Genetic moderation of the effects of the Family Check-Up intervention on children's internalizing symptoms: A longitudinal study with a racially/ethnically diverse sample

KATHRYN LEMERY-CHALFANT,[a] SIERRA CLIFFORD,[a] THOMAS J. DISHION,[a] DANIEL S. SHAW,[b] AND MELVIN N. WILSON[c]
[a]*Arizona State University;* [b]*University of Pittsburgh;* and [c]*University of Virginia*

Abstract

Development involves synergistic interplay among genotypes and the physical and cultural environments, and integrating genetics into experimental designs that manipulate the environment can improve understanding of developmental psychopathology and intervention efficacy. Consistent with differential susceptibility theory, individuals can vary in their sensitivity to environmental conditions including intervention for reasons including their genotype. As a consequence, understanding genetic influences on intervention response is critical. Empirically, we tested an interaction between a genetic index representing sensitivity to the environment and the Family Check-Up intervention. Participants were drawn from the Early Steps Multisite randomized prevention trial that included a low-income and racially/ethnically diverse sample of children and their families followed longitudinally ($n = 515$). As hypothesized, polygenic sensitivity to the environment moderated the effects of the intervention on 10-year-old children's symptoms of internalizing psychopathology, such that children who were genetically sensitive and were randomly assigned to the intervention had fewer symptoms of child psychopathology than genetically sensitive children assigned to the control condition. A significant difference in internalizing symptoms assessed with a clinical interview emerged between the intervention and control groups for those 0.493 *SD* above the mean on polygenic sensitivity, or 25% of the sample. Similar to personalized medicine, it is time to understand individual and sociocultural differences in treatment response and individualize psychosocial interventions to reduce the burden of child psychopathology and maximize well-being for children growing up in a wide range of physical environments and cultures.

Development involves synergistic interplay among genotypes and the physical and sociocultural environments. These influences are interdependent, with reciprocal causal pathways and feedback, correlations, and interactions across levels of analysis. Until recently, the ability to study such transactions has been limited by technology. With the completion of the Human Genome Project (International Human Genome Sequencing Consortium, 2004), development of new methods of measuring brain architecture and functioning (Roalf & Gur, 2017), and the birth of new fields of science (e.g., epigenetics and proteomics), levels of analysis that previously existed in "the black box" can now be measured and developmental theories integrating genetic and sociocultural levels of analysis can be rigorously tested. Such research has the potential not only to advance our understanding of child development but also to improve the design and testing of interventions. Because individuals vary in their sensitivity to psychosocial interventions for reasons including their genotype, understanding genetic influences on intervention response is critical for an accurate judgment of efficacy. However, genetically informed intervention research is still rare and largely based on a handful of candidate genes (e.g., Bakermans-Kranenburg & van IJzendoorn, 2015). It is likely that psychosocial interventions have large effects on developmental psychopathology for some individuals, but no significant effects for other children. When we do not acknowledge developmental theory and only consider main effects of intervention and effectively average responses across all individuals in the sample, we may erroneously conclude that interventions have small and sometimes nonsignificant effects. A tailored and personalized medicine approach is needed to identify which interventions work for which people.

Further, genetically informed research as a whole still lags behind in the area of cultural and ethnic diversity, with most studies relying on samples of Northern European descent and extremely limited representation of Latino, Indigenous, and African or African American participants (Popejoy & Fullerton, 2016). Given that both genetic and environmental influences can differ across populations, genetics research based on participants from a narrow range of cultures and ethnicities cannot be assumed to generalize to humanity as a whole, and in particular, to racial and ethnic minorities who are often left

out of research participation (Oquendo, Canino, Lehner, & Licinio, 2010).

We address these limitations by testing genetic moderation of the effects of a parenting-based intervention on internalizing symptoms in a racially/ethnically diverse sample of US children recruited based on economic and family risk, using a polygenic score based on an existing genome-wide association study intended to tap sensitivity to the environment (Keers et al., 2016). Participants underwent a randomized control trial with the Family Check-Up, a family-based intervention that incorporates motivational interviewing that has been successfully utilized to reduce child problem behavior across a wide range of socioeconomic and sociocultural groups (Dishion, Kavanagh, Schneiger, Nelson, & Kaufman, 2002; Dishion et al., 2008; Shaw, Dishion, Supplee, Gardner, & Arnds, 2006). We indexed genetic sensitivity to the environment using a polygenic score based on genetic variants that were associated with identical twin differences in internalizing symptoms in a previous genome-wide association study (Keers et al., 2016). Because it was derived from the prediction of identical twin differences, this score is an aggregation, not of small genetic main effects, but of variants associated with the magnitude of response to nonshared environmental factors across all environments the twins encounter, in essence acting as a genetic index of sensitivity to the environment broadly defined. Accordingly, we hypothesized that the Family Check-Up intervention would have stronger effects on internalizing psychopathology for children who were higher on the polygenic sensitivity score.

Theoretical Background

Developmental theories suggest that we cannot understand genetic influences on individual behavior and outcomes without simultaneous consideration of sociocultural influences, and conversely, that we cannot understand sociocultural influences without understanding genetic influences. Yet cultural and genetic fields of study are still very much distinct, with few collaborations between cultural and genetic scientists, and no established training programs that combine these theoretical perspectives and methods. Behavioral genetics focuses on how genotypes are related to phenotypes, with *phenotypes* defined as observable characteristics and behaviors that are the result of both genes and environments. Cultural psychologists often study relations between components of culture and individual phenotypes, with *culture* defined as a system of behaviors and cognitions shared by a community that informs values, goals, practices, traditions, and institutions; this knowledge is transmitted from one generation to the next (Cohen, 2009). The emerging field of cultural genomics combines these areas by studying the interplay of genomics, cultures, physical environments, and phenotypes (Causadias, Telzer, & Gonzales, 2018).

According to this framework, individual development is dynamically shaped by families and communities (e.g., race, religion, and social group), creating unique realities across cultures

(Causadias et al., 2018). At the same time, person-level factors including genetic variation can influence the way individuals actively construct and interact with their own cultures and environments, with all of these factors mutually influencing each other over both developmental and evolutionary time. Genetic variation also enables humans to adapt to large geographic ranges and changes in the environment, with a genotype that may be disadvantageous in one environment potentially offering an advantage in another (e.g., the genetic mutation that causes sickle cell anemia is protective against malaria), and natural selection maintaining genetic variation when there is high variability in environments. Further, both the strength and the nature of genetic effects can vary considerably for individuals with similar genotypes based on factors including the physical environment and access to resources, social and cultural influences, and individual experiences (Shanahan & Hofer, 2005).

Transactions between genes and environments can take many forms (Shanahan & Hofer, 2005). Genetic influences on a phenotype may become a more important source of individual variation when socialization does not constrain expression of the phenotype (e.g., generational increases in the heritability of Swedish women's tobacco use; Kendler, Thornton, & Pedersen, 2000). Many heritable traits are not realized without environmental support, such that intelligence is on average both lower and less heritable in children growing up in extreme poverty in the United States (Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003). Conversely, some genetic vulnerabilities may be mitigated by high levels of environmental support or enrichment even when they would manifest under typical conditions (Shanahan & Hofer, 2005), and some genetic variants may increase sensitivity to environmental stress (i.e., diathesis-stress; Rende & Plomin, 1992), or to both stressful and positive sociocultural factors as a whole (e.g., differential susceptibility; Ellis, Boyce, Belsky, Bakermans-Kranenburg, & van Ijzendoorn, 2011), such that genetic risk manifests only as a function of the environment. Although behavioral genetics research has increasingly acknowledged the importance of gene–environment interplay and nonadditivity, research integrating insights from a cultural genomics perspective is still very limited.

The largest barrier to adopting a cultural genomics approach is study design and measurement. Behavior geneticists have adopted study designs with very large racially and ethnically homogeneous samples, extensive genotyping, but very limited phenotypic and environmental measurement. These behavior genetics designs are supported by outdated "medical models" that suggest phenotypes are directly caused by genetic mutations, and that this direct effect does not vary across physical and sociocultural environments. Conversely, cultural scientists have adopted study designs with small samples, rich environmental and phenotypic measurement, but no or very limited genotyping. These cultural designs are supported by models that suggest interplay across levels of analysis (Bronfenbrenner & Ceci, 1994), but the impact of children's own biology is discounted. Variations and limitations in study design and mea-

surement have prohibited interdisciplinary collaborations across cultural and behavior genetics scientists, generated from differences in perspective and the unique methodological challenges facing each field. For example, differences in allelic frequencies across cultural groups must be accounted for in genetics research to avoid spurious associations, making research with culturally diverse samples challenging. In addition, small genetic effects require large sample sizes for adequate power. However, with continuing advances in genotyping and genetic methodology, as well as the replication and extension of existing behavioral genetics research in more diverse samples with richer cultural measurement, the strengths of the two fields can be combined.

One major issue facing behavior genetics research today is the lack of racial and ethnic diversity in genetically informed studies, perhaps especially those taking a molecular genetics approach. Although cultural genomic theorists caution against using race and ethnicity as proxies for culture, they emphasize their importance as culturally linked social factors with the potential to shape individual and social-level experiences (e.g., identity formation and discrimination), and highlight the need to increase representation of racial and ethnic minorities in research (Causadias, 2013). Unfortunately, although matters have improved in recent years, molecular genetics research is still largely based on a narrow range of samples of European, and in particular, Northern European ancestry. A 2016 study found that 81% of the participants included in the genome-wide association study were of European ancestry, with much of the remaining 19% accounted for by participants of Asian ancestry living in Asian countries, and participants of African, Latino, and Native or Indigenous ancestry comprising less than 4% of the total (Popejoy & Fullterton, 2016). This lack of diversity in genetics research is a problem for several reasons. First, it is unfair to people of racial and ethnic minority groups themselves, as it will result in failure to detect and account for genetic risk factors that are more common or only present in these groups, or, just as important, differences in drug safety and treatment efficacy. Second, it is a wasted opportunity for elucidating genetic risk for disorders that do differ across populations, and runs the risk of filtering out novel variants that have strong effects on a phenotype but happen to be rare in European populations. Third, findings from samples of European and European American participants may not replicate in other racial and ethnic groups, and this nonreplication may be a sign of false-positive findings. As such, failure to address this bias not only will result in unequal distribution of the potential benefits of genetics research but also is detrimental to the quality of the research itself.

Improving the representation of racial and ethnic minorities in genetics research will require a concentrated effort to recruit and retain participants from backgrounds other than Northern European, including the use of culturally sensitive methods of recruitment and interaction with participants, as well as strategies such as oversampling or the use of multisite studies to recruit demographically representative samples of individuals from less well-represented groups (Oquendo et al., 2009). In addition, as noted earlier, accounting for genetic diversity across populations presents a statistical chal-

lenge for molecular genetics research. Addressing this challenge will require genotyping and imputation that is sensitive to variation in allelic frequencies and haplotypes across populations, as well as statistical methods of accounting for population stratification (Oquendo et al., 2009). The present Early Steps Multisite study was designed to recruit low-income racially/ethnically diverse families from urban, suburban, and rural regions of the United States, who also scored high on family or child risk factors, to better represent children at risk of developing psychopathology.

Developmental Cascades of Externalizing and Internalizing Psychopathology

Internalizing and externalizing symptoms are moderately to highly positively correlated across childhood, with support for a general shared genetic etiology, and more evidence of externalizing in childhood leading to later internalizing problems than vice versa. Specifically, McDonough-Caplan, Klein, and Beauchaine (2018) report that externalizing problems in childhood lead to comorbid internalizing/externalizing problems in adolescence, but internalizing problems in childhood did not increase risk of later externalizing problems. Similarly, Moilanen, Shaw, and Maxwell (2010) tested developmental cascades of internalizing and externalizing problems across childhood in low-income boys, controlling for moderate to high autoregressive associations. They reported that higher externalizing at ages 6 and 11 predicted higher internalizing at ages 8 and 12, respectively. There is also evidence of bidirectional influences between parental sensitivity and child psychopathology, with child externalizing problems affecting maternal sensitivity, and maternal sensitivity influencing later internalizing problems in females (Zvara, Sheppard, & Cox, 2018). Paternal sensitivity was reciprocally related to both internalizing and externalizing problems.

An intervention that focuses on increasing parents' positive behavioral management such as the Family Check-Up is thought to decrease children's internalizing symptoms as well as externalizing problem behaviors because of shared risk factors, including the long-studied association between parenting and children's psychopathology. The present study focuses on children's internalizing symptoms utilizing a structured clinical interview in the home with the children in middle childhood. A recent meta-analysis concluded that parent harsh and psychological control predicted increases in internalizing symptoms, whereas authoritative parenting, autonomy granting, behavioral control, and warmth predicted decreases in internalizing symptoms across childhood, with some of these associations being bidirectional (Pinquart, 2017).

Parenting as a Target of Intervention

Perhaps the risk/protective factor that has received the most attention across cultures in the field of developmental psychopathology is parenting. Parenting is a proximal process impacting child psychopathology, mediating the effects of family

socioeconomic adversity (Dodge, Pettit, & Bates, 1994), and parent psychopathology (Harold, Rice, Hay, Boivin, van Den Bree, & Thapar, 2011). Parenting, although partially genetically influenced, has been related to child behavior after accounting for genetic influences (Oliver, Trzaskowski, & Plomin, 2014). Research across cultures suggests that some factors, such as warm, positive parenting, are associated with positive child adjustment, whereas detached or abusive parenting is related to maladjustment (Smith, Knoble, Zerr, Dishion, & Stormshak, 2014). It is these factors that the Family Check-Up intervention aims to target, in a way that is sensitive to family and cultural differences.

The Family Check-Up intervention

The Family Check-Up was developed as an intervention framework that is flexible and adaptive to diverse cultural groups and is individually tailored to each family context (Dishion & Stormshak, 2007). The original purpose of the intervention was to reduce oppositional and aggressive behavior by improving positive parent management skills, especially during times of developmental transition (e.g., early adolescence or the "terrible twos"). However, it has also been shown to have positive effects on children's broader development, including increases in inhibitory control and verbal skills (Lunkenheimer et al., 2008) and decreases in symptoms of internalizing disorders (Shaw, Connell, Dishion, Wilson, & Gardner, 2009). The Family Check-Up has been tested in a series of randomized controlled trials from ages 2 to 18 years and found to be effective in the prevention of internalizing and externalizing problem behaviors (e.g., Connell & Dishion, 2017; Dishion et al., 2002, 2008; Stormshak et al., 2011; Stormshak, Fosco, & Dishion, 2010; Van Ryzin, Stormshak, & Dishion, 2012). The menu of service options fits within a variety of cultural frameworks, with a focus on contextual stressors and parental factors that may lead to the emergence and maintenance of child psychopathology based on coercive family processes (Patterson, Reid, & Dishion, 1992) and social learning theory.

In general practice, the Family Check-Up begins with an initial get-to-know-you interview during which time the family coach finds out about the parents' strengths and challenges, their aspirations for their child, and their family values (30–40 min). This is followed by a brief home-based ecological assessment to formally assess caregiving and child and family functioning, the focus of which is tailored to the child's developmental status. Using data from the assessment and initial interview, the caregiver(s) and family coach then meet for a feedback session (1.25–1.5 hr), during which family coach provides the parents with information on family and child functioning relative to normative data. To retain blindness, in research studies such as the present one, the assessment precedes the initial interview, to minimize potential bias resulting from the intervention group having a session prior to the assessment (i.e., both control and intervention families have the same assessment). Family stories at the

core of family intervention is important for the majority of cultural groups in the United States (McGoldrick & Hardy, 2008), and during the initial interview, family coaches ask open-ended questions to foster a trusting relationship and give caregivers an opportunity to tell their family story. Caregivers are thus acknowledged as the respected authority on their children and their family, which is culturally congruent across diverse families. Within this collaborative framework, the family coach's questions illuminate contextual factors that contribute to children's mood and behavior problems (e.g., family roles or discrimination stress) and motivate the caregiver to change family management strategies. The feedback sessions also are adapted to focus on both parenting strengths and challenges within the cultural context. The extent to which the Family Check-Up is effective in improving parenting and children's adjustment across European American, African American, and Latino groups has been empirically evaluated, with participation in the Family Check-Up reducing antisocial behavior through reducing family conflict across all three groups (Smith et al., 2014).

Gene × Intervention interactions

There is wide variation in the efficacy of psychological preventive interventions such as the Family Check-Up, with response to treatment varying substantially between individuals, and genetics are likely to be one source of these individual differences. Understanding the role of genetics elucidates mechanisms underlying treatment response, and identifying genetic predictors of treatment response allows one to match treatment to individuals at the outset to improve outcomes, such as in personalized medicine. Accordingly, a parenting-based intervention such as the Family Check-Up could work best with children who are genetically sensitive to the environment. Conversely, a more cognitively based intervention may work best with those who are less sensitive to contexts. In addition, randomizing participants to an intervention is one of the best ways to understand interactions between genetics and the physical and sociocultural environments. Testing gene–environment interaction in the context of random assignment to an intervention increases power because of the absence of *gene–environment correlation*, or the extent to which the genetic score is correlated with the environmental exposure (Plomin, DeFries, & Loehlin, 1977). Gene–environment correlation is widespread, as individuals live in environments that are partially created by kin (*passive gene–environment correlation*). In addition, they evoke differential responses from others based on their heritable characteristics (*evocative gene–environment correlation*), and they niche pick, or actively seek out environments that match their heritable traits (*active gene–environment correlation*). Experimental designs with random assignment of children and families to treatment and control conditions eliminates confounding because of gene–environment correlation, affording stronger causal inference. To that end, Bakermans-Kranenburg and van IJzendoorn (2015) suggest that statistical

power of studies that consider genetic moderation of intervention effects is much higher than correlational studies that require up to 13 times more participants to reach similar levels of power.

Molecular Behavior Genetics Methodology

The majority of genetic effects in humans are homogeneous and do not contribute to individual differences. However, dbSNP (www.ncbi.nlm.nih.gov/projects/SNP/) contains over 12 million single nucleotide polymorphisms (SNPs) that vary across individuals, with some differences in *allele* (i.e., specific genetic variant) frequencies across cultural groups. With 23 paired chromosomes (1 from each parent), each individual has two copies of each strand of DNA, and thus carry 0, 1, or 2 copies of a particular allele at a SNP (e.g., AA, AG, or GG nucleotides make up the genotype at one SNP across individuals). Early molecular genetics research focused on examining relations between single genetic polymorphisms in *candidate genes*, chosen for their theoretical relevance, and outcomes such as psychopathology (see Bakermans-Kranenburg & van IJzendoorn, 2011; van IJzendoorn, Belsky, & Bakermans-Kranenburg, 2012, for meta-analyses). However, this approach has multiple limitations. Genetic influences of single common genetic variants on complex psychological traits are likely to be small, especially as psychopathology is biologically distal from an intra- or intercellular genetic effect. As a consequence, most candidate gene research has been underpowered to detect true effects. The lack of power, combined with multiple testing across studies and *the file drawer effect* (i.e., studies with nonsignificant findings are likely not published), has led to a high false-positive rate in the field as a whole (Ioannidis, 2005).

To address the high false-positive rate, behavior genetics research has turned increasingly to the genome-wide association study, a hypothesis-free analysis of the additive predictive power of each individual SNP across the genome on a particular outcome (i.e., presence or absence of a diagnosis, or a composite of symptoms). Because so many statistical tests are performed in a genome-wide association study, the significance threshold is typically set at a stringent 5×10^{-8} ($p < .00000005$) to statistically correct for the number of tests, rather than the standard 5×10^{-2} ($p < .05$). Because we expect very small effect sizes for individual SNPs, using a *polygenic score* to aggregate the small effects of SNPs identified in a discovery genome-wide association study has utility. Typically, results from a discovery genome-wide association study are used to identify a large group of SNPs based on their individual association with the outcome of interest, using discovery genome-wide association study cutoffs of $p < .001$, $p < .01$, $p < .05$, $p < .1$, and $p < .5$, or all SNPs studied. The so-called risk alleles are then weighted based on their effect size in the discovery genome-wide association study and summed to produce a quantitative polygenic score for use in an independent validation sample. However, like candidate gene research, the genome-wide association study has its limitations, some of which make it challenging to combine with a cultural approach. For example, a genome-wide association study is designed to detect small main effects of SNPs, without regard for interaction with the environment, whereas many models of gene–environment interplay (e.g., diathesis-stress and differential susceptibility) posit that some genetic variants are associated with an outcome *only* in particular environmental contexts. In addition, the hypothesis-free multiple testing and small effects lead to both high Type 1 and Type II error, and aggregation into polygenic scores mitigates but does not eliminate the false-positive problems. Thus, the findings from discovery genome-wide association studies should be taken as a first step from which other research can draw on to more clearly elucidate biological and sociocultural pathways. In addition, although genome-wide association studies are designed for the detection of small additive effects, innovative research designs such as the prediction of identical twin differences (Keers et al., 2016) make it possible to address gene–environment interplay within a genome–wide association study framework.

Empirical Support of Genetic Moderation of Intervention Effects

As with gene–environment interplay as a whole, most research testing genetic moderation of randomized intervention has been limited to studies of a few candidate genes (e.g., serotonin transporter gene, serotonin receptors, dopamine transporter gene, and dopamine receptors). One meta-analysis of more than 20 experiments supports genetic moderation of the effect of family-based intervention on externalizing problems, but findings for internalizing were more equivocal (Bakermans-Kranenburg & van IJzendoorn, 2015), and effect sizes are likely inflated by the combination of underpowered individual studies and publication bias. Nevertheless, to the extent that these findings prove replicable in adequately powered samples, genetic moderation of intervention is monumental in that it suggests that the effects of early intervention are likely underestimated or go undetected for children whose genotypes support higher environmental sensitivity.

To date, there are only two studies that consider moderation of intervention effects utilizing polygenic scores based on findings from a previous genome-wide association study rather than focusing on one or a few candidate genes (Keers et al., 2016; Musci et al., 2015, 2018). Musci et al. found a Polygenic Risk × Intervention interaction predicting age of first tobacco use and age of first marijuana use in a sample of 556 primarily African American individuals followed from sixth grade to age 18 (Musci et al., 2015, 2018). Specifically, Musci et al. report that a classroom-based behavioral intervention targeting aggressive and disruptive behavior was most strongly associated with later onset of smoking and marijuana use for individuals scoring high on a polygenic score based on 12,058 SNPs previously associated with smoking cessation and lower substance use (alcohol, marijuana, and tobacco) in adults.

Rather than focus on risk, Keers et al. (2016) for the first time formed a polygenic sensitivity to the environment score from SNPs that predicted identical twin differences in childhood emotional problems in a discovery genome-wide association study, and found that it moderated the effects of cognitive behavioral therapy on emotional disorders in children. The sample of twins was 93% White, which is representative of the population of the United Kingdom (Haworth, Davis, & Plomin, 2013). It is this research that we draw on in the present study, with the goal of forming a polygenic score that includes SNPs that differentiate individuals who are more or less sensitive to the physical and sociocultural environments.

Polygenic Sensitivity Score Based on Identical Twin Differences

Unlike other genome-wide association studies that aim to detect genetic variants associated with main effects on a phenotype, Keers et al. (2016) took advantage of the unique nature of an identical twin sample to pinpoint genetic variants associated with sensitivity to the environment itself. Because identical twins are genetically identical and share the same family environment, all differences between identical twins in a pair are due to environmental factors not shared between twin siblings, controlling for all genetic main effects. Identical twin differences can result not only from environmental main effects but also from interaction between genetic factors and the nonshared environment, such that genes that increase sensitivity to broad or specific environmental factors will also act to increase differences between identical twins whose experiences of those environments differ. Using this logic, Keers et al. (2016) predicted identical twin differences in emotional disorder symptoms at age 12 as the outcome of a discovery genome-wide association study, allowing for the identification of SNPs associated with sensitivity to the environment while simultaneously controlling for all other genetic and shared environmental effects.

The discovery sample included 1,026 identical twin pairs from the United Kingdom who participated in the Twins Early Development Study. At 12 years of age, the Strengths and Difficulties Questionnaire was given to parents and children, and a mean composite of the standardized emotional symptoms scale was used to index emotional problems. Identical twin differences in emotional problems were operationally defined as the absolute difference in scores between cotwins, after regressing out age, sex, and the twin pair's mean score on emotional symptoms. Linear regressions using the software PLINK were conducted, covarying the first 10 principal components representing population admixture, and p values were obtained for each SNP examined.

Based on this discovery genome-wide association study, a polygenic sensitivity score was formed in a separate Twins Early Development Study sample of 1,406 individuals to examine polygenic moderation of parenting effects on children's emotional problems. As is standard practice, multiple significance thresholds were used to create eight polygenic

sensitivity scores from a total of 155,019 SNPs (after linkage disequilibrium pruning, or dropping some SNPs because they were highly correlated with others) at $p < .001$ ($n = 400$ SNPs), $p = .01$ ($n = 3,161$), $p = .05$ ($n = 13,632$), $p = .1$ ($n = 25,384$), $p = .2$ ($n = 46,752$), $p = .3$ ($n = 66,205$), $p = .4$ ($n = 84,025$), and $p = .5$ ($n = 100,111$). Polygenic sensitivity moderated the influence of parenting on emotional problems, with five of the eight polygenic scores formed ($p < .1$ through $p < .5$) yielding significant interactions. For those with higher polygenic sensitivity, positive parenting was associated with decreased emotional problems and negative parenting was associated with more emotional problems. At lower levels of polygenic sensitivity, parenting was unrelated to emotional problems.

Next, with the Genes for Treatment sample, a polygenic sensitivity score was formed in the same fashion and was used to predict response to cognitive behavioral therapy in children with anxiety disorders (Keers et al., 2016). Genes for Treatment includes 973 children who met DSM-IV criteria for a primary diagnosis of an anxiety disorder and were genotyped. After linkage disequilibrium pruning, there were 72,375 SNPs in common between the Twins Early Development Study discovery sample and Genes for Treatment, so the polygenic sensitivity scores were smaller. The polygenic scores at thresholds of $p < .05$ and above significantly moderated treatment response such that individual cognitive behavioral therapy (compared to group cognitive behavioral therapy or brief parent-led cognitive behavioral therapy) had a larger effect for children with higher polygenic sensitivity scores, although there was no moderation of overall treatment response.

The Present Study

We formed polygenic sensitivity scores based on Keers et al.'s (2016) discovery genome-wide association study results associating SNPs with identical twin differences in emotional problems, and then used these scores to examine genetic moderation of intervention effects in a high economic and family risk and culturally diverse US sample. The overarching goal was to examine whether a genetic index of environmental sensitivity moderated the effects of the Family Check-Up intervention on symptoms of internalizing psychopathology in middle childhood. The sample was the Early Steps Multisite Study, a large randomized controlled trial of the Family Check-Up in early childhood that was designed to reduce aggressive and oppositional behavior in early and middle childhood by increasing positive parenting (Dishion et al., 2014; Shaw et al., 2017). Children's psychopathology was assessed with a broad index of symptoms and diagnoses of internalizing disorders with a structured clinical interview. Advances in nosology suggest that psychopathology can be parsed into broad dimensions rather than disorder-specific categories (Beauchaine & Cicchetti, 2016), and findings from twin studies have suggested that a shared latent genetic factor accounts for comorbidity and co-occurrence within and between anxiety disorders and depression (Middeldorp, Cath,

Van Dyck, & Boomsma, 2005). Thus, we initially tested genetic association using a sum of all symptoms across the internalizing diagnostic categories, rather than focusing on diagnoses or their symptom clusters as distinct analyses. Our first hypothesis was that children with higher polygenic sensitivity would have larger intervention effects. Children who scored higher on polygenic sensitivity and were in the intervention condition would have fewer internalizing symptoms than those who had higher polygenic sensitivity and were randomly assigned to the control condition. We also tested Polygenic Score × Intervention effects on "presence versus absence of any internalizing diagnosis" as an outcome, but we did not expect significant effects given the low rates of diagnoses at age 10 years. Because our sample was at high economic and family risk, our second hypothesis was that those randomly assigned to the control group who scored higher on genetic sensitivity to the environment would have more internalizing symptoms than those lower on polygenic sensitivity.

Method

Participants in the Early Steps Multisite Study

Seven hundred and thirty-one culturally diverse, low-income families with 2-year-old children were recruited between 2002 and 2003 from the Special Supplemental Nutrition Program for Women, Infants, and Children in Eugene, Oregon (suburban), within and outside Charlottesville, Virginia (rural), and in metropolitan Pittsburgh, Pennsylvania (urban). Screening procedures were used to recruit families of toddlers at high risk for conduct problems. Recruitment risk criteria were defined as 1 *SD* above the mean on screening measures in at least two of the following three domains: (a) sociodemographic risk (low education achievement—less than or equal to a mean of 2 years of post-high school education between parents and low family income using Special Supplemental Nutrition Program for Women, Infants, and Children criterion); (b) primary caregiver risk (maternal depression—Center for Epidemiological Studies on Depression Scale; Radloff, 1977; daily parenting challenges; Parenting Daily Hassles; Crnic & Greenberg, 1990; or self-report of substance or mental health diagnosis, or adolescent parent at birth of first child); and (c) toddler behavior problems (conduct problems—Eyberg Child Behavior Inventory; Robinson, Eyberg, & Ross, 1980; or high-conflict relationships with adults; Adult Child Relationship Scale; adapted from Pianta, Steinberg, & Rollins, 1995). Participation rates were high across the three sites (83.2% total [49% female]; 84% in Eugene [$n = 271$], 76% in Charlottesville [$n = 188$], and 88% in Pittsburgh [$n = 272$]). Primary caregivers (97% mothers) self-identified as belonging to the following ethnic groups: 13% Latino, 28% African American, 50% European American, 13% biracial, and 9% other groups (e.g., Native American or Asian American). More than two-thirds of the families reported an annual income of less than $20,000, with 24% of primary caregivers having less than a high school education,

41% having a high school diploma or general education diploma (GED), and an additional 32% having 1–2 years of post-high school education. For more information about sample characteristics, see Dishion et al. (2008).

Families were randomly assigned to the control condition or the intervention condition after the baseline assessment at child age 2 years. Those in the control condition received Special Supplemental Nutrition Program for Women, Infants, and Children services as usual. Those in the intervention condition received services implementing the Family Check-Up. The Family Check-Up is composed of three sessions: *assessment*, where research staff and parents completed questionnaires about the child's behavior and family factors, and parents and children were videotaped while taking part in tasks that varied in terms of stress level (e.g., free play vs. cleanup task); *initial interview*, where intervention staff and parents discussed their child's strengths and challenges as well as aspirations the parents had for their child; and *feedback*, where intervention staff provided feedback to the parents based on the assessment and initial interview, and encouraged reflection on behavioral change and engagement in further intervention services. All families were recontacted at child age 3, 4, 5, 7.5, 8.5, 9.5, and 10.5 years for home-based assessments, with intervention families also being offered the same Family Check-Up services through age 10.5. In terms of engagement, 76% of families engaged at age 2, with over 90% of the families engaging in at least one Family Check-Up by child age 5. Families also were seen at their homes at youth age 14, with 81% participating.

Adolescents who were genotyped at age 14 years ($n = 515$, or 86.7% of the sample who participated in home visits at age 14) make up the sample for the current study. These adolescents were 50% female and belonged to the following racial/ethnic groups: 10% Latino, 30% African American, 48% European American, 5% Native American, Native Hawaiian or Pacific Islander, 1% Asian American, and 6% other race or unknown race. Selective attrition analyses revealed no significant differences between members of the initial sample who did not give a saliva sample for genotyping at age 14, and those who did give a saliva sample with respect to parental education (high school diploma vs. no high school diploma), $\chi^2 (1) = 0.40$, $p = .53$; minority racial status (Black vs. non-Black), $\chi^2 (1) = 2.73$, $p = .10$; gender (male vs. female), $\chi^2 (1) = 0.45$, $p = .50$; intervention status (control vs. Family Check-Up), $\chi^2 (1) = 0.023$, $p = .88$; study site (Pennsylvania vs. non-Pennsylvania), $\chi^2 (1) = 2.27$, $p = .13$ (Virginia vs. non-Virginia), $\chi^2 (1) = 1.02$, $p = .31$; parental depression (assessed at child age 2 before the intervention), $t (590) = -0.003$, $p = .998$; child behavioral inhibition (assessed at child age 2), $t (562) = -0.99$, $p = .32$; and child conduct problems (assessed at child age 2), $t (591) = -1.36$, $p = .17$.

Procedure

The computerized self-report version of the National Institutes of Mental Health Diagnostic Interview Schedule for

Children—IV was administered to the child using a laptop computer at the age 10.5 years home visit (but not at other ages). Interviewers underwent several days of formal training with certified DISC-IV administrators. The interview took 30–45 min to complete with select modules, sometimes longer if many symptoms were endorsed.

During the age 14 home visit, participants provided saliva samples with Oragene kits for genotyping. RUCDR Infinite Biologics at Rutgers University extracted and normalized the DNA, and then genotyped the samples using the Affymetrix Axiom Biobank1 Array. Any SNP or individual with a missing data rate greater than or equal to 5% was removed (no participants met this criteria), and any SNP with a minor allele frequency less than 1% was removed. SNPs not in Hardy–Weinberg equilibrium at $p < 10^{-6}$ were also removed. To reduce correlation among the SNPs, or linkage disequilibrium, we did not impute the data; we screened out regions of long-range linkage disequilibrium, as well as local linkage disequilibrium, using the software PLINK's sliding window procedure.

Measures

Diagnostic Interview Schedule for Children (Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, *2000;* Shaffer, Fisher, Lucas, & NIMH DISC Editorial Board, *1998).* The National Institute of Mental Health Diagnostic Interview Schedule for Children—IV is a structured psychiatric interview for children age 6 years and older. Child responses are Yes or No for most questions, and follow-up questions are determined by previous answers in the module. The interrater reliability ($r = .93$) and test–retest reliability ($r = .64$) of the past-year diagnoses have been well established. Furthermore, the Diagnostic Interview Schedule for Children—IV showed moderate validity when compared to diagnoses generated from symptom ratings made after a clinical-style interview ($\kappa = .52$). The following seven internalizing modules were selected for use based on age-relevant disorders that corresponded to the research foci: generalized anxiety disorder, separation anxiety disorder, social anxiety disorder, specific phobia, obsessive–compulsive disorder, major depression, and manic disorder. A sum score representing total number of symptoms across modules was computed with an α reliability of 0.83 and was used in primary analyses. Follow-up analyses used disorder-specific symptom counts.

Polygenic sensitivity to the environment score. We formed polygenic scores based on the Twins Early Development Study discovery genome-wide association study indexing genetic influences on environmental sensitivity by predicting identical twin differences in emotional problems (Keers et al., 2016). After quality control, 318,549 SNPs remained in our data, with 53,010 present in both the discovery sample summary statistics and our data. We then filtered out synonymous SNPS resulting in 51,102 SNPs, and used PLINK's clumping procedure to account for nonindependence among the SNPs (threshold of $r^2 = .1$ and 250 kb), resulting in

36,246 independent SNPs. We formed polygenic sensitivity scores with p value thresholds of .001, .01, .05, and .10, unit-weighting each SNP. We chose not to form a score based on $p < .50$ because liberal scores often have high overlap with population admixture. The polygenic sensitivity scores contained 47 SNPs for $p = .001$; 503 SNPs for $p = .01$; 2,372 SNPs for $p = .05$; and 4,606 SNPs for $p = .10$.

Population admixture. We conducted a principal components analysis of all autosomal SNPs across all participants in the sample to represent population admixture using PLINK. We extracted the first 20 components, with the first component (PC1) having an eigenvalue of 28.84 and largely differentiating European American, Latino, and Native American groups from African American groups, with biracial participants falling in the middle. The second component (PC2) had an eigenvalue of 5.62 and largely differentiated non-Latino participants (European and African American) from Latino and Native American participants. The remaining components had eigenvalues ranging from 1.45 to 1.21 and were excluded from further analyses. Every participant had a score on each principal component and thus his or her genetic race/ethnicity was represented when controlling for population admixture.

Covariates. Covariates included in all models included age in months ($M = 128.560$, $SD = 3.362$, mean centered prior to analysis), gender (*females* = 0, *males* = 1; $M = 0.509$, $SD = 0.500$), family monthly income ($M = \$2,456.110$, $SD = \$1,594.626$, Z-scored), study site location (Eugene and Charlottsville compared to Pittsburgh indexed with two dummy codes), and the first two ancestry principal components, PC1 and PC2.

Statistical approach

We examined whether the interaction between intervention status and polygenic score was related to Diagnostic Interview Schedule for Children—IV symptoms and diagnoses using MPlus software version 7.4 (Muthén & Muthén, 1998–2015). Outcome variables were counts, but they were normally distributed and not zero inflated, and thus we used ordinary regression. Polygenic scores and the first two ancestry principle components (PC1 and PC2) were Z-scored prior to running models, and covariates were Z-scored or mean centered when appropriate. Initial models included the main effects of all covariates, two-way interactions between polygenic score and age, sex, and income (i.e., best practices for testing Gene × Environment interaction; Keller, 2014), the main effects of polygenic score and intervention, and a product term representing the interaction between polygenic score and intervention entered last. Main effects of all covariates were retained in final models regardless of significance, but two-way interactions between polygenic score and covariates were trimmed if they were nonsignificant.

We ran four regression models examining Polygenic × Intervention interaction in relation to total symptoms for each of

the four polygenic scores ($p = .001$, $p = .01$, $p = .05$, and $p = .10$). When interaction models were significant, we followed up by testing individual Diagnostic Interview Schedule for Children—IV symptom scales to see which carried the effect. Finally, we used logistic regression in Mplus with robust maximum likelihood estimation and Monte Carlo integration to examine whether significant interactions held when predicting diagnoses rather than symptoms ($0 = no\ diagnosis$, $1 = at\ least\ one\ diagnosis\ on\ any\ Diagnostic\ Interview\ Schedule\ for\ Children—IV\ scale$). Regions of significance and simple slopes for all significant interactions were calculated using Preacher, Curran, and Bauer's (2006) online utility for probing interaction effects.

Missing data. Full information maximum likelihood estimation was used to handle missing data. All 515 individuals in the genotyped sample had complete data for intervention status, polygenic score, PC1 and PC2, gender, and study site. Because there was complete data for both variables used in the product term, concerns regarding centering decisions when using full information maximum likelihood estimation with product terms do not apply.

Results

Descriptive statistics and correlations

Descriptive statistics for study variables are presented in Table 1. Skew did not exceed +/– 2.00 and kurtosis did not

exceed +/–7.00 (Curran, West, & Finch, 1996) for all variables except PC2, which was transformed by adding a constant equal to 1 plus the minimum value to ensure that all values were positive, then square root transforming prior to Z-scoring. Total symptoms ranged from 0 to 64 symptoms, out of a possible 88 symptoms. At the scale level, the maximum number of symptoms ranged from 6 for obsessive–compulsive disorder (out of a possible 7) to 21 for a major depressive episode (out of a possible 22). Diagnoses at the scale level were infrequent and are presented in Table 1 for descriptive purposes. Out of 418 children with Diagnostic Interview Schedule for Children—IV data, 100 had at least one diagnosis (23.9%). Of these 100 individuals, 74 had only one diagnosis (most commonly specific phobia, followed by obsessive–compulsive disorder), 21 had two diagnoses, 4 had three diagnoses, and 1 had four diagnoses.

Correlations among study variables were computed in Mplus using full information maximum likelihood estimation, with key correlations presented in Table 2, and others presented in text. Intervention status was uncorrelated with polygenic score, ancestry PCs, or other covariates, as expected based on random assignment, but also uncorrelated with total symptoms and all Diagnostic Interview Schedule for Children—IV symptom scales. Correlations between covariates and symptoms were modest and typically nonsignificant, although individuals with higher scores on PC1 (largely those with African American ancestry) showed significantly more symptoms of general anxiety ($r = .11$, $p < .05$), obsessive–compulsive disorder ($r = .19$, $p < .01$), and specific

Table 1. *Descriptive statistics and diagnoses*

	N	Min	Max	Mean	SD	Skewness	Kurtosis
Intervention	515	.000	1.000	0.500	0.500	−0.012	−2.008
Polygenic score ($p = .05$)	515	.479	0.519	0.498	0.006	−0.044	0.086
Polygenic score ($p = .10$)	515	.484	0.513	0.500	0.005	−0.119	0.016
Ancestry PC 1	515	−.039	0.081	0.000	0.044	0.611	−1.422
Ancestry PC 2	515	−.245	0.027	0.000	0.044	−3.184	10.664
Total symptoms	418	.000	64.000	19.845	12.824	0.765	0.146
General anxiety symptoms	418	.000	10.000	2.520	2.293	0.836	−0.021
Separation anxiety symptoms	418	.000	11.000	2.870	2.423	0.840	0.034
Social anxiety symptoms	417	.000	12.000	2.670	3.128	0.947	−0.345
Specific phobia symptoms	414	.000	9.000	1.880	1.815	1.166	1.175
Obsessive–compulsive symptoms	407	.000	6.000	1.210	1.295	1.116	0.737
Major depressive symptoms	414	.000	21.000	5.320	4.417	0.875	0.127
Manic symptoms	414	.000	12.000	3.520	2.433	0.781	0.237

				Diagnoses			
	General anxiety	Separation anxiety	Social anxiety	Specific phobia	Obsessive–compulsive	Major depression	Manic episode
No diagnosis	400	306	284	259	332	398	392
Diagnostic criteria not met, but symptoms and/or impairment present	13	94	123	95	47	7	20
Positive diagnosis	1	18	11	63	28	9	2

Note: Intervention is coded $0 = control$, $1 = intervention$. Polygenic scores are calculated as proportion of environmental sensitivity alleles below a particular p threshold. Ancestry PC 1 and 2 are principle components accounting for genetic variation due to race/ethnicity.

Table 2. *Zero-order correlations*

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Intervention	1.00												
2. Polygenic score ($p = .05$)	.01	1.00											
3. Polygenic score ($p = .10$)	.01	.73**	1.00										
4. Ancestry PC 1	.03	.09*	.02	1.00									
5. Ancestry PC 2	.00	.31**	.01	.00	1.00								
6. Total symptoms	.04	.04	.04	.08	.02	1.00							
7. General anxiety symptoms	.05	.06	.08	.11*	.01	.78**	1.00						
8. Sep. anxiety symptoms	.03	.06	.09	.07	.03	.77**	.59**	1.00					
9. Social anxiety symptoms	.00	.05	.03	.00	.09	.58**	.36**	.44**	1.00				
10. Specific phobia symptoms	.01	.07	.11*	.12*	.01	.55**	.40**	.41**	.21**	1.00			
11. Obsess. comp. symptoms	.00	.08	.11*	.19**	.01	.63**	.49**	.44**	.24**	.33**	1.00		
12. Major dep. symptoms	.07	.16	.01	.02	.08	.84**	.56**	.53**	.28**	.33**	.47**	1.00	
13. Manic symptoms	.02	.03	.03	.04	.05	.76**	.50**	.47**	.23**	.34**	.45**	.71**	1.00

Note: Intervention is coded $0 = control$, $1 = intervention$. Ancestry PC 1 and 2 are principle components accounting for genetic variation due to race/ethnicity. Sep. anxiety, separation anxiety. Obsess. comp., obsessive–compulsive. Major dep., major depression. *$p \leq .05$. **$p \leq .01$.

phobia ($r = .12$, $p < .05$). Males reported fewer social anxiety ($r = -.11$, $p < .05$) and specific phobia symptoms than females ($r = -.19$, $p < .01$), and older children reported fewer symptoms of separation anxiety ($r = -.10$, $p < .05$) and specific phobia ($r = -.17$, $p < .01$). As expected, all symptom scales were significantly and moderately to highly correlated with each other, although specific phobia symptoms were only modestly related to obsessive–compulsive, major depressive, and manic symptoms. Of note, polygenic scores across the different p thresholds were modestly to highly correlated with each other (with the exception of polygenic scores $p = .10$ and $p = .001$, which were not significantly correlated), and some polygenic scores were correlated with PC1 but not PC2 (see Table 2).

Testing Polygenic Sensitivity × Intervention Status interactions

Results for regression models testing the first hypothesis examining polygenic sensitivity, intervention status, and their interaction in relation to total symptoms are presented in Table 3. All two-way interactions between polygenic scores and covariates predicting total symptoms were initially included, but were nonsignificant. Specifically, Polygenic $p = .05 \times$ Gender, $p < .953$; Polygenic $p = .05 \times$ Income, $p < .216$; Polygenic $p = .05 \times$ Age, $p < .499$; Polygenic $p = .05 \times$ PC1, $p < .308$; Polygenic $p = .05 \times$ PC2, $p < .056$; and Polygenic $p = .05 \times$ Study Site, $p < .737$ and $p < .864$ for the two dummy codes. Similarly, Polygenic $p = .10 \times$ Gender, $p < .925$; Polygenic $p = .10 \times$ Income, $p < .114$; Polygenic $p = .10 \times$ Age, $p < .281$; Polygenic $p = .10 \times$ PC1, $p < .173$; Polygenic $p = .10 \times$ PC2, $p < .564$; and Polygenic $p = .10 \times$ Study Site, $p < .986$ and $p < .969$ for the two dummy codes. Interactions with covariates were removed from final models because none were significant, and excluding them from analyses did not alter the

significance of any model. Thus, we present results for models including only the main effect of covariates.

Total symptoms. The $p = .05$ polygenic score significantly interacted with intervention status, such that the intervention effect was stronger at higher levels of polygenic sensitivity, and a similar but weaker interaction was found for the $p = .10$ score (see Table 3). These results support the first hypothesis. There were no significant effects for either the $p = .001$ or the $p = .01$ polygenic scores, and they were not considered further. Regions of significance for the interaction effect across different values of polygenic sensitivity, and simple slopes of polygenic score in the control and intervention groups, are presented in Table 4 for significant interactions, and Figure 1 depicts regions of significance and simple slopes for the $p = .05$ and $p = .10$ scores. For the $p = .05$ score, there was a significant main effect of polygenic sensitivity, such that for individuals in the control group, higher polygenic score was significantly associated with higher total symptom count. This result supports the second hypothesis. However, the simple slope of polygenic score in the intervention group was nonsignificant. Testing regions of significance for the polygenic score indicated that the control and intervention groups differed significantly from each other on total symptoms for values of the polygenic score greater than 0.493 *SD* above the mean and less than 2.019 *SD* below the mean. As only 10 individuals scored lower than 2.00 *SD* below the mean (6 in the control group, 4 in the intervention), the lower bound of the region of significance should be interpreted with caution.

Unlike the $p = .05$ score, the simple slope of the $p = .10$ polygenic score was not significant in either the intervention or the control group when examining total symptoms. However, the full interaction model indicated that these two slopes did differ significantly from each other, and testing regions of significance indicated that the control and intervention groups differed significantly from each other on symptom count for

Table 3. *Testing polygenic moderation of intervention effects on total symptoms*

	Est.	*SE*	*p*	Lower CI	Upper CI	R^2
Polygenic score (p = .05)						
Gender	−2.268	1.253	.070	−4.723	0.187	.033
Age	−0.076	0.208	.716	−0.484	0.332	
Income	−0.040	0.676	.953	−1.364	1.285	
Site 1	−0.484	1.657	.770	−3.731	2.763	
Site 2	−0.763	1.671	.648	−4.037	2.512	
PC1	0.748	0.755	.322	−0.732	2.227	
PC2	0.184	0.634	.772	−1.059	1.426	
Intervention	−1.093	1.236	.377	−3.516	1.330	
Polygenic score	2.066*	0.891	.020	0.319	3.813	
Polygenic × Intervention	−3.261**	1.237	.008	−5.685	−0.836	
Polygenic score (p = .10)						
Gender	−2.165	1.257	.085	−4.629	0.298	.028
Age	−0.095	0.209	.648	−0.504	0.314	
Income	−0.098	0.678	.885	−1.427	1.230	
Site 1	−0.263	1.675	.875	−3.546	3.021	
Site 2	−0.674	1.678	.688	−3.963	2.615	
PC1	0.744	0.782	.341	−0.788	2.276	
PC2	0.188	0.636	.767	−1.059	1.436	
Intervention	−1.156	1.240	.351	−3.586	1.274	
Polygenic score	1.466	0.901	.104	−0.300	3.232	
Polygenic × Intervention	−2.670*	1.229	.030	−5.080	−0.260	

Note: PC1 and PC2, principle components 1 and 2, respectively.

values of polygenic sensitivity higher than 0.660 *SD* above the mean.

Symptom scales. For the two polygenic scores that interacted significantly with intervention status in relation to total symptoms, we followed up by testing the interaction separately for each symptom scale (see Table 5 for regression results, and Table 4 for regions of significance and simple slopes). Findings were similar across the two scores, with the interaction

between polygenic score and intervention status significantly associated with symptoms of general anxiety, obsessive–compulsive disorder, and mania. These results support the first hypothesis. In terms of regions of significance and simple slopes for the *p* = .05 score, significant findings at the scale level were largely similar to findings for total symptoms; however, for obsessive–compulsive symptoms, the control and intervention groups only differed significantly for individuals higher than 1.848 *SD* above the mean or lower

Table 4. *Regions of significance and simple slopes for significant interactions*

		Regions of significance for the polygenic score		Simple slopes for the Polygenic Score × Intervention Status	
		Lower bound	Upper bound	Control	Intervention
Polygenic score (p = .05)	Total symptoms	−2.019	0.493	2.066* (0.891)	−1.195 (0.865)
	General anxiety	−1.941	0.494	0.403** (0.159)	−0.186 (0.154)
	Obsessive–compulsive symptoms	−1.015	1.847	0.248** (0.090)	−0.052 (0.086)
	Manic symptoms	−1.031	0.934	0.391* (0.170)	−0.271 (0.165)
Polygenic score (p = .10)	Total symptoms	−5.380	0.660	1.466 (0.901)	−1.204 (0.894)
	General anxiety	−2.161	0.501	0.371* (0.160)	−0.197 (0.159)
	Obsessive–compulsive symptoms	−0.970	1.669	0.240** (0.090)	−0.067 (0.089)
	Manic symptoms	−4.088	2.497	0.112 (0.172)	−0.371* (0.171)

Note: Polygenic risk scores were *Z*-scored, so regions of significance are on a metric of standard deviations above and below the mean. The control and intervention groups are significantly different outside the upper and lower bounds of the region of significance for all models. *$p \leq .05$. **$p \leq .01$.
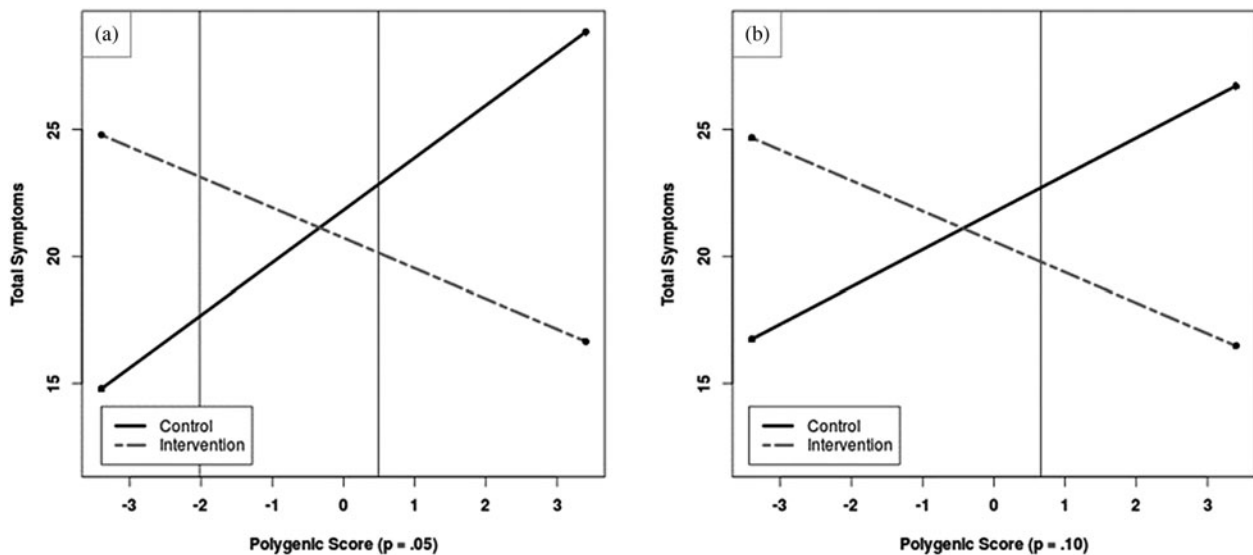
Figure 1. Polygenic Score × Intervention interaction. (a) Polygenic *p* = .05 score and (b) polygenic *p* = .10 score. Polygenic score is on a Z-score metric, and thus units on the *x*-axis correspond to standard deviations from the mean. The range of the polygenic scores observed in our sample is from –3.036 to 3.356 for the *p* = .05 score and from –3.291 to 2.763 for the *p* = .10 score. Vertical lines denote regions of significance. The control and intervention groups are significantly different from each other outside the vertical lines for the *p* = .05 score, and to the right of the vertical line for the *p* = .10 score.

than 1.015 *SD* below the mean on polygenic score. For the *p* = .10 score, the simple slope of polygenic score was significant in the control group for both general anxiety and obsessive–compulsive symptoms, such that a higher genetic score was associated with higher symptoms in the control group, with no significant association in the intervention group. This result supports the second hypothesis. For symptoms of mania, the simple slope of polygenic score was significant in the intervention but not the control group, such that polygenic score was associated with lower manic symptoms in the intervention group. However, regions of significance indicated that the control and intervention groups only differed significantly at values of polygenic score greater than 2.499 *SD* above the mean (three individuals in our sample, two in the intervention and one in the control), suggesting that results should be interpreted with caution.

Diagnoses. For the two polygenic scores that interacted significantly with intervention status to predict total symptoms, we used logistic regression in Mplus to examine whether polygenic sensitivity interacted with intervention status in relation to the presence versus absence of any diagnosis. The interaction between polygenic score and intervention status was not significantly related to diagnosis for either the *p* = .05 score (estimate = –.170, *SE* = .246, *p* = .489, odds ratio = 0.843) or the *p* = .10 score (estimate = .027, *SE* = .243, *p* = .912, odds ratio = 1.027).

Discussion

We argue that interdisciplinary approaches that capture the synergy between behavior genetics and cultural psychology

under the realm of cultural genomics are needed to test contemporary developmental theories and elucidate risk and resilience for child psychopathology. Each of these fields on its own captures important linear influences, but modeling nonlinear coaction across levels of analysis is needed to represent the many feedback loops that exist within and across biological and environmental levels. Empirically, we tested an interaction between a genetic index representing sensitivity to the environment and an efficacious intervention that focused on parenting. The major finding was that polygenic sensitivity moderated the effects of the Family Check-Up intervention on children's symptoms of internalizing psychopathology in a culturally diverse, high economic and family risk sample. Specifically, children who were genetically sensitive to the environment and were randomly assigned to the intervention group had fewer symptoms of internalizing psychopathology than genetically sensitive children assigned to the control condition.

This finding of significant moderation of intervention effects is very important because the intervention group did not linearly predict the child-report Diagnostic Interview Schedule for Children—IV symptoms of internalizing psychopathology in this sample. Thus, earlier in middle childhood (ages 7–8), investigators concluded that the Family Check-Up only demonstrated an indirect effect on the development of parent-reported internalizing symptoms by improving maternal depression during early childhood, with no direct effects (Reuben, Shaw, Brennan, Dishion, & Wilson, 2015). However, when genetic moderation was examined, we saw important effects for approximately 25% of the sample, with significant differences between the control and intervention groups beginning at 0.493 *SD* above the

Table 5. *Testing polygenic moderation of intervention effects on symptom scales*

Polygenic score ($p = .05$)							
		Est.	SE	p	Lower CI	Upper CI	R^2
General anxiety	Intervention	−.192	.221	.385	−0.626	0.242	.040
	Polygenic score	.403*	.159	.011	0.091	0.714	
	Polygenic × Intervention	−.589**	.221	.008	−1.022	−0.157	
Separation anxiety	Intervention	−.143	.234	.543	−0.601	0.316	.031
	Polygenic score	.273	.169	.106	−0.058	0.603	
	Polygenic × Intervention	−.301	.234	.199	−0.760	0.158	
Social anxiety	Intervention	−.035	.301	.908	−0.625	0.555	.038
	Polygenic score	.059	.217	.786	−0.366	0.484	
	Polygenic × Intervention	−.463	.301	.125	−1.053	0.128	
Specific phobia	Intervention	−.030	.171	.859	−0.365	0.304	.085
	Polygenic score	.258*	.123	.036	0.017	0.499	
	Polygenic × intervention	−.313	.171	.066	−0.648	0.021	
Obsessive–compulsive symptoms	Intervention	.044	.125	.727	−0.201	0.288	.059
	Polygenic score	.248**	.090	.006	0.073	0.424	
	Polygenic × Intervention	−.300*	.124	.015	−0.543	−0.058	
Major depressive symptoms	Intervention	−.468	.431	.278	−1.313	0.377	.018
	Polygenic score	.450	.309	.145	−0.156	1.056	
	Polygenic × Intervention	−.748	.430	.082	−1.590	0.094	
Manic symptoms	Intervention	−.013	.237	.957	−0.476	0.451	.027
	Polygenic score	.390*	.170	.021	0.058	0.723	
	Polygenic × Intervention	−.662**	.236	.005	−1.123	−0.200	

Polygenic score ($p = .10$)							
		Est.	SE	p	Lower CI	Upper CI	R^2
General anxiety	Intervention	−.201	.222	.364	−0.635	0.233	.038
	Polygenic score	.371*	.160	.021	0.057	0.685	
	Polygenic × Intervention	−.568**	.219	.009	−0.997	−0.139	
Separation anxiety	Intervention	−.144	.234	.537	−0.602	0.314	.034
	Polygenic score	.342*	.170	.044	0.009	0.675	
	Polygenic × Intervention	−.369	.232	.111	−0.823	0.085	
Social anxiety	Intervention	−.035	.302	.909	−0.626	0.557	.032
	Polygenic score	−.055	.219	.802	−0.485	0.375	
	Polygenic × Intervention	−.182	.299	.543	−0.769	0.404	
Specific phobia	Intervention	−.035	.171	.839	−0.369	0.300	.083
	Polygenic score	.238	.124	.055	−0.005	0.481	
	Polygenic × Intervention	−.272	.169	.108	−0.603	0.059	
Obsessive– compulsive symptoms	Intervention	.040	.125	.751	−0.205	0.284	.059
	Polygenic score	.240**	.090	.008	0.063	0.416	
	Polygenic × Intervention	−.307*	.123	.012	−0.548	−0.066	
Major depressive symptoms	Intervention	−.488	.432	.258	−1.334	0.358	.016
	Polygenic score	.232	.312	.457	−0.380	0.844	
	Polygenic × Intervention	−.600	.426	.160	−1.436	0.236	
Manic symptoms	Intervention	−.035	.237	.882	−0.500	0.430	.021
	Polygenic score	.112	.172	.515	−0.225	0.448	
	Polygenic × Intervention	−.483*	.234	.040	−0.942	−0.023	

Note: Child gender, age, monthly family income, and site were included as covariates in all models, but results are omitted to save space. No covariate significantly predicted symptom outcomes in any model. *$p \leq .05$. **$p \leq .01$.

mean on polygenic sensitivity with the $p = .05$ score. Results support the theoretically based assertion that parenting-based interventions such as the Family Check-Up have large effects on internalizing psychopathology for environmentally sensitive children, but no significant effects for children who are less sensitive to the environment. By averaging across all individuals in the sample, traditional approaches that do not account for individual differences do a disservice to the field and to the population by concluding that interventions have small and sometimes nonsignificant effect sizes on everyone. The children who do not respond to traditional preventive interventions focused on parenting are in need of interventions tailored to their characteristics and circumstances, and this work can proceed once they are identified. Overall, these

findings move us closer to using a tailored and personalized medicine approach when providing interventions for child psychopathology.

Despite theoretical support, only one other group has formed a polygenic sensitivity score based on genome-wide association study findings representing children's genetic predisposition to environmental sensitivity. Keers et al. (2016) conceptualized and empirically tested this polygenic sensitivity score utilizing two samples of children from the United Kingdom, the Twins Early Development Study sample of 12-year-old twin children, and the Genes for Treatment sample of children with diagnosed anxiety disorders. We replicated and expanded the generalization of their findings to a culturally diverse sample of children in the United States using a parenting-based intervention. However, our results focus exclusively on 10-year-olds, and genetic association may vary by chronological age or developmental period. With our follow-up scale-level findings, we reported stronger associations with anxiety disorders and obsessive–compulsive disorder symptoms than depression symptoms. Depression generally has a low base rate in middle childhood, and early life anxiety creates vulnerability for later depressive disorders and impairment across the life span (Emmelkamp & Wittchen, 2009). Thus, it may be that genetic association with depression would be stronger in adolescence. This hypothesis should be empirically examined, as should association with developmental outcomes, such as growth in child psychopathology across childhood and adolescence.

Besides Keers et al. (2016), only one other group has tested moderation of intervention effects using any polygenic score based on a discovery genome-wide association study, with the extant literature largely relying on candidate gene approaches (Bakermans-Kranenburg & van IJzendoorn, 2015). Musci and colleagues formed the tobacco polygenic risk score and found that it moderated the effects of an elementary school-based intervention on age of first cigarette smoked (Musci et al., 2015) and age of first marijuana use (Musci et al., 2018). Besides the importance of identifying individuals for whom the intervention worked, experimental studies hold great advantages for testing genetic moderation of physical environment and sociocultural effects. Most important, the polygenic score and the measure of the environment are uncorrelated in experimental designs, ensuring independence between genetic influences and changes in the environment. In addition, experimental studies typically use state-of-the-art concurrent and proximal measures of the environment that reduce measurement error. Partially because of these strengths, the power to test genetic moderation is much higher using experimental designs. In one set of simulations, a correlational study requiring 1,300 participants would only need 100 participants with an experimental design to have the same power to test genetic moderation (Bakermans-Kranenburg & van IJzendoorn, 2015). Thus, meaningful genetic effects can be obtained with traditional sample sizes that are powered to test moderation.

Theoretical implications of study findings

Our results are consistent with differential susceptibility theory (Belsky & Pluess, 2009; Ellis et al. 2011) in that children in the intervention group had significantly fewer internalizing symptoms than those in the control group if they were higher on polygenic sensitivity to the environment (Hypothesis 1). In addition, children in the control group with higher polygenic sensitivity had more internalizing symptoms than those also in the control group with lower polygenic sensitivity (Hypothesis 2). These findings support the theoretical assertion that differentially susceptible children do worse in high-risk environments (i.e., control group experiencing economic and family risk) and do better in protective or promotive environments (i.e., intervention group experiencing the same economic risk, but with improved parent mental health and parents utilizing more positive behavioral management after the intervention). Note that we could not compare the diathesis-stress (i.e., an interaction between a dispositional diathesis and environmental stress when predicting negative outcomes), differential susceptibility, and vantage sensitivity theories (i.e., an interaction between a dispositional diathesis and environmental support when predicting positive outcomes) because these models cannot be differentiated unless child outcomes are assessed across the full continuum from negative to positive (Clifford & Lemery-Chalfant, 2015). Our outcome was symptoms of child psychopathology, which does not capture positive adaptation or flourishing.

The overall finding of genetic differences in response to intervention may explain why past attempts to identify genetic variants linearly associated with child psychopathology have been largely unsuccessful. It also explains other patterns observed in the behavior genetics literature, such as the discrepancy between "too low" SNP heritability (i.e., the degree to which SNP variation accounts for phenotypic variation) compared to heritability estimates obtained from quantitative genetic approaches such as twin and adoption studies.

Polygenic sensitivity to the environment based on identical twin differences

It is standard practice when creating polygenic scores from genome-wide association study results to form the scores at multiple p value cutoffs and test association with all of these scores, as individual SNP effect sizes are unknown and are thought to be small. In line with our findings, Keers et al. (2016) also reported that the $p < .05$ and $p < .10$ polygenic sensitivity scores, but not the $p < .01$ or $p < .001$ scores, significantly moderated the cognitive behavioral therapy intervention on children's anxiety. These are not independent findings as these two scores have a high degree of overlap and are correlated, $r = .73$, $p < .01$, in our sample. The fact that two relatively conservative scores captured the important variance is encouraging as association with more liberal scores (e.g., $p = .50$) could substantially increase the chance of spurious results. It is likely that the use of identical twins in the discovery

genome-wide association study led to scores that were relatively free of environmental contamination, as the identical twin difference design controls for all genetic (i.e., identical twins have identical genotypes) and family-level influences (e.g., family socioeconomic and sociocultural status) on the phenotypes (Lemery & Goldsmith, 1999).

At the same time, this score was limited to identical twin differences in composited parent report and self-report of emotional problems in 12-year-olds growing up in the United Kingdom. It is likely that twin differences in emotional problems do not capture all of the genetic influences on environmental sensitivity. Thus, future work should examine genetic association with identical twins of different ages and with different phenotypes including externalizing psychopathology to help inform polygenic sensitivity to the environment.

Study limitations

In addition to many strengths, including the theoretically-based design that included random assignment to an intervention, there are several limitations. First, our results do not generalize to everyone. We examined internalizing psychopathology only at age 10 years, the only age when the Diagnostic Interview Schedule for Children—IV interview was conducted with children. Genetic and environmental influences can vary in their relevance to development at different ages, and thus our findings might be specific to middle childhood (Dick et al., 2017). Although our sample had adequate representation of European Americans, African Americans, and Latinos, we had little representation of other groups such as Native Americans and Asian Americans. All families were recruited from Women, Infant, and Children nutrition supplement programs for low-income families. Families were additionally screened for family (e.g., maternal depression) and/or child risk (e.g., problem behaviors). Based on these restrictions, the results are most directly applicable to families with economic and family risk for child psychopathology. Second, the polygenic sensitivity score was formed based on a discovery genome-wide association study conducted on genotyped identical twins participating on the longitudinal Twins Early Development Study. The sample of twins was 93% White, which is representative of the population of the United Kingdom (Haworth et al., 2013). However, it is unknown the extent to which SNPs identified in this largely White sample may generalize to more diverse samples such as ours. Third, we relied on children's self-reports during a structured clinical interview to assess outcomes. Although self-report can be potentially biased, we think that self-reported outcomes are more conservative than caregiver-reported outcomes when testing the effects of the Family Check-Up, as the primary focus of this intervention is on changing parenting (Van Ryzin et al., 2012). A fourth limitation is that our analysis conflated assignment to the intervention group with receipt of treatment. Each year, 25% to 35% of those assigned to the intervention did not engage in services (Dishion et al., 2014). Thus, true effect sizes are likely to be larger than reported here.

Implications and future directions

Forming polygenic scores based on biological function. An important implication of the theoretical foundation and empirical findings from this study is that important SNPs should not be limited to those that show linear associations with phenotypes in a genome-wide association study. Through consortiums and data harmonization, sample sizes for genome-wide association studies have increased from thousands of participants to tens of thousands and sometimes hundreds of thousands of participants, which ensures that we can identify SNPs with significant main effects, even if they have very small effect sizes. Nearly all of these participants are of European descent in an attempt to control for population stratification of genotypes. With these large samples, measurement of the phenotype is very limited, sometimes involving only a few self- or parent-reported items, or a single dichotomous diagnosis. Furthermore, the genome-wide association study model is elementary, including separate linear main effects of each SNP across the genome, with few covariates included. One could imagine that the relation between a single genetic variant and a complex psychological outcome such as a psychiatric diagnosis might be more complex and involve additional predictors or levels of analysis. Moreover, there is some limited evidence that when the genetically informed model goes beyond main effects and includes gene–environment correlation and interaction, some combinations yield moderate and large effects (Bakermans-Kranenburg & van IJzendoorn, 2015). The problem with ignoring the environmental and cultural levels of analysis in behavior genetics studies is that genetic association likely varies by components of the environment and culture, such that the same genotype can lead to different phenotypes in different physical and sociocultural contexts.

Bioinformatics provides methods of incorporating additional knowledge into the selection of genetic polymorphisms to assist in forming polygenic scores for use in genetic association analyses. Bioinformatic tools are available to help identify gene networks and biological pathways with known function in brain regions linked to phenotypes such as developmental psychopathology. One example is the use of gene set enrichment analysis to identify biologically meaningful gene sets from a list of initial SNPs (Holden, Deng, Wojnowski, & Kulle, 2008). The initial list of SNPs can come from a targeted microarray, from proteomics, or from a genome-wide association study with SNPs showing association with the phenotype, for example. Gene set enrichment analysis helps identify phenotype-relevant SNPs, and also identifies mechanisms of action by identifying corresponding biological pathways. Future studies in cultural genomics should use bioinformatics tools to select functionally meaningful genetic variants for further study.

Genetically informative studies of culturally diverse groups. Biased sampling and the neglect of cultural contributions to health and development is threatening the generalizability of what we know about human behavior and social processes gen-

erally, and developmental psychopathology more specifically. Over 80% of research participants are Western, educated, industrialized, rich, democratic (WEIRD) people who make up only 12% of the world's population (Henrich, Heine, & Norenzayan, 2010). By turning scientific attention toward variable in traditions and lifeways (VITAL) peoples, we can begin to understand cultural and biological interplay and paint a more accurate picture of risk and resilience processes in development.

The study of genetic influences on social, behavioral and health processes is similarly afflicted with a focus on WEIRD populations, threatening what we know about genetic contributions to development. In this field, findings from studies of these populations are considered human universals, while extensive cultural variation in these processes are systematically understudied. Furthermore, the field of behavior genetics has laws based largely on findings from studies of WEIRD populations. Turkheimer (2000) presented the three laws of behavior genetics, with Chabris, Lee, Cesarini, Benjamin, and Laibson (2015) adding a fourth law. The first law is that all human behavioral traits are affected by genetic variation. The fourth law of behavior genetics added by Chabris et al. (2015) expands this first law to state that human traits are influenced by many genetic variants of small effect ($<1\%$) rather than a few genes of large effect. The second law is that the effect of being raised in the same family is smaller than the effect of genes, with twin studies of WEIRD samples yielding genetic effects around 50% with negligible effects of the shared environment, or aspects of the environment that create similarities between individuals (Polderman et al., 2015).

With truncated trait relevant environmental differences present when studying WEIRD samples, these "laws" are not surprising. For example, when giving the presidential address at the biannual meetings of the Society for Research in Child Development, behavior geneticist Dr. Sandra Scarr made the shocking argument that parenting within the normal range does not matter for child development (Scarr, 1992). With research on WEIRD children, rearing environments are likely less variable and thus genetics accounts for a larger proportion of individual differences than does the environment and culture. However, when we expand the reach of behavior genetics studies to other cultural groups or populations, or we improve our models to better allow for individual differences, we see a different pattern.

Relevant to child psychopathology, a meta-analysis reported that additive genetic influences accounted for nearly half of the variance in self-reported Child Behavior Checklist symptoms of anxiety (48%) and depression (44%), with smaller shared environmental influences (12% and 14%, respectively) with samples made up of WEIRD children (Burt, 2009). Longitudinal studies also suggest that stability is largely genetically mediated, with some new genetic influence emerging in adolescence. These findings are in line with the laws of behavior genetics. In contrast, a recent study of Chinese twins who self-reported anxiety and depression on the Child Behavior Checklist longitudinally at approximately 11 and 14 years of age yielded unexpected findings (Zheng, Rijsdijk, Pingault, McMahon, &

Unger, 2016). Heritability estimates were substantially lower (23% and 20%, respectively) at 11 years and decreased to negligible by 14 years. Shared environmental influences were similar at age 11 (20% and 27%, respectively) and increased substantially (57% and 60%, respectively) by 14 years. Stability was largely explained by the shared environment. In China, the transition to adolescence corresponds to environmental changes such that family and neighborhood experiences could have a larger impact on developmental psychopathology. It also could be that active gene–environment correlation occurs more frequently with WEIRD populations as children have more opportunity to select environments that match their genetic dispositions. Active gene–environment correlation would decrease fraternal twin correlations relative to identical twin correlations and thus increase estimates of genetic influence, and decrease estimates of shared environmental influence. Other studies of WEIRD samples utilizing moderated heritability models also elucidate the hidden importance of the sociocultural environment for some traits and some individuals (children's temperament; Lemery-Chalfant, Kao, Swann, & Goldsmith, 2013).

Finally, the third law of behavior genetics states that a large portion of individual differences are not due to genetics nor the family, that is, they are *nonshared environmental influences* that create differences rather than similarities between individuals (Turkheimer, 2000). Behavior geneticists have labeled the attempt to identify what aspects of the trait-relevant environment (e.g., differential treatment by parents and different classrooms) make up the nonshared environment "the Gloomy Prospect," as most studies yielded null results (Plomin & Daniels, 1987). However, differential susceptibility theory indicates that one way genetics may influence health and development is by representing variability in sensitivity to the environment (Belsky & Pluess, 2009). Keers et al.'s (2016) polygenic sensitivity score based on SNPs that predicted identical twin differences in emotional problems suggests that genetic variants can be meaningfully related to identical twin differences. Thus, the difficulty in accounting for estimates of nonshared environmental influence found in twin studies are likely explained in part by environmental influences acting in concert with individual genetic predispositions over time, rather than the simple linear effect of any single environmental factor. With differential susceptibility theory to guide the identification of genetic variants associated with environmental sensitivity, there may be nothing gloomy about "the Gloomy Prospect" at all.

Thus, the laws for WEIRD peoples should be updated to incorporate developmental processes in VITAL peoples to elucidate risk and resilience for developmental psychopathology. The field of behavior genetics is young and stands to gain much by coupling with the field of cultural psychology in the new field of cultural genomics to better represent the full range of genetic and environmental trait-relevant variation in VITAL peoples. Meanwhile, cultural psychology can benefit from integrating genetic and physiological theory and analysis into cultural research, especially as an understanding of genetics research and methodology can provide some of the most

compelling arguments against an overly simplistic, deterministic view of the role of genetics in shaping individual outcomes. In line with differential susceptibility theory, the extent to which risk and promotive factors contribute to phenotypes may be a function of a genetic predisposition to environmental sensitivity, or developmental plasticity. Few studies have tested this theory in minority groups.

There are racial–ethnic disparities in developmental outcomes for youth and families, with some initial evidence that genetics moderates the links, for example, between racial discrimination and the development of conduct problems (Brody et al., 2011) and criminal arrests (Schwartz & Beaver, 2011). We have an ethical obligation to pursue the public health priority of designing and testing multicultural interventions (Smith et al., 2014), and examining putative moderators of intervention efficacy.

Conclusion

We argue for an integration of cultural psychology and behavior genetics under contemporary developmental theories to bridge these perspectives and make developmental science more representative of local cultural conditions and more valid across cultures of the world. Based on our findings, it is likely that intervention effects on environmentally sensitive children are underestimated. We found that the parenting-based Family Check-Up intervention was effective in children with higher polygenic sensitivity to the environment, such that a significant difference in self-reported Diagnostic Interview Schedule for Children—IV symptoms emerged between the intervention and control groups for those 0.493 *SD* above the mean on polygenic sensitivity to the environment, or 25% of the sample. It is useful to understand which children benefit from parenting-based interventions and which children do not. Children with lower environmental sensitivity may benefit from an alternative intervention, perhaps a cognitive-based intervention. Similar to personalized medicine, it is time to understand individual differences in treatment response and individualize psychosocial interventions to reduce the burden of child psychopathology and maximize well-being for children growing up in a wide range of physical environments and cultures.

References

Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2011). Differential susceptibility to rearing environment depending on dopamine-related genes: New evidence and a meta-analysis. *Developmental Psychopathology*, 23, 39–52. doi:10.1017/ S0954579410000635

Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2015). The hidden efficacy of interventions: Gene-environment experiments from a differential susceptibility perspective. *Annual Review of Psychology*, 66, 381–409. doi:10.1146/annurev-psych-010814-015407

Beauchaine, T. P., & Cicchetti, D. (2016). A new generation of comorbidity research in the era of neuroscience and research domain criteria. *Development and Psychopathology*, 28, 891–894. doi:10.1017/ S0954579416000602

Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin*, 135, 885–908. doi:10.1037/a0017376

Brody, G. H., Beach, S. R., Chen, Y. F., Obasi, E., Philibert, R. A., Kogan, S. M., & Simons, R. L. (2011). Perceived discrimination, serotonin transporter linked polymorphic region status, and the development of conduct problems. *Development and Psychopathology*, 23, 617–627. doi:10.1017/S0954579411000046

Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review*, 101, 568–586. doi:10.1037/0033-295X.101.4.568

Burt, S. A. (2009). Rethinking environmental contributions to child and adolescent psychopathology: A meta-analysis of shared environmental influences. *Psychological Bulletin*, 135, 608–637. doi:10.1037/a0015702

Causadias, J. M. (2013). A roadmap for the integration of culture into developmental psychopathology. *Development and Psychopathology*, 25, 1375–1398. doi:10.1017/S0954579413000679

Causadias, J. M., Telzer, E. H., & Gonzales, N. A. (2018). Introduction to culture and biology interplay. In J. M. Causadias, E. H. Telzer, & N. A. Gonzales (Eds.), *The handbook of culture and biology* (pp. 3–29). New York: Wiley.

Chabris, C. F., Lee, J. J., Cesarini, D., Benjamin, D. J., & Laibson, D. I. (2015). The fourth law of behavior genetics. *Current Directions in Psychological Science*, 24, 304–312. doi:10.1177/0963721415580430

Clifford, S., & Lemery-Chalfant, K. (2015). Molecular genetics of resilience. In M. Pluess (Ed.), *Genetics of psychological well-being: The role of heritability and genetics in positive psychology* (pp. 177–192). Oxford: Oxford University Press.

Cohen, A. B. (2009). Many forms of culture. *American Psychologist*, 64, 194–204. doi:10.1037/a0015308

Connell, A. M., & Dishion, T. J. (2017). Long-term effects of the Family Check-Up in public secondary school on diagnosed major depressive disorder in adulthood. *Journal of Youth and Adolescence*, 46, 570–581. doi:10.1007/s10964-016-0482-6

Crnic, K. A., & Greenberg, M. T. (1990). Minor parenting stresses with young children. *Child Development*, 61, 1628–1637. doi:10.1111/ j.1467-8624.1990.tb02889.x

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1, 16–29. doi:10.1037/1082-989X.1.1.16

Dick, D. M., Barr, P. B., Cho, S. B., Cooke, M. E., Kuo, S. I., Lewis, T. J., . . . Su, J. (2017). Post-GWAS in psychiatric genetics: A developmental perspective on the "other" next steps. *Genes, Brain and Behavior*. Advance online publication. doi:10.1111/gbb.12447

Dishion, T. J., Brennan, L. M., Shaw, D. S., McEachern, A. D., Wilson, M. N., & Jo, B. (2014). Prevention of problem behavior through annual family check-ups in early childhood: Intervention effects from home to early elementary school. *Journal of Abnormal Child Psychology*, 42, 343–354. doi:10.1007/s10802-013-9768-2

Dishion, T. J., Kavanagh, K., Schneiger, A., Nelson, S., & Kaufman, N. K. (2002). Preventing early adolescent substance use: A family-centered strategy for the public middle school. *Prevention Science*, 3, 191–201. doi:10.1023/A:1019994500301

Dishion, T. J., Shaw, D., Connell, A., Gardner, F., Weaver, C., & Wilson, M. (2008). The Family Check-Up with high-risk indigent families: Preventing problem behavior by increasing parents' positive behavior support in early childhood. *Child Development*, 79, 1395–1414. doi:10.1111/ j.1467-8624.2008.01195.x

Dishion, T. J., & Stormshak, E. (2007). *Intervening in children's lives: An ecological, family-centered approach to mental health care*. Washington, DC: American Psychological Association.

Dodge, K. A., Pettit, G. S., & Bates, J. E. (1994). Socialization mediators of the relation between socioeconomic status and child conduct problems. *Child Development*, 65, 649–665. doi:10.2307/1131407

Ellis, B. J., Boyce, W. T., Belsky, J., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2011). Differential susceptibility to the environment: An evolutionary–neurodevelopmental theory. *Development and Psychopathology*, 23, 7–28. doi:10.2307/1131407

Emmelkamp, P. M., & Wittchen, H. U. (2009). Specific phobias. In G. Andrews, D. S. Charney, P. J. Sirovatka, & D. A. Regier (Eds.), *Stress-induced and fear circuitry disorders. Refining the research Agenda for DSM-V* (pp. 77–101). Arlington, VA: American Psychological Association.

Harold, G. T., Rice, F., Hay, D. F., Boivin, J., van Den Bree, M., & Thapar, A. (2011). Familial transmission of depression and antisocial behavior symptoms: Disentangling the contribution of inherited and environmental factors and testing the mediating role of parenting. *Psychological Medicine*, *41*, 1175–1185. doi:10.1017/S0033291710001753

Haworth, C. M., Davis, O. S., & Plomin, R. (2013). Twins Early Development Study (TEDS): A genetically sensitive investigation of cognitive and behavioral development from childhood to young adulthood. *Twin Research and Human Genetics*, *16*, 117–125. doi:10.1017/thg.2012.91

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61–83. doi:10.1017/S0140525X0999152X

Holden, M., Deng, S., Wojnowski, L., & Kulle, B. (2008). GSEA-SNP: Applying gene set enrichment analysis to SNP data from genome-wide association studies. *Bioinformatics*, *24*, 2784–2785. doi:10.1016/j.ygeno.2011.04.006

International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. *Nature*, *431*, 931. doi:10.1038/nature03001

Ioannidis, J. P. (2005). Why most published research findings are false. *PLOS Medicine*, *2*, e124. doi:10.1371/journal.pmed.0020124

Keers, R., Coleman, J. R., Lester, K. J., Roberts, S., Breen, G., Thastum, M., . . . Nauta, M. (2016). A genome-wide test of the differential susceptibility hypothesis reveals a genetic predictor of differential response to psychological treatments for child anxiety disorders. *Psychotherapy and Psychosomatics*, *85*, 146–158. doi:10.1159/000444023

Keller, M. C. (2014). Gene × Environment interaction studies have not properly controlled for potential confounders: The problem and the (simple) solution. *Biological Psychiatry*, *75*, 18–24. doi:10.1016/j.biopsych.2013.09.006

Kendler, K. S., Thornton, L. M., & Pedersen, N. L. (2000). Tobacco consumption in Swedish twins reared apart and reared together. *Archives of General Psychiatry*, *57*, 886–892. doi:10.1001/archpsyc.57.9.886

Lemery, K. S., & Goldsmith, H. H. (1999). Genetically informative designs for the study of behavioural development. *International Journal of Behavioral Development*, *23*, 293–317. doi:10.1080/016502599383838

Lemery-Chalfant, K., Kao, K., Swann, G., & Goldsmith, H. H. (2013). Childhood temperament: Passive gene–environment correlation, gene–environment interaction, and the hidden importance of the family environment. *Development and Psychopathology*, *25*, 51–63. doi:10.1017/S0954579412000892

Lunkenheimer, E. S., Dishion, T. J., Shaw, D. S., Connell, A. M., Gardner, F., Wilson, M. N., & Skuban, E. M. (2008). Collateral benefits of the family check-up on early childhood school readiness: Indirect effects of parents' positive behavior support. *Developmental Psychology*, *44*, 1737–1752. doi:10.1037/a0013858

McDonough-Caplan, H., Klein, D. N., & Beauchaine, T. P. (2018). Comorbidity and continuity of depression and conduct problems from elementary school to adolescence. *Journal of Abnormal Psychology*, *127*, 326–337. doi:10.1037/abn0000339

McGoldrick, M., & Hardy, K. V. (Eds.) (2008). *Re-visioning family therapy: Race, culture, and gender in clinical practice*. New York: Guilford Press.

Middeldorp, C. M., Cath, D. C., Van Dyck, R., & Boomsma, D. I. (2005). The co-morbidity of anxiety and depression in the perspective of genetic epidemiology. A review of twin and family studies. *Psychological Medicine*, *35*, 611–624. doi:10.1017/S003329170400412X

Moilanen, K. L., Shaw, D. S., & Maxwell, K. L. (2010). Developmental cascades: Externalizing, internalizing, and academic competence from middle childhood to early adolescence. *Development and Psychopathology*, *22*, 635–653. doi:10.1017/S0954579410000337.

Musci, R. J., Fairman, B., Masyn, K. E., Uhl, G., Maher, B., Sisto, D. Y., . . . Ialongo, N. S. (2018). Polygenic Score × Intervention Moderation: An application of discrete-time survival analysis to model the timing of first marijuana use among urban youth. *Prevention Science*, *19*, 6–14. doi:10.1007/s11121-016-0729-1

Musci, R. J., Masyn, K. E., Uhl, G., Maher, B., Kellam, S. G., & Ialongo, N. S. (2015). Polygenic Score × Intervention Moderation: An application of discrete-time survival analysis to modeling the timing of first tobacco use among urban youth. *Development and Psychopathology*, *27*, 111–122. doi:10.1017/S0954579414001333

Muthén, L. K., & Muthén, B. O. (1998–2015). *Mplus user's guide: Statistical analysis with latent variables* (Version 7.4). Los Angeles: Author.

Oliver, B. R., Trzaskowski, M., & Plomin, R. (2014). Genetics of parenting: The power of the dark side. *Developmental Psychology*, *50*, 1233–1240. doi:10.1037/a0035388

Oquendo, M. A., Canino, G., Lehner, T., & Licinio, J. (2010). Genetic repositories for the study of major psychiatric conditions: What do we know about ethnic minorities' genetic vulnerability? *Molecular Psychiatry*, *15*, 970–975. doi:10.1038/mp.2010.11

Patterson, G. R., Reid, J. B., & Dishion, T. J. (1992). *Antisocial boys* (Vol. 4). Eugene, OR: Castalia.

Pianta, R. C., Steinberg, M. S., & Rollins, K. B. (1995). The first two years of school: Teacher child relationships and deflections in children's classroom adjustment. *Development and Psychopathology*, *7*, 295–312. doi:10.1017/S0954579400006519

Pinquart, M. (2017). Associations of parenting dimensions and styles with externalizing problems of children and adolescents: An updated meta-analysis. *Developmental Psychology*, *53*, 873–932. doi:10.1037/dev0000295

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences*, *10*, 1–60. doi:10.1093/ije/dyq148

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin*, *84*, 309–322. doi:10.1037/0033-2909.84.2.309

Polderman, T. J., Benyamin, B., De Leeuw, C. A., Sullivan, P. F., Van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nature Genetics*, *47*, 702–709. doi:10.1038/ng.3285

Popejoy, A. B., & Fullerton, S. M. (2016). Genomics is failing on diversity. *Nature News*, *538*, 161–164. doi:10.1038/538161a

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*, 437–448. doi:10.3102/10769986031004437

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401. doi:10.1177/014662167700100306

Rende, R., & Plomin, R. (1992). Diathesis–stress models of psychopathology: A quantitative genetic perspective. *Applied and Preventative Psychology*, *1*, 177–182. doi:10.1016/S0962-1849(05)80123-4

Reuben, J. D., Shaw, D. S., Brennan, L. M., Dishion, T. J., & Wilson, M. N. (2015). A family-based intervention for improving children's emotional problems through effects on maternal depressive symptoms. *Journal of Consulting and Clinical Psychology*, *83*, 1142–1148. doi:10.1037/ccp0000049

Roalf, D. R., & Gur, R. C. (2017). Functional brain imaging in neuropsychology over the past 25 years. *Neuropsychology*, *31*, 954–971. doi:10.1037/neu0000426

Robinson, E. A., Eyberg, S. M., & Ross, A. W. (1980). The standardization of an inventory of child conduct problem behaviors. *Journal of Clinical Child and Adolescent Psychology*, *9*, 22–28. doi:10.1080/15374418009532938

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development*, *63*, 1–19. doi:10.1111/j.1467-8624.1992.tb03591.x

Schwartz, J. A., & Beaver, K. M. (2011). Evidence of a gene × environment interaction between perceived prejudice and MAOA genotype in the prediction of criminal arrests. *Journal of Criminal Justice*, *39*, 378–384. doi:10.1016/j.jcrimjus.2011.05.003

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children, Version IV (NIMH DISC-IV): Description, differences from previous versions and reliability of some common diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*, 28Y38. doi:10.1097/00004583-200001000-00014

Shaffer, D., Fisher, P. I., Lucas, C., & NIMH DISC Editorial Board. (1998). *Diagnostic Interview Schedule for Children, Version IV*. New York: Columbia University, Division of Psychiatry.

Shanahan, M. J., & Hofer, S. M. (2005). Social context in gene -environment interactions: Retrospect and prospect. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *60*, 65–76. doi:10.1093/geronb/60.Special_Issue_1.65

Shaw, D. S., Connell, A., Dishion, T. J., Wilson, M. N., & Gardner, F. (2009). Improvements in maternal depression as a mediator of intervention ef-

fects on early childhood problem behavior. *Development and Psychopathology*, *21*, 417–439. doi:10.1017/S0954579409000236

Shaw, D. S., Dishion, T. J., Supplee, L., Gardner, F., & Arnds, K. (2006). Randomized trial of a family-centered approach to the prevention of early conduct problems: 2-year effects of the Family Check-Up in early childhood. *Journal of Consulting and Clinical Psychology*, *74*, 1–9. doi:10.1037/0022-006X.74.1.1

Shaw, D. S., Sitnick, S. L., Brennan, L. M., Choe, D. E., Dishion, T. J., Wilson, M. N., & Gardner, F. (2017). The long-term effectiveness of the Family Check-Up on school-age conduct problems: Moderation by neighborhood deprivation. *Development and Psychopathology*, *28*, 1471–1486. doi:10.1017/S0954579415001212

Smith, J. D., Berkel, C., Hails, K. A., Dishion, T. J., Shaw, D. S., & Wilson, M. N. (2017). Predictors of participation in the Family Check-Up program: A randomized trial of yearly services from age 2 to 10 years. *Prevention Science*, *88*, 427–435. doi:10.1007/s11121-016-0679-7

Smith, J. D., Knoble, N. B., Zerr, A. A., Dishion, T. J., & Stormshak, E. A. (2014). Family Check-Up effects across diverse ethnic groups: Reducing early-adolescence antisocial behavior by reducing family conflict. *Journal of Clinical Child and Adolescent Psychology*, *43*, 400–414. doi:10.1080/15374416.2014.888670

Stormshak, E. A., Connell, A. M., Véronneau, M. H., Myers, M. W., Dishion, T. J., Kavanagh, K., & Caruthers, A. S. (2011). An ecological approach to promoting early adolescent mental health and social adaptation: Family-centered intervention in public middle schools. *Child Development*, *82*, 209–225. doi:10.1111/j.1467-8624.2010.01551.x

Stormshak, E. A., Fosco, G. M., & Dishion, T. J. (2010). Implementing interventions with families in schools to increase youth school engagement: The Family Check-Up model. *School Mental Health*, *2*, 82–92. doi:10.1007/s12310-009-9025-6

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science*, *9*, 160–164. doi:10.1111/1467-8721.00084

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. doi:10.1111/1467-8721.00084

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

Van Ryzin, M. J., Stormshak, E. A., & Dishion, T. J. (2012). Engaging parents in the Family Check-Up in middle schools: Longitudinal effects through the transition to high school. *Journal of Adolescent Health*, *50*, 627–633. doi:10.1016/j.jadohealth.2011.10.255

Zheng, Y., Rijsdijk, F., Pingault, J. B., McMahon, R. J., & Unger, J. B. (2016). Developmental changes in genetic and environmental influences on Chinese child and adolescent anxiety and depression. *Psychological Medicine*, *46*, 1829–1838. doi:10.1017/S0033291716000313

Zvara, B. J., Sheppard, K. W., & Cox, M. (2018). Bidirectional effects between parenting sensitivity and child behavior: A cross-lagged analysis across middle childhood and adolescence. *Journal of Family Psychology*, *32*, 484–495. doi:10.1037/fam0000372