing data except for adoption openness, where participants included in the analysis had significantly higher levels of openness ($M = .08$) compared with those excluded because of missing data ($M = -.11$, $p = .01$). Table 3 provides a comparison of families with missing data and families with full outcome data. There were no substantial differences in any reported results between the imputed dataset including all families and the reduced samples (see Supplementary Tables 1 and 2).

Preliminary test of DST

In the main effects model, maternal parenting stress was a significant predictor of paternal reports of child externalizing (Table 4) and social competence (Table 5), after controlling for the other predictors in the model. Each unit increase in the maternal parenting stress factor score was associated with an average increase of 3.12 units ($SE = 1.43$, $p = .03$) in child externalizing problems (CBCL) and an average decrease of -4.98 units ($SE = 2.14$, $p = .02$) in social competence (SSRS).

We found three significant environmental interactions with the p-factor, including mother parenting stress for both externalizing problems ($\beta = -7.55$, $SE = 2.87$, $p = .009$) and social competence ($\beta = 12.29$, $SE = 4.27$, $p = .004$), and parental internalizing for social competence only ($\beta = -76.89$, $SE = 35.21$, $p = .03$). The predictors included in the interaction models were significantly related with both outcomes according to the analysis of variance F -test (all p -values $< .03$). Additionally, the three two-way interaction terms added a significant increase in explained variance compared with the main effects models (all p values $< .03$) and each had lower AIC and BIC estimates than the main effects models. For the control variables, child sex was negatively associated with child social competence ($p < .05$), such that boys had higher levels of social competence than girls at age 54 months. Other control variables, including adoption openness, birth mother obstetric complications, and birth mother specific externalizing and internalizing factors, were not significantly associated with child externalizing problems and social competence at 54 months.

However, when we controlled for prior levels of child behavior (i.e., externalizing behavior or social competence at 27 months) in a second set of models, controlling for prior levels of social competence produced a negligible change, while the interaction between the p-factor and parenting stress on child externalizing was no longer significant ($\beta = -3.83$, $SE = 2.59$, $p = .14$). Therefore, in order to better elucidate when in development the significant G×Es occurred for child externalizing, we only performed post hoc analyses for parenting stress on child externalizing. In our model including an average of 9–18 months of our three environmental variables predicting 27-month child outcomes, the Gene × Parenting Stress interaction was significant ($p = .04$) but became nonsignificant once we included 18-month child behavior ($p = .16$; see Supplementary Table 1). In our second set of models including age 9 months environmental data predicting 18-month child externalizing behavior, there was no interaction present at all, regardless of whether we controlled for infant behavior (see Supplementary Table 2). We found evidence of evocative effects of child externalizing on

Table 3. Comparison of the full sample versus families with missing data

Characteristic	Missing data		p^b
	Yes ($n = 233$) ^a	No ($n = 328$) ^a	
Adoptive mother age	37.84 ± 5.27	38.60 ± 6.04	.12
Adoptive father age	38.89 ± 5.77	39.60 ± 5.92	.18
Adoption openness ^c	0.11 ± 0.95	0.08 ± .86	.01
Financial difficulty ^d	0.12 ± 0.09	0.14 ± .11	.09
Child sex			.91
Male	134 (57.51)	187 (57.01)	
Female	99 (42.49)	141 (42.99)	
Child race/ethnicity			.12
Non-Hispanic white	120 (51.50)	192 (58.54)	
Non-Hispanic black	37 (15.88)	36 (10.98)	
Hispanic	22 (9.44)	39 (11.89)	
Other	54 (23.18)	61 (18.60)	
Adoptive mother education			.44
Less than 4-year college degree	16 (8.94)	19 (6.96)	
4-year college degree or higher	163 (91.06)	254 (93.04)	
Adoptive father education			.13
Less than 4-year college degree	20 (13.42)	23 (8.65)	
4-year college degree or higher	129 (86.58)	243 (91.35)	

Notes: We used listwise deletion in order to allow for full SAS functionality to explore significant interactions, currently unavailable with multiple imputation. Table values are mean ± SD for continuous variables and n (column %) for categorical variables.

^a n may not sum to 100 due to missing demographics, percentages may not sum to 100% due to rounding.

^b p -value is for t -test (continuous variables) or chi-square test (categorical variables).

^cAdoption openness is a standardized variable with a mean of 0.

^dFinancial difficulty is the average of material worry, making ends meet, and financial cutbacks on a scale of 0–1.

parenting stress from 18 to 27 months in a cross-lagged analysis (see Supplementary Figure 1).

As shown in Figures 2 and 3, follow-up analyses to explore the nature of the significant interactions revealed that mothers' higher levels of parenting stress were not associated with externalizing problems (slope = .57, $SE = 1.72$, $p = .74$) or social competence (slope = $-.86$, $SE = 2.56$, $p = .74$) for individuals with high levels of the p -factor, while mothers' parenting stress was significantly associated with child outcomes for those with low levels of the p -factor (externalizing problems: slope = 8.12, $SE = 2.37$, $p < .001$; social competence: slope = -13.15 , $SE = 3.54$, $p < .001$). These findings are not in line with either DST or diathesis–stress theory. Figure 4 shows the opposite pattern for parent internalizing symptoms on child social competence: children with high levels of the p -factor had decreasing levels of social competence as parental internalizing increased (slope = -42.54 , $SE = 19.90$, $p = .03$). Adoptive mothers' internalizing symptoms were not associated with children's social competence for those with low levels of the p -factor (slope = 34.34, $SE = 30.57$, $p = .26$). All interactions had cross-over points within the observed values of the environmental variables. However, the only significant interaction that resembled DST was parental internalizing symptoms on social competence. Therefore, this was the only interaction we formally tested statistically.

Parameter estimates between our full-case and combined imputation datasets were consistent, although the interaction terms were

only nonsignificant trends ($p = .06$ – $.07$; see Supplementary Tables 3 and 4). Despite nonsignificant interactions with the imputed models, we chose to explore the nature of the interactions considering (a) the uncertainty of imputation and the facts that (b) the interactions were still below a significance level of $p < .10$, (c) our primary hypothesis involved product terms, and (d) parameter estimates were consistent between datasets.

Confirmatory statistical test of DST

We examined the nature of interaction of the p -factor and adoptive mother internalizing problems on child social competence, comparing results across four models: strong and weak forms of DST, and strong and weak forms of diathesis–stress (Table 6), following Widaman et al. (2012) and Belsky et al. (2013).

The strong form of DST fitted the data well ($F = 10.21$, $df = 2$, 327, $p < .001$) with an R^2 of .06. However, the lower bound of the cross-over point for the strong DST model was at the minimum possible value of X (95% CI = $[-0.04, 0.11]$). The weak form of DST also fitted the data well ($F = 7.26$, $df = 3, 236$, $p < .001$) with an R^2 of .06. The cross-over point (mother internalizing = .03) was close to the mean value of the environmental variable X (mother internalizing = .02), and its CI (95% CI = $[-0.02, 0.07]$) was within bounds of the observable range of X (range = $-.04$ to $.24$). The strong DST model did not result in a significant reduction in explained variance compared with the less restricted

Table 4. Standard regression models for child externalizing problems ($n = 337$)

Parameter	G and E main effects: Model 1		p-factor×Internalizing symptoms: Model 2		p-factor×Parenting stress: Model 3		p-factor×Marital warmth/hostility: Model 4	
	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p
Intercept	49.93 (1.90)	<.001	50.18 (1.93)	<.001	49.96 (1.88)	<.001	49.66 (1.90)	<.001
Internalizing symptoms	16.58 (11.16)	.14	2.62 (20.42)	.90	17.37 (11.07)	.12	15.41 (11.15)	.17
Marital warmth/hostility	-2.35 (4.69)	.62	-2.2 (4.70)	.64	-0.76 (4.69)	.87	10.51 (8.73)	.23
Parenting stress	3.12 (1.43)	.03	3.06 (1.44)	.03	8.12 (2.37)	.001	3.41 (1.44)	.02
p-factor	1.63 (1.31)	.21	1.32 (1.37)	.34	1.65 (1.30)	.21	1.95 (1.32)	.14
p-factor×Environmental factor	—	—	19.02 (23.29)	.41	-7.55 (2.87)	.009	-17.34 (9.95)	.08
Model fit								
R^2	.04		.05		.06		.05	
F	1.7		1.6		2.25		1.85	
df	9,327		10,326		10,326		10,326	
P	.09		.11		.01		.05	
Model comparisons								
F versus 1 ^a	—		.67		6.93		3.04	
df	—		1,326		1,326		1,326	
P	—		.41		.009		.08	
AIC	1482.05		1483.36		1476.96		1480.92	
BIC	1484.66		1486.10		1479.70		1483.66	

Notes: Reduced sample size is due to listwise deletion; estimates are unstandardized.

^a F versus 1 is the analysis of variance F -test comparing difference in explained variance for listed model versus Model 1.

Table 5. Standard regression models for child social competence ($n = 330$)

Parameter	G and E main effects: Model 1		p-factor×Internalizing symptoms: Model 2		p-factor×Parenting stress: Model 3		p-factor×Marital warmth/hostility: Model 4	
	β (SE)	p	β (SE)	p	β (SE)	p	β (SE)	p
Intercept	102.81 (2.86)	<.001	101.83 (2.88)	<.001	102.81 (2.83)	<.001	102.97 (2.88)	<.001
Internalizing symptoms	−20.80 (17.34)	.23	34.34 (30.57)	.26	−21.75 (17.15)	.21	−20.09 (17.38)	.25
Marital warmth/hostility	.08 (7.06)	.99	−.20 (7.02)	.98	−2.42 (7.04)	.73	−8.05 (13.52)	.55
Parenting stress	−4.98 (2.14)	.02	−4.75 (2.13)	.03	−13.15 (3.54)	<.001	−5.14 (2.15)	.02
p-factor	.77 (1.98)	.70	2.05 (2.06)	.32	.71 (1.96)	.72	0.58 (2.00)	.77
p-factor×Environmental factor	—		−76.89 (35.21)	.03	12.29 (4.27)	.004	10.77 (15.25)	.48
Model fit								
R^2	.05		.06		.07		.05	
F	1.82		2.13		2.5		1.68	
df	9, 320		10, 319		10, 319		10, 319	
P	.06		.02		.007		.08	
Model comparisons								
F versus 1 ^a	—		4.77		8.28		.5	
df	—		1,319		1,319		1,319	
P	—		.03		.004		.48	
AIC	1714.83		1711.93		1708.37		1716.31	
BIC	1717.45		1714.69		1711.12		1719.07	

Notes: reduced sample size is due to listwise deletion; Estimates are unstandardized.

^a F versus 1 is the analysis of variance F -test comparing difference in explained variance for listed model versus Model 1.

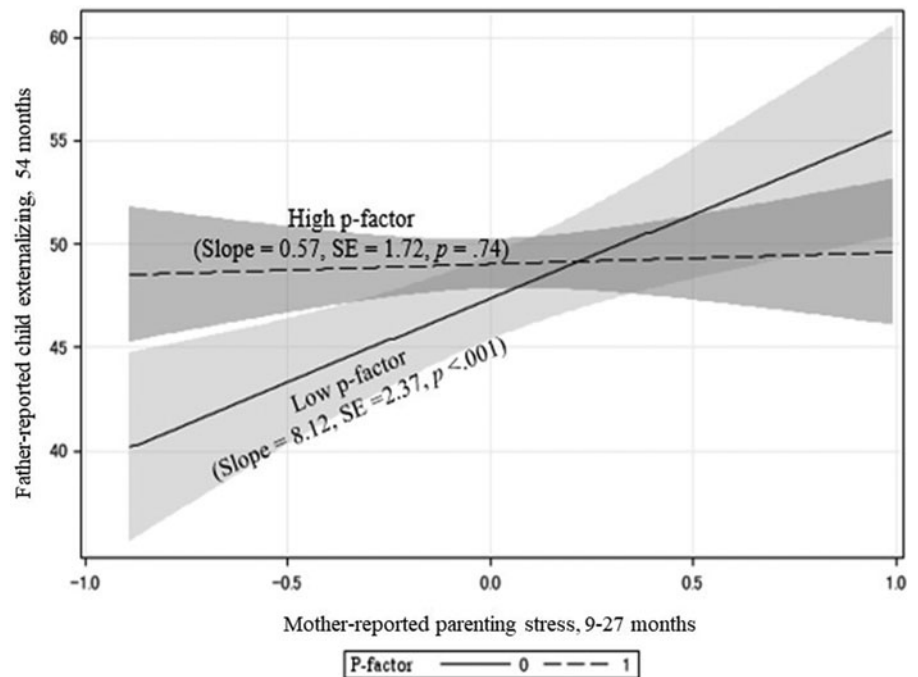


Figure 2. Interaction between parenting stress and p-factor (general psychopathology factor) on child externalizing problems. *Note:* Children with high genetic risk were not susceptible to parenting stress, while children with low risk were susceptible to parenting stress, both positively and negatively.

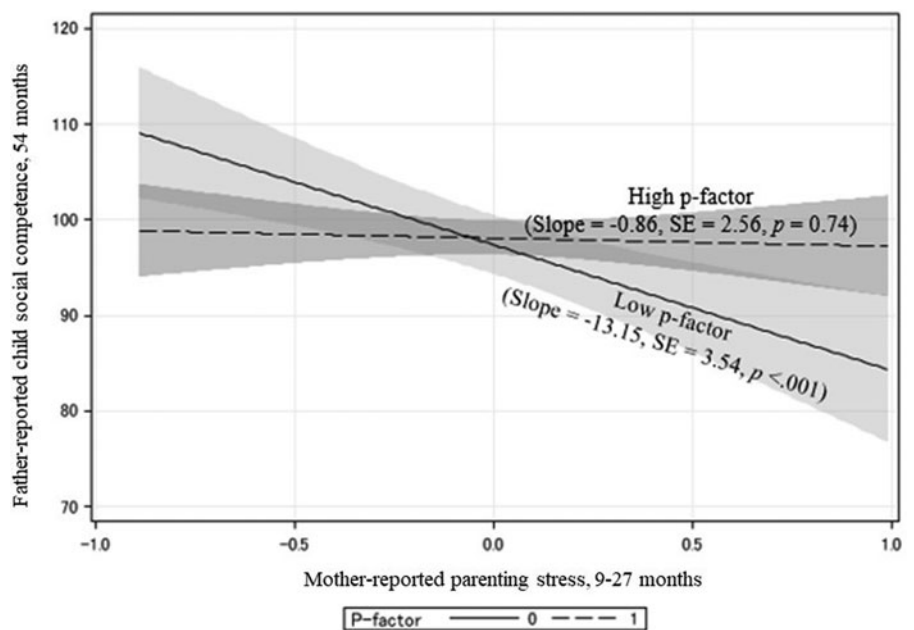


Figure 3. Interaction between parenting stress and p-factor (general psychopathology factor) on child social competence. *Note:* Children with high genetic risk were not susceptible to parenting stress, while children with low risk were susceptible to parenting stress, both positively and negatively.

weak DST model ($F = 1.36$, $df = 1,136$, $p = .25$). Our data somewhat supported the strong form of DST over the weak form of DST; however, diathesis–stress could not be ruled out as the CI of the strong DST model was on the lower bound of the cross-over point.

The strong form of diathesis–stress fitted the data well ($F = 17.71$, $df = 1,328$, $p < .001$) with an R^2 of .05. The weak form of diathesis–stress fitted the data similarly well ($F = 8.85$, $df = 2,327$, $p < .001$) with an R^2 of .05. Our data supported the strong over the weak form of diathesis–stress as it did not result in significant loss of explained variance (comparison not shown; $F = .03$, $df = 1,327$, $p = .85$), and also had a slightly lower

AIC (1697.93 versus 1699.90) and BIC (1705.53 vs. 1711.30) as compared with the weak form of diathesis–stress. Additionally, the strong diathesis–stress model did not result in a significant reduction in explained variance as compared with either differential susceptibility model (strong DST: $F = 2.61$, $df = 1,327$, $p = .11$; weak DST: $F = 1.99$, $df = 2,326$, $p = .14$), despite having a higher overall proportion of variance explained ($R^2 = .06$ versus .05) as well as higher AIC and BIC statistics.

In summary, as the CI of the cross-over point for the strong DST model was at the lower bound of the observable range of the environmental variable in question and the strong diathesis–stress model did not result in a significant loss in

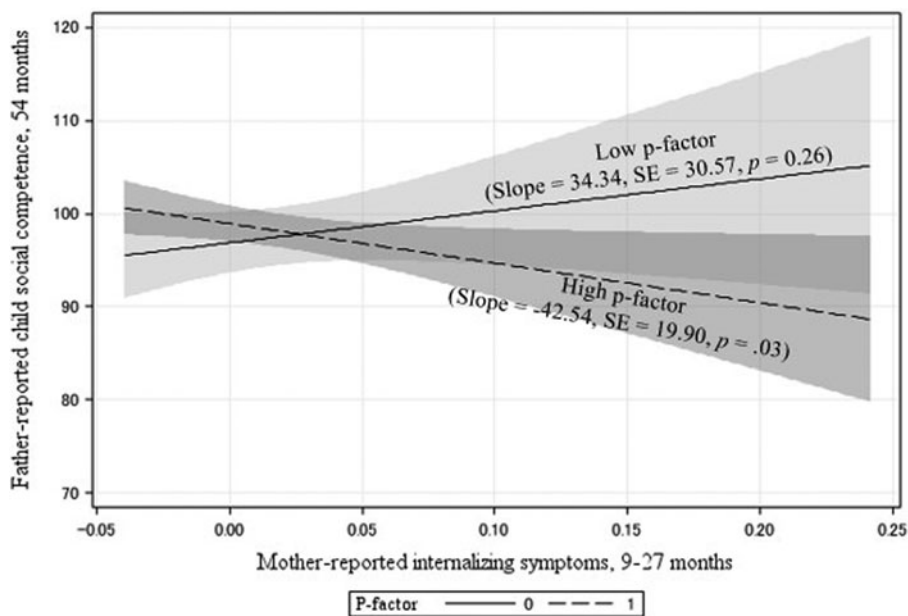


Figure 4. Interaction between parental internalizing symptoms and p-factor (general psychopathology factor) on child social competence. *Note:* Children with high genetic risk exposed to greater parent internalizing problems had lower levels of social competence, while children with low genetic risk had slightly higher levels of social competence.

explained variance over either DST model, our results supported the strong form of diathesis–stress as the most likely theory for explaining the interaction between the p-factor and adoptive mother internalizing symptoms on child social competence. Further, we did not find any interactions consistent with diathesis–stress or DST for child externalizing problems.

Post hoc sensitivity analyses

We conducted a series of post hoc sensitivity analyses. First, we examined the same models using father-reported home environmental variables (parenting, psychopathology, and marital quality) and mother-reported outcomes. However, no significant interactions between the birth mother p-factor score and the home environmental variables were found for these models, even when including only mother reports or only father reports in the analyses. Second, we explored other possibilities for capturing the variability in the p-factor by re-examining the same models using the continuous p-factor score. The results were consistent with the findings reported above using the binary p-factor score, except that the interaction between maternal parenting stress and the continuous p-factor score was no longer significant.

Discussion

DST predicts that individuals with plasticity genes will be more susceptible to their environments, such that those with plasticity genes will have the best outcomes when placed in promotive environments and the worst outcomes when placed in negative environments, as compared with those without these genes (Belsky et al., 2007; Reiss et al., 2013). Our study did not find evidence for this theory using a broad, general liability for psychopathology, a range of promotive and adverse home environmental variables, and two important child outcomes at 54 months (i.e., externalizing problems and social competence). We used a parent–offspring adoption design, which is a particularly useful method for disentangling heritable and prenatal influences from postnatal rearing environmental effects because (a) children are not genetically related to their adoptive parents who provide the

rearing environment and (b) birth parents provide genes to the adopted child but do not provide the rearing environment. This eliminates the confounding of postnatal rearing environment and shared genes between children and parents.

Findings for exposure to maternal internalizing problems were consistent with the diathesis–stress model (Goforth, Pham, & Carlson, 2011) for child social competence. Children with high levels of the p-factor were vulnerable and children with low levels of the p-factor were not susceptible to their adoptive mother’s internalizing symptoms. We found results opposite in form from what we anticipated for parenting stress. For these models, children with lower levels of the p-factor appeared to be responsive to their negative environments while those with higher levels of the p-factor were not. Finally, marital warmth and hostility were not associated with child outcomes regardless of level of the p-factor.

In contrast to predictions from DST or the diathesis–stress model, we found a positive association between parenting stress and child externalizing problems in the low p-factor group, but not in the high p-factor group. These findings are consistent with previous work examining candidate G×E, which showed that children with the low-risk genotype (e.g., DRD4 4-repeat homozygous variant) were more susceptible to both positive (e.g., social wellbeing) and negative environments (e.g., peer victimization, perinatal complications) compared to children with the high-risk genotype (e.g., DRD4 7-repeat allele) (Bersted & DiLalla, 2016; Kretschmer, Dijkstra, Ormel, Verhulst, & Veenstra, 2013). It is possible that genetic influences may only be seen at positive ends of environmental influences, considering that the CIs overlapped across both p-factor groups, except at low levels of parenting stress. Adverse environments (here, high parenting stress) might make children relatively high in externalizing; their genetic differences only apparent in favorable environments. These results are consistent with previous work with the EGDS sample by Lipscomb et al. (2012), where birth mother negative affect was only related to child negative emotionality at 9 months at low levels of over-reactive parenting. Both sets of findings are in line with the social push hypothesis, where genetic liability becomes most apparent under positive environmental conditions

Table 6. Re-parameterized regression model: p-factor moderation of parental internalizing problems on child social competence

Parameter	Differential susceptibility theory (DST)				Diathesis–stress (DS)			
	Strong DST ^a		Weak DST ^b		Strong DS ^c		Weak DS ^d	
	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI	β (SE)	95% CI
Intercept	102.40 (1.34)	99.80, 105.10	102.70 (.85)	101.10, 104.40	104.00 (.96)	102.10, 105.90	104.10 (1.16)	101.80, 106.40
Cross-over point	0.03 (0.04)	–0.04, 0.11	0.03 (0.02)	–0.02, 0.07	–0.04 (–) ^e	– ^e	–0.04 (–) ^e	– ^e
Slope, low p-factor	0.00 (–) ^e	– ^e	34.34 (29.46)	–23.61, 92.29	0.00 (–) ^e	– ^e	–4.11 (22.30)	–47.97, –59.89
Slope, high p-factor	–42.55 (18.21)	–78.37, –6.72	–42.52 (18.20)	–78.35, –6.74	–26.05 (15.13)	–55.79, 3.70	–27.29 (16.57)	–59.89, 5.32
Model fit								
R^2	.06		.06		.05		.05	
F	10.21		7.26		17.71		8.85	
df	2,327		3,326		1,328		2,327	
p	<.001		<.001		<.001		<.001	
Model comparisons								
F versus strong DST ^g	–		1.36		2.61		– ^f	
df	–		1,326		1,327		– ^f	
P	–		.25		.11		– ^f	
Model comparisons								
F versus weak DST ^h	1.36		–		1.99		3.94	
df	1, 326		–		2, 326		1, 326	
p	.25		–		.14		.05	
AIC	1697.3		1697.93		1697.93		1699.90	
BIC	1708.7		1713.13		1705.53		1711.30	

Notes: ^aStrong DST: We subtracted the cross-over point \hat{C} from the observed value of the environmental variable of interest X . We then further re-parameterized the model by setting the slope of the low p-factor group to 0 and re-ran the regression equation. This model represents a more extreme form of DST where the low p-factor group is not at all reactive to their home environment.

^bWeak DST: We subtracted the cross-over point \hat{C} from the observed value of the environmental variable of interest X and re-ran the regression equation.

^cStrong DS: We subtracted the most positive extreme value of X (or most adaptive environment) from the observed value of the environmental variable of interest. We then further re-parameterized the model by setting the slope of the low p-factor group to 0 and re-ran the regression equation.

^dWeak DS: We subtracted the most positive extreme value of X (or most adaptive environment) from the observed value of the environmental variable of interest and re-ran the regression equation.

^eFixed at reported value, no standard error or CI.

^fThe strong DST and weak DS models were not nested and could not be directly compared to one-another using the F -test, so total R^2 and the AIC and BIC were used to compare models.

^g F versus strong DST is the analysis of variance F -test comparing difference in explained variance for listed model versus strong DST.

^h F versus weak DST is the analysis of variance F -test comparing difference in explained variance for listed model versus weak DST.

(Bronfenbrenner & Ceci, 1994; Raine, 2002). Similar results have been found looking at childhood IQ, where genetic influences are most pronounced at higher levels of socioeconomic status (Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003); however, not all reports confirm these findings (Hanscombe et al., 2012).

Interestingly, when we controlled for prior levels of child functioning, the G×E for parenting stress was no longer significant for the externalizing behavior model, suggesting unmeasured child evocative processes could be operating. In our cross-lagged analysis to further explore this possibility, we found evocative effects of child externalizing on parenting stress from age 18 to 27 months. It is possible that unmeasured, potentially heritable, characteristics of the child unrelated to the general liability for psychopathology are producing the results shown in Figure 2 (Reiss, 2016).

Of note, we chose to measure the child's home environment using data reported by the adoptive mother and child outcomes using father-reported data as a strict control of shared method variance. However, our post hoc analysis using father-reported home environmental variables (parenting, psychopathology, and marital quality) and mother-reported outcomes did not reveal significant interactions on child outcomes. This was a surprise given that source effects should bias associations further from unity. It is possible that child experiences of the father environmental variables we measured are not as salient to young children; similar findings have been reported in the EGDS (Natsuaki et al., 2010) and in other samples of toddlers (Meadows, McLanahan, & Brooks-Gunn, 2007) and young adolescents (Tully, Iacono, & McGue, 2008), while fathers may have more of an impact on child behavior in later adolescence (Marmorstein, Iacono, & McGue, 2012). However, other studies have found effects of father depressive symptoms on very young children (e.g., Ramchandani et al., 2008), and it is unclear and perplexing as to why we did not find effects for mother-reported outcomes in analyses where mothers were reporting their own parenting stress and internalizing.

We did not find evidence that the p-factor was associated with child outcomes at 54 months. However, the p-factor was associated with child outcomes at 27 months, justifying its use as a measure of genetic risk in modeling G×Es from child age 27 months to 54 months. Similarly, prior work examining the influence of DRD4 7-repeat allele on child externalizing and social competence found nonsignificant genetic main effects among children of a similar age to the current study (Belsky et al., 2013). Additionally, the p-factor was associated with other measures of birth parent functioning not included in the measurement of this construct reflecting antisocial personality symptoms, depressive symptoms, and lower life satisfaction, supporting the validity of the p-factor for use as a susceptibility factor.

It is also important to note the relatively small proportion of the variance explained by G×Es in the current sample; the addition of interaction terms increased the variance explained by 1–2%. It is possible that our low R^2 value might be explained by limited variation in outcome or exposure among our sample of adoptive parents and children, or low data precision, despite our use of highly validated measures. Low data precision likely explains our low R^2 values at least in part, as shown by the relatively large standard errors for the interaction term effect estimates.

Limitations

Although the bifactor structure of psychopathology has been replicated in several studies since Lahey et al. (2012) and Caspi et al.

(2014) brought it to the forefront of psychological research, Bonifay, Lane, and Reise (2016) warn that results obtained from models using the p-factor as a measure of general liability should be interpreted with caution. Some studies have shown that models with a bifactor structure fit best because of model over-fitting (Murray & Johnson, 2013). Additionally, strong correlations at the phenotypic level may not implicate an identical structure at the genotypic level. However, this concern is partially alleviated by recent work in children using multiple informants to estimate the bifactor model along with genome-wide complex trait analysis (GCTA), showing significant p-factor SNP heritability estimates of 38% (Neumann et al., 2016). The authors concluded that “common autosomal SNPs underlie a general psychopathology factor in childhood,” indicating that the p-factor may be capturing the true structure of psychopathology.

Additionally, the p-factor variable created in this study included only externalizing and internalizing dimensions, and did not account for other disorders (e.g., autism or psychotic disorders). Murray et al. (2016) note that future studies including the full spectrum of mental disorders would be beneficial in fully exploring the p-factor. Another consideration is that, according to DST, birth mothers with susceptibility genes exposed to enriched environments would have fewer symptoms of psychopathology compared with birth mothers without susceptibility genes. Therefore, it is possible a subset of children with susceptibility genes were misclassified as not having genetic susceptibility. Furthermore, we only included psychopathology information from birth mothers because our study was not adequately powered to examine whether the results hold for birth fathers, for whom we only had psychiatric data on 40% of the total sample. However, as we only captured half of the child's genetic risk, we would expect that the results would only be attenuated. Additionally, it would have been beneficial to have multiple points of data to calculate the p-factor variable, although a recent study has shown the p-factor to be very stable even during developmental periods marked by instability and change (Murray et al., 2016).

Furthermore, restricted variations in environmental exposures and child outcomes in our sample of adoptive families (i.e., average family income exceeding US\$100,000) and a lack of bipolar single measurements capturing a spectrum of truly positive and negative exposures and outcomes contributed to our inability to model a true continuum of child and parent mental health in a single analysis, which may have interfered with our ability to distinguish between diathesis–stress and differential susceptibility. While parenting stress and marital warmth consisted of high efficacy, low hassles, low over-reactivity, and high marital warmth on one end of the spectrum and low efficacy, high hassles, high over-reactivity and low marital warmth on the other, variability on these scales was not extensive and we did not have a truly bipolar dimension for parent or child mental health. Child social competence may be more representative of a bipolar dimension, as low social competence could indicate worse psychological functioning than would high externalizing. This could explain why we found qualitative evidence of DST for this child outcome (although not statistical evidence, perhaps because parent internalizing is a primarily negative environmental exposure).

Another limitation is that it is possible that maternal parenting stress is partially influenced by child behaviors at an earlier time. However, similar to all studies of parents and children, we were unable to eliminate entirely the confounding of child evocative effects on maternal parenting stress. Specifically, maternal

parenting stress, including over-reactivity, self-efficacy in caring for the child, and perceived hassle intensity, were assessed at child ages of 9, 18, and 27 months, whereas the earliest assessments of child problem behaviors or social competence were at 18 months. Thus, we did not have information on child externalizing problems and social competence at 9 months or earlier than 9 months, limiting the extent to which we can eliminate the confounding of child evocative effects on maternal stress. Nevertheless, we were able to control for some evocative effects within the limits of our measurement and study design.

An additional consideration is that ethnicity has been found to be an important moderator of differential susceptibility. A meta-analysis including over 9,361 adolescents found that differential susceptibility was more pronounced in samples with >80% whites as compared with >20% mixed-ethnicity samples because white children with the nonsusceptible variant of the serotonin transporter genotype 5HTTLPR were less affected by their home environments at both ends of the spectrum than were children of other ethnicities (van Ijzendoorn, Belsky, & Bakermans-Kranenburg, 2012). The effects we observed from our data, which were largely from European-American children (55.6%), may have been attenuated. However, we expect the p-factor is robust against population stratification because genetic risk is estimated without the measurement of individual polymorphisms.

We were unable to replicate our findings in our post hoc analyses, which found nonsignificant interactions using the continuous p-factor variable. One potential explanation is that the continuous p-factor variable was highly skewed, which provided justification for our a priori decision to dichotomize the variable. Using a highly skewed p-factor variable could have contributed to our nonsignificant findings. Second, it is possible that birth parents with more than one psychopathology do not confer additional genetic risk to their children compared with birth parents with a single psychopathology, at the time in development when the measurements were obtained. However, it is possible that analyses using the binary p-factor score may exaggerate group differences in reacting to low versus high maternal parenting stress, leading to cautions in interpreting the findings examining the interaction effects between maternal parenting stress and the binary p-factor score on child problem behaviors and social competence. Further, despite a larger, yet comparable sample size when testing our models with mother-only reported environmental factors and outcomes, interactions were no longer significant. This finding was a surprise as source effects should bias associations further from unity and could suggest unreliable results.

Finally, the adoption design does not fully disentangle prenatal and heritable influences. However, by carefully measuring birth mothers' lifetime psychopathology characteristics and birth mothers' prenatal environments separately, this design helps to clarify the effects of prenatal environments and heritable factors on child development.

Conclusions

In our sample of adoptive families we did not find evidence supporting the differential susceptibility hypothesis in relation to child externalizing problems or social competence using a broad general liability for psychopathology and a range of environmental exposures. Rather, we found differing results depending on the adoptive parent environmental factor studied. High levels of the p-factor, as measured by adoptive parent psychopathology, appear

to confer a liability when adoptive mother internalizing symptoms are present, consistent with a diathesis-stress hypothesis, for social competence but not for externalizing problems. Conversely, high levels of the p-factor appear to confer a resistance to both positive and negative maternal parenting stress.

Our results, coupled with mixed findings in the literature in relation to DST in a variety of settings (e.g., Belsky & Pluess, 2013b; Fong, Measelle, Conrath, & Ablow, 2017; Leve *et al.*, 2009), do not provide evidence supporting tailoring interventions to children based on their genetic susceptibility or "sensitivity" alone. However, genetic differences can be considered in cases where promising family interventions show weak or no significant effects, as highlighted by van Ijzendoorn & Bakermans-Kranenburg (2015). The American Academy of Pediatrics recommendation to screen all children for a range of environmental risk factors, including parental mental health history (American Academy of Pediatrics, 2017), is necessary, and may be even more important for children who are genetically vulnerable to parental psychiatric symptoms. This work expands the differential susceptibility literature using a broad measure of genetic susceptibility and a genetically informed design. Although this study by no means disproves DST, the results suggest further investigation of DST is warranted before revising screening or intervention recommendations in young children.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S0954579420000450>

Acknowledgments. Work on this study was supported by a training grant to the first author from the National Institute of Mental Health (T32 MH014235; PI: Heping Zhang PhD) and the Yale School of Public Health. This project was supported by grant R01 HD042608 from the National Institute of Child Health and Human Development and the National Institute on Drug Abuse, NIH, U.S. PHS (PI years 1–5: David Reiss MD; PI years 6–10: Leslie Leve PhD), grant R01 DA020585 from the National Institute on Drug Abuse, the National Institute of Mental Health and OBSSR, NIH, U.S. PHS (PI: Jenae Neiderhiser PhD), grant R01 MH092118 from the National Institute of Mental Health, NIH, U.S. PHS (PIs: Jenae Neiderhiser PhD and Leslie Leve PhD) and grant UH3 OD023389 from the Office of the Director, National Institutes of Health (PIs: Leslie Leve PhD, Jenae Neiderhiser PhD, and Jody Ganiban PhD). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development or the National Institutes of Health (NIH). We extend special thanks to the families and children who participated in this study. We would also like to thank Sally Guyer and the EGDS staff and researchers whose work helped make this study possible. Correspondence concerning this article should be addressed to Robyn Cree, 60 College Street, New Haven, CT 06511, USA (robyncree@gmail.com) or David Reiss, PO Box 207900, New Haven, CT 06520, USA (david.reiss@yale.edu).

References

- Achenbach, T. M. (2009). *The Achenbach system of empirically based assessment (ASEBA): Development, findings, theory, and applications*. Burlington, VT: University of Vermont Research Center for Children, Youth and Families.
- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- Aiken, L., & West, S. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- American Academy of Pediatrics. (2017). *Bright futures: Guidelines for health supervision of infants, children and adolescents* (4th ed., p. 868). Itasca, IL: American Academy of Pediatrics.

- Arnold, D. S., O'Leary, S. G., Wolf, L. S., & Acker, M. M. (1993). The parenting scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment*, 9, 137–144. doi:10.1037/1040-3590.5.2.137
- Bagner, D. M., Pettit, J. W., Lewinsohn, P. M., Seeley, J. R., & Jaccard, J. (2013). Disentangling the temporal relationship between parental depressive symptoms and early child behavior problems: A transactional framework. *Journal of Clinical Child & Adolescent Psychology*, 42, 78–90. doi:10.1080/15374416.2012.715368
- Beck, A. T., & Steer, R. A. (1993). *Beck anxiety inventory manual*. San Antonio, TX: Psychological Corporation.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.
- Belsky, J. (2005). Differential susceptibility to rearing influence: An evolutionary hypothesis and some evidence. In B. Ellis & D. Bjorklund (Eds.), *Origins of the social mind: Evolutionary psychology and child development* (pp. 139–163). New York, NY: Guilford Press.
- Belsky, J., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). For better and for worse: Differential susceptibility to environmental influences. *Current Directions in Psychological Science*, 16, 300–304. doi:10.1111/j.1467-8721.2007.00525.x
- Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin*, 135, 885–908. doi:10.1037/a0017376
- Belsky, J., & Pluess, M. (2013a). Beyond risk, resilience, and dysregulation: Phenotypic plasticity and human development. *Development and Psychopathology*, 25, 1243–1261. doi:10.1017/S095457941300059X
- Belsky, J., & Pluess, M. (2013b). Genetic moderation of early child-care effects on social functioning across childhood: A developmental analysis. *Child Development*, 84, 1209–1225. doi:10.1111/cdev.12058
- Belsky, J., Pluess, M., & Widaman, K. F. (2013). Confirmatory and competitive evaluation of alternative gene-environment interaction hypotheses. *Journal of Child Psychology and Psychiatry*, 54, 1135–1143. doi:10.1111/jcpp.12075
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238.
- Bersted, K. A., & DiLalla, L. F. (2016). The influence of DRD4 genotype and perinatal complications on preschoolers' negative emotionality. *Journal of Applied Developmental Psychology*, 42, 71–79. doi:10.1016/j.appdev.2015.12.001
- Bonifay, W., Lane, S. P., & Reise, S. P. (2016). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science*, 5, 184–186. doi:10.1177/2167702616657069
- Boyce, W. T., & Ellis, B. J. (2005). Biological sensitivity to context: I. An evolutionary-developmental theory of the origins and functions of stress reactivity. *Development and Psychopathology*, 17, 271–301. doi:10.1017/S0954579405050145
- Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: a bioecological model. *Psychological Review*, 101, 568–586. doi:10.1037/h0044895
- Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, 21, 230–258. doi:10.1177/0049124192021002005
- 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 Psychology Review*, 2, 119–137. doi:10.1177/2167702613497473
- Crnic, K. A., Gaze, C., & Hoffman, C. (2005). Cumulative parenting stress across the preschool period: Relations to maternal parenting and child behaviour at age 5. *Infant and Child Development*, 14, 117–132. doi:10.1002/icd.384
- Crnic, K. A., & Greenberg, M. T. (1990). Minor parenting stresses with young children. *Child Development*, 61, 1628–1637. doi:10.2307/1130770
- Cross-Disorder Group of the Psychiatric Genomics Consortium. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature Genetics*, 45, 984–994. doi:10.1038/ng.2711
- Deater-Deckard, K. (1998). Parenting stress and child adjustment: Some old hypotheses and new questions. *Clinical Psychology: Science and Practice*, 5, 314–332. doi:10.1111/j.1468-2850.1998.tb00152.x
- Demaray, M. K., Ruffalo, S. R., Carlson, J., Busse, R. T., Olson, A. E., McManus, S. M., & Leventhal, A. (1995). Social skills assessment: A comparative evaluation of six published rating scales. *School Psychology Review*, 24, 648–671.
- Dick, D. M., Agrawal, A., Keller, M. C., Adkins, A., Aliev, F., Monroe, S., ... Sher, K. J. (2015). Candidate gene-environment interaction research: Reflections and recommendations. *Perspectives on Psychological Science*, 10, 37–59.
- Dodge, K. A., & Crick, N. R. (1990). Social information-processing bases of aggressive behavior in children. *Personality and Social Psychology Bulletin*, 16, 8–22. doi:10.1177/0146167290161002
- Dohrenwend, B. S., Askenasy, A. R., Krasnoff, L., & Dohrenwend, B. P. (1978). Exemplification of a method for scaling life events: The PERI life events scale. *Journal of Health Social Behavior*, 19, 205–229.
- Duncan, L. E., & Keller, M. C. (2011). A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *The American Journal of Psychiatry*, 168, 1041–1049.
- Eley, T., Keers, R., Coleman, J., Lester, K., Thastum, M., Schneider, S., ... Breen, G. (2017). Polygenic scores for differential susceptibility to emotional symptoms predict response to Cbt in child anxiety. *European Neuropsychopharmacology*, 27, S502–S503.
- Elliott, D. S., & Huizinga, D. J. C. (1983). Social class and delinquent behavior in a national youth panel: 1976–1980. *Criminology*, 21, 149–177.
- Fearon, R. M., Reiss, D., Leve, L. D., Shaw, D. S., Scaramella, L. V., Ganiban, J. M., & Neiderhiser, J. M. (2015). Child-evoked maternal negativity from 9 to 27 months: Evidence of gene-environment correlation and its moderation by marital distress. *Development and Psychopathology*, 27, 1251–1265. doi:10.1017/S0954579414000868
- Feng, X., Shaw, D. S., Kovacs, M., Lane, T., O'Rourke, F. E., & Alarcon, J. H. (2008). Emotion regulation in preschoolers: The roles of behavioral inhibition, maternal affective behavior, and maternal depression. *Journal of Child Psychology and Psychiatry*, 49, 132–141.
- Fong, M. C., Measelle, J., Conrath, E., & Ablow, J. C. (2017). Links between early baseline cortisol, attachment classification, and problem behaviors: A test of differential susceptibility versus diathesis-stress. *Infant Behavior & Development*, 46, 158–168. doi:10.1016/j.infbeh.2017.01.005
- Gartstein, M. A., & Fagot, B. I. (2003). Parental depression, parenting and family adjustment, and child effortful control: Explaining externalizing behaviors for preschool children. *Journal of Applied Developmental Psychology*, 24, 143–177. doi:10.1016/S0193-3973(03)00043-1
- Ge, X., Conger, R. D., Cadoret, R. J., Neiderhiser, J., Yates, W., Troughton, E., & Stewart, M. A. (1996). The developmental interface between nature and nurture: Mutual influence model of child antisocial behavior and parent behaviors. *Developmental Psychology*, 32, 574–589.
- Ge, X., Natsuaki, M. N., Martin, D. M., Leve, L. D., Neiderhiser, J., Shaw, D. S., ... Reiss, D. (2008). Bridging the divide: Openness in adoption and post-adoption psychosocial adjustment among birth and adoptive parents. *Journal of Family Psychology*, 22, 529–540. doi:10.1037/a0012817
- Goforth, A. N., Pham, A. V., & Carlson, J. S. (2011). Diathesis-stress model. In S. Goldstein & J. A. Naglieri (Eds.), *Encyclopedia of child behavior and development* (pp. 502–503). Boston, MA: Springer.
- Goodman, S. H. (2007). Depression in mothers. *Annual Review of Clinical Psychology*, 3, 107–135. doi:10.1146/annurev.clinpsy.3.022806.091401
- Goodman, S. H., & Gotlib, I. H. (1999). Risk for psychopathology in the children of depressed mothers: A developmental model for understanding mechanisms of transmission. *Psychological Review*, 106, 458–490. doi:10.1037/0033-295X.106.3.458
- Goodman, S. H., Rouse, M. H., Connell, A. M., Broth, M. R., Hall, C. M., & Heyward, D. (2011). Maternal depression and child psychopathology: A meta-analytic review. *Clinical Child and Family Psychology Review*, 14, 1–27. doi:10.1007/s10567-010-0080-1
- Gresham, F., & Elliott, S. (1990). *Social skills rating system manual*. Minneapolis, MN: NCS Pearson, Inc.
- Hankin, B. L., Nederhof, E., Oppenheimer, C. W., Jenness, J., Young, J. F., Abela, J. R. Z., ... Oldehinkel, A. J. (2011). Differential susceptibility in youth: evidence that 5-HTTLPR x positive parenting is associated with positive affect 'for better and worse'. *Translational Psychiatry*, 1, e44.
- Hanscombe, K. B., Trzaskowski, M., Haworth, C. M. A., Davis, O. S. P., Dale, P. S., & Plomin, R. (2012). Socioeconomic status (SES) and children's

- intelligence (IQ): In a UK-representative sample SES moderates the environmental, not genetic, effect on IQ. *PLoS One*, 7, e30320. doi:10.1371/journal.pone.0030320
- Jones, T. L., & Prinz, R. J. (2005). Potential roles of parental self-efficacy in parent and child adjustment: A review. *Clinical Psychology Review*, 25, 341–363. doi:10.1016/j.cpr.2004.12.004
- Kendler, K. S., & Eaves, L. J. (1986). Models for the joint effect of genotype and environment on liability to psychiatric illness. *The American Journal of Psychiatry*, 143, 279–289.
- Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992). Major depression and generalized anxiety disorder. Same genes, (partly) different environments? *Archives of General Psychiatry*, 49, 716–722.
- Kendler, K. S., Prescott, C. A., Myers, J., & Neale, M. C. (2003). The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of General Psychiatry*, 60, 929–937. doi:10.1001/archpsyc.60.9.929
- Kessler, R. C., Chiu, W. T., Demler, O., Merikangas, K. R., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62, 617–627.
- Kretschmer, T., Dijkstra, J. K., Ormel, J., Verhulst, F. C., & Veenstra, R. J. (2013). Dopamine receptor D4 gene moderates the effect of positive and negative peer experiences on later delinquency: The Tracking Adolescents' Individual Lives Survey study. *Development and Psychopathology*, 25, 1107–1117. doi:10.1017/S0954579413000400
- Krueger, R. F., & Markon, K. E. (2006). Reinterpreting comorbidity: A model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology*, 2, 111–133. doi:10.1146/annurev.clinpsy.2.022305.095213
- Laceulle, O. M., Vollebergh, W. A. M., & Ormel, J. (2015). The structure of psychopathology in adolescence: Replication of a general psychopathology factor in the TRAILS study. *Clinical Psychological Science*, 3, 850–860. doi:10.1177/2167702614560750
- 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 Child Psychology*, 121, 971–977.
- 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, 415–422. doi:10.1111/jcpp.12300
- Leve, L. D., Harold, G. T., Ge, X., Neiderhiser, J. M., Shaw, D. S., Scaramella, L. V., & Reiss, D. (2009). Structured parenting of toddlers at high versus low genetic risk: Two pathways to child problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, 48, 1102–1109. doi:10.1097/CHI.0b013e3181b8bfc0
- Leve, L. D., Neiderhiser, J. M., Ganiban, J. M., Natsuaki, M. N., Shaw, D. S., & Reiss, D. (2019). The Early Growth and Development Study: A dual-family adoption study from birth through adolescence. *Twin Research and Human Genetics*, 22, 1–12. doi:10.1017/thg.2019.66
- Leve, L. D., Neiderhiser, J. M., Ge, X., Scaramella, L. V., Conger, R. D., Reid, J. B., ... Reiss, D. (2007). The Early Growth and Development Study: A prospective adoption design. *Twin Research and Human Genetics*, 10, 84–95.
- Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Ganiban, J., Natsuaki, M. N., & Reiss, D. (2012). The Early Growth and Development Study: A prospective adoption study from birth through middle childhood. *Twin Research and Human Genetics*, 16, 412–423. doi:10.1017/thg.2012.126
- Lipscomb, S. T., Leve, L. D., Shaw, D. S., Neiderhiser, J. M., Scaramella, L. V., Ge, X., ... Reiss, D. (2012). Negative emotionality and externalizing problems in toddlerhood: Overreactive parenting as a moderator of genetic influences. *Development and Psychopathology*, 24, 167–179. doi:10.1017/S0954579411000757
- Liu, J. (2004). Childhood externalizing behavior: Theory and implications. *Journal of Child and Adolescent Psychiatric Nursing*, 17, 93–103.
- Lovejoy, M. C., Graczyk, P. A., O'Hare, E., & Neuman, G. (2000). Maternal depression and parenting behavior: A meta-analytic review. *Clinical Psychology Review*, 20, 561–592. doi:10.1016/S0272-7358(98)00100-7
- Luoma, I., Tamminen, T., Kaukonen, P., Laippala, P., Puura, K., Salmelin, R., & Almqvist, F. (2001). Longitudinal study of maternal depressive symptoms and child well-being. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40, 1367–1374. doi:10.1097/00004583-200112000-00006
- Marceau, K., Hajal, N., Leve, L. D., Reiss, D., Shaw, D. S., Ganiban, J. M., ... Neiderhiser, J. M. (2013). Measurement and associations of pregnancy risk factors with genetic influences, postnatal environmental influences, and toddler behavior. *International Journal of Behavioral Development*, 37, 366–375.
- Marmorstein, N. R., Iacono, W. G., & McGue, M. (2012). Associations between substance use disorders and major depression in parents and late adolescent-emerging adult offspring: An adoption study. *Addiction*, 107, 1965–1973. doi:10.1111/j.1360-0443.2012.03934.x
- Mathis, E. T., & Bierman, K. L. (2015). Dimensions of parenting associated with child prekindergarten emotion regulation and attention control in low-income families. *Social Development*, 24, 601–620. doi:10.1111/sode.12112
- McCoy, K., Cummings, E. M., & Davies, P. T. (2009). Constructive and destructive marital conflict, emotional security and children's prosocial behavior. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 50, 270–279. doi:10.1111/j.1469-7610.2008.01945.x
- Meadows, S. O., McLanahan, S. S., & Brooks-Gunn, J. (2007). Parental depression and anxiety and early childhood behavior problems across family types. *Journal of Marriage and Family*, 69, 1162–1177. doi:10.1111/j.1741-3737.2007.00439.x
- Melby, J. N., Conger, R. D., Ge, X., & Warner, T. (1995). The use of structural equation modeling in assessing the quality of marital observations. *Journal of Family Psychology*, 9, 280–293. doi:10.1037/0893-3200.9.3.280
- Messer, B. J., & Harter, S. (1986). *Manual for the adult self-perception profile*. Denver, CO: University of Denver.
- Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: Implications for the depressive disorders. *Psychological Bulletin*, 110, 406–425.
- Murray, A. L., Eisner, M., & Ribeaud, D. (2016). The development of the general factor of psychopathology 'p Factor' through childhood and adolescence. *Journal of Abnormal Child Psychology*, 44, 1573–1586. doi:10.1007/s10802-016-0132-1
- Murray, A. L., & Johnson, W. (2013). The limitations of model fit in comparing the bi-factor versus higher-order models of human cognitive ability structure. *Intelligence*, 41, 407–422. doi:10.1016/j.intell.2013.06.004
- Muthén, L., & Muthén, B. (2019). *Mplus. The comprehensive modelling program for applied researchers: user's guide*, 5. Muthén & Muthén.
- Natsuaki, M. N., Ge, X., Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Conger, R. D., ... Reiss, D. (2010). Genetic liability, environment, and the development of fussiness in toddlers: The roles of maternal depression and parental responsiveness. *Developmental Psychology*, 46, 1147–1158. doi:10.1037/a0019659
- Natsuaki, M. N., Shaw, D. S., Neiderhiser, J. M., Ganiban, J. M., Harold, G. T., Reiss, D., & Leve, L. D. (2014). Raised by depressed parents: Is it an environmental risk? *Clinical Child and Family Psychology Review*, 17, 357–367. doi:10.1007/s10567-014-0169-z
- Neiderhiser, J. M., Marceau, K., De Araujo-Greecher, M., Ganiban, J. M., Mayes, L. C., Shaw, D. S., ... Leve, L. D. (2016). Estimating the roles of genetic risk, perinatal risk, and marital hostility on early childhood adjustment: Medical records and self-reports. *Behavior Genetics*, 46, 334–352. doi:10.1007/s10519-016-9788-0
- 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, 1038–1045.e1034. doi:10.1016/j.jaac.2016.09.498
- Newman, D. L., Moffitt, T. E., Caspi, A., & Silva, P. A. (1998). Comorbid mental disorders: Implications for treatment and sample selection. *Journal of Abnormal Psychology*, 107, 305–311.
- O'Connor, T. G., Caspi, A., Defries, J. C., & Plomin, R. (2003). Genotype-environment interaction in children's adjustment to parental separation. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 44, 849–856.
- Owen, M. T., & Cox, M. J. (1997). Marital conflict and the development of infant-parent attachment relationships. *Journal of Family Psychology*, 11, 152–164. doi:10.1037/0893-3200.11.2.152

- Patalay, P., Fonagy, P., Deighton, J., Belsky, J., Vostanis, P., & Wolpert, M. (2015). A general psychopathology factor in early adolescence. *The British Journal of Psychiatry*, 207, 15.
- Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, 84(2), 309–322. <https://doi.org/10.1037/0033-2909.84.2.309>
- Pluess, M., & Belsky, J. (2010). Differential susceptibility to parenting and quality child care. *Developmental Psychology*, 46, 379–390. doi:10.1037/a0015203
- Raine, A. (2002). Biosocial studies of antisocial and violent behavior in children and adults: A review. *Journal of Abnormal Child Psychology*, 30, 311–326.
- Ramchandani, P. G., O'Connor, T. G., Evans, J., Heron, J., Murray, L., & Stein, A. (2008). The effects of pre- and postnatal depression in fathers: A natural experiment comparing the effects of exposure to depression on offspring. *Journal of Child Psychology and Psychiatry*, 49, 1069–1078. doi:10.1111/j.1469-7610.2008.02000.x
- Reiss, D. (2016). Genetics and family systems: articulation and disarticulation. In D. Cicchetti (Ed.), *Developmental psychopathology* (3rd ed., Vol. 2, pp. 339–377). New York, NY: Wiley.
- Reiss, D., Leve, L. D., & Neiderhiser, J. (2013). How genes and the social environment moderate each other. *American Journal of Public Health*, 103, S111–S121. doi:10.2105/AJPH.2013.301408
- Respler-Herman, M., Mowder, B. A., Yasik, A. E., & Renee Shamah, R. (2012). Parenting beliefs, parental stress, and social support relationships. *Journal of Child Family Studies*, 21, 190–198. doi:10.1007/s10826-011-9462-3
- Rhoades, K. A., Leve, L. D., Harold, G. T., Neiderhiser, J. M., Shaw, D. S., & Reiss, D. (2011). Longitudinal pathways from marital hostility to child anger during toddlerhood: Genetic susceptibility and indirect effects via harsh parenting. *Journal of Family Psychology*, 25, 282–291. doi:10.1037/a0022886
- Roben, C. K., Moore, G. A., Cole, P. M., Molenaar, P., Leve, L. D., Shaw, D. S., ... Neiderhiser, J. M. (2015). Transactional patterns of maternal depressive symptoms and mother-child mutual negativity in an adoption sample. *Infant and Child Development*, 24, 322–342. doi:10.1002/icd.1906
- Robins, L. N., Helzer, J. E., Croughan, J., & Ratcliff, K. S. (1981). National institute of mental health diagnostic interview schedule. Its history, characteristics, and validity. *Archives of General Psychiatry*, 38, 381–389.
- Rosenthal, D., Wender, P. H., Kety, S. S., Schulsinger, F., Welner, J., & Østergaard, L. (1968). Schizophrenics' offspring reared in adoptive homes. *Journal of Psychiatric Research*, 6, 377–391. doi:10.1016/0022-3956(68)90028-9
- Rutter, M., Moffitt, T. E., & Caspi, A. (2006). Gene-environment interplay and psychopathology: Multiple varieties but real effects. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 47, 226–261. doi:10.1111/j.1469-7610.2005.01557.x
- SAS Institute Inc. (2013). *SAS/STAT Software (Version 9.4)*. Cary, NC: SAS Institute Inc.
- Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype greater than environment effects. *Child Development*, 54, 424–435. doi:10.2307/1129703
- Scarr, S., & Weinberg, R. A. (1977). Intellectual similarities within families of both adopted and biological children. *Intelligence*, 1, 170–191.
- Scarr, S., & Weinberg, R. A. (1983). The Minnesota adoption studies: Genetic differences and malleability. *Child Development*, 54, 260–267. doi:10.2307/1129689
- Schultz, D., & Shaw, D. (2003). Boys' maladaptive social information processing, family emotional climate, and pathways to early conduct problems. *Social Development*, 12, 440–460.
- Selzam, S., Coleman, J. R., Caspi, A., Moffitt, T. E., & Plomin, R. (2018). A polygenic p factor for major psychiatric disorders. *Translational Psychiatry*, 8, 205.
- Semrud-Clikeman, M. (2007). *Social competence in children*. New York, NY: Springer.
- Sjöberg, R. L., Nilsson, K. W., Nordquist, N., Öhrvik, J., Leppert, J., Lindström, L., & Oreland, L. (2006). Development of depression: Sex and the interaction between environment and a promoter polymorphism of the serotonin transporter gene. *International Journal of Neuropsychopharmacology*, 9, 443–449.
- Stover, C. S., Connell, C. M., Leve, L. D., Neiderhiser, J. M., Shaw, D. S., Scaramella, L. V., ... Reiss, D. (2012). Fathering and mothering in the family system: Linking marital hostility and aggression in adopted toddlers. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 53, 401–409. doi:10.1111/j.1469-7610.2011.02510.x
- 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, 1142–1153. doi:10.1037/a0034151
- Teti, D. M., & Gelfand, D. M. (1991). Behavioral competence among mothers of infants in the first year: The mediational role of maternal self-efficacy. *Child Development*, 62, 918–929. doi:10.1111/j.1467-8624.1991.tb01580.x
- Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1–10. doi:10.1007/BF02291170
- Tully, E. C., Iacono, W. G., & McGue, M. (2008). An adoption study of parental depression as an environmental liability for adolescent depression and childhood disruptive disorders. *The American Journal of Psychiatry*, 165, 1148–1154. doi:10.1176/appi.ajp.2008.07091438
- Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, 14, 623–628.
- van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, 16, 219–242. doi:10.1177/0962280206074463
- van Ijzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2015). Genetic differential susceptibility on trial: Meta-analytic support from randomized controlled experiments. *Development and Psychopathology*, 27, 151–162.
- van Ijzendoorn, M. H., Belsky, J., & Bakermans-Kranenburg, M. J. (2012). Serotonin transporter genotype 5HTTLPR as a marker of differential susceptibility? A meta-analysis of child and adolescent gene-by-environment studies. *Translational Psychiatry*, 2, e147. doi:10.1038/tp.2012.73
- Von Hippel, P. T. (2009). How to impute interactions, squares, and other transformed variables. *Sociological Methodology*, 39, 265–291. doi:10.1111/j.1467-9531.2009.01215.x
- Wahlberg, K. E., Wynne, L. C., Oja, H., Keskitalo, P., Pykalainen, L., Lahti, I., ... Tienari, P. (1997). Gene-environment interaction in vulnerability to schizophrenia: Findings from the Finnish Adoptive Family Study of Schizophrenia. *The American Journal of Psychiatry*, 154, 355–362.
- Waldman, I. D., Poore, H. E., van Hulle, C., Rathouz, P. J., & Lahey, B. B. (2016). External validity of a hierarchical dimensional model of child and adolescent psychopathology: Tests using confirmatory factor analyses and multivariate behavior genetic analyses. *Journal of Abnormal Psychology*, 125, 1053–1066. doi:10.1037/abn0000183
- Ware, J. E., Kosinski, M., & Dewey, J. E. (2000). *How to score version 2 of the SF-36 health survey (standard & acute forms)*. Lincoln, RI: QualityMetric Incorporated.
- Widaman, K. F., Helm, J. L., Castro-Schilo, L., Pluess, M., Stallings, M. C., & Belsky, J. (2012). Distinguishing ordinal and disordinal interactions. *Psychological Methods*, 17, 615–622.
- Wilson, S., & Durbin, C. E. (2010). Effects of paternal depression on fathers' parenting behaviors: A meta-analytic review. *Clinical Psychology Review*, 30, 167–180. doi:10.1016/j.cpr.2009.10.007
- Wright, A. G. C., Krueger, R. F., Hobbs, M. J., Markon, K. E., Eaton, N. R., & Slade, T. (2013). The structure of psychopathology: Toward an expanded quantitative empirical model. *Journal of Abnormal Psychology*, 122, 281–294.