Three phases of Gene × Environment interaction research: Theoretical assumptions underlying gene selection

Xiaoya Zhang and Jay Belsky
Department of Human Ecology, University of California, Davis, CA, USA

Abstract
Some Gene × Environment interaction (G×E) research has focused upon single candidate genes, whereas other related work has targeted multiple genes (e.g., polygenic scores). Each approach has informed efforts to identify individuals who are either especially vulnerable to the negative effects of contextual adversity (diathesis stress) or especially susceptible to both positive and negative contextual conditions (differential susceptibility). A critical step in all such molecular G×E research is the selection of genetic variants thought to moderate environmental influences, a subject that has not received a great deal of attention in critiques of G×E research (beyond the observation of small effects of individual genes). Here we conceptually distinguish three phases of G×E work based on the selection of genes presumed to moderate environmental effects and the theoretical basis of such decisions: (a) single candidate genes, (b) composited (multiple) candidate genes, and (c) GWAS-derived polygenic scores. This illustrative, not exhaustive, review makes it clear that implicit or explicit theoretical assumptions inform gene selection in ways that have not been clearly articulated or fully appreciated.

Keywords: diathesis stress, differential susceptibility, G×E interaction, gene selection

Introduction
Ever since the turn of the century, much attention has been paid to Gene × Environment (G×E) interplay (Rutter, 2006). Such interplay has been studied for one of two purposes: to illuminate the differential influence of the environment on individuals who differ in their genetic makeup (e.g., Caspi et al., 2002, 2003, 2005) or to highlight the role the environment plays in determining whether genetic differences are manifested in psychological and behavioral functioning (e.g., Harden, Turkheimer, & Loehlin, 2007; Turkheimer, Haley, Waldron, D’Onofrio, & Gottesman, 2003). It is the former G×E interaction perspective—the one that regards genes as moderating environmental influence, rather than the environment moderating genetic influence—that is the focus herein.

Multiple critiques of G×E research (as well as of genotype—phenotype work) have emerged in recent years (e.g., Bakermans-Kranenburg & van IJzendoorn, 2015; Manuck & McCaffery, 2014; Moffitt & Caspi, 2014; Moffitt, Caspi, & Rutter, 2005, 2006; Salvatore & Dick, 2015). Perhaps most important among the criticisms has been limited statistical power due to both small samples and the extremely small effects of individual genes vis-à-vis most phenotypes of interest to psychological and behavioral scientists (e.g., psychopathology, educational achievement). Concern has also been expressed for multiple testing without adjustment for chance results. Here, though, we raise an issue that we believe has been under-discussed, if considered at all, despite its apparent importance for the investigation of G×E interaction—the selection of genes hypothesized to moderate environmental effects and the associated theoretical assumptions on which they are implicitly or explicitly based.

A critical—and perhaps most important first—step in molecular G×E studies involves the selection of genetic variants to investigate. Three distinct, even if overlapping, phases of such molecular–genetic G×E inquiry can be distinguished, as well as several subphases. Our goal in doing so is to illuminate often unstated—or at least under-appreciated—conceptual assumptions underlying each phase and subphase when it comes to selecting putative moderating genes. The three major phases target (a) single candidate genes and (b) composited (multiple) candidate genes, sometimes based on the “biological plausibility” that they will be involved in G×E interaction; and (c) polygenic scores based on genome-wide association studies (GWAS). The goal of highlighting what we regard as under-discussed issues raised by each of the three approaches seems especially called for at a time when the prevailing view seems to be that GWAS-based—and theory-free—polygenic scores should be the strategy of choice when it comes to studying molecular genetics and human development, including G×E inquiry. Here we don’t so much reject that view as qualify it.

More specifically, we develop the argument that the selection of genes across the three major phases just mentioned—and to
be discussed in detail herein—has been implicitly or explicitly guided by one or both of two contrasting theoretical frameworks: diathesis stress (Monroe & Simons, 1991; Zuckerman, 1999) and differential susceptibility (Belsky, 1997a, 1997b, 2005; Belsky, Bakermans-Kranenburg, & van IJzendoorn, 2007; Belsky & Pluess, 2009, 2013a, 2013b; Ellis, Boyce, Belsky, Bakermans-Kranenburg, & van IJzendoorn, 2011). Thus, (a) we first outline the two theoretical frameworks (implicitly or explicitly) guiding much G×E inquiry before considering (b) their statistical implications when testing G×E interaction, and (c) examples of each of phase (and subphase) of G×E research. In a fourth and final section we discuss some issues of existing G×E research. This effort should not be regarded as an exhaustive review of G×E literature, but instead as an illustrative analysis of approaches adopted across the aforementioned phases when selecting genetic variants for G×E investigation.

**G×E Theoretical Frameworks**

The earliest G×E studies were guided by diathesis stress thinking, which posits that some individuals are particularly likely to develop problematically as a result of two conditions: their genetic makeup and their exposure to adverse contextual conditions (e.g., negative life events, divorce). Genes, as well as other organismic factors (e.g., temperament), may thus operate as “diatheses,” making individuals especially vulnerable to negative effects of contextual adversity (Monroe & Simons, 1991; Zuckerman, 1999). This dual focus on contextual adversity and organismic vulnerability, including genotype, which developmentalists discuss in terms of “dual risk” rather than diathesis stress (Belsky & Pluess, 2009), first emerged in an attempt to understand the etiology of psychopathology. Such concern no doubt accounts for the focus on putative “risk genes” and adverse experiences/exposures that predispose individuals to develop problematically.

Genes involved in G×E research informed by the diathesis stress framework are typically selected based on either (a) an inferential analysis of biological plausibility—for example, a particular gene is related to a neurotransmitter, which is related to a phenotype—and/or (b) genotype–phenotype studies documenting a statistical association between some genetic variant and the problematical functioning that is to be predicted (e.g., depression). Although diathesis stress research has contributed to the understanding of genetic and environmental contributors to psychopathology, the following question seems to rarely be considered: Why would natural selection preserve genetic variants that contribute to carriers being especially vulnerable to the negative effects of contextual adversity, especially in ways that might be expected to undermine reproductive fitness (Belsky, 2005; Belsky et al., 2007; Belsky & Pluess, 2013a, 2013b; Ellis et al., 2011)?

Differential susceptibility theory addressed this issue, thereby generating an alternative conceptual framework for G×E studies. Inspired by evolutionary thinking, differential susceptibility theory stipulates that individuals vary along a continuum of susceptibility to environmental influences due to their genetic makeup (and/or other organismic factors; Belsky & Pluess, 2009, 2013a, 2013b). Thus some are more developmentally plastic or malleable and others less so. This view is based on the argument that because the future is uncertain, there is no way of knowing whether development shaped by early-life experiences will pay off—reproductively. When the future environment is consistent with the childhood one, benefits should accrue to those shaped by their early-life experiences because such individuals would be especially likely to “fit” or “match” the environment in which they find themselves later in life. But when a mismatch occurs between childhood and future environments—given that the future is uncertain—such individuals highly susceptible to early-life environmental effects should pay a (reproductive) cost due to their being poorly prepared for their later-life contexts (Belsky, 1997a, 1997b, 2005; Belsky et al., 2007; Belsky & Pluess, 2009, 2013a, 2013b; Ellis et al., 2011).

Differential susceptibility theorizing is thus based on the idea that, due to the inherent uncertainty of the future, nature has “hedged its bet,” making some individuals more and others less susceptible to developmental and other contextual influences. What this implies, then, is that those individuals regarded as most vulnerable to adverse effects of environment according to diathesis-stress theorizing are the same ones who reap the most benefit from supportive or enriched environments, or even just benign ones (Belsky & van IJzendoorn, 2017). According to this view, what diathesis-stress theorizing regards as “risk alleles” (Burmeister, McInnis, & Zöllner, 2008), making individuals especially susceptible to the negative effects of adversity, are redefined as more general “plasticity alleles,” making individuals especially susceptible to environmental influences “for-better-and-for-worse” (Belsky et al., 2007).

Given the evolutionary foundations of differential susceptibility theory theorizing (Belsky, 1999; Belsky & Pluess, 2009), we would be remiss if we failed to acknowledge that the modern world—where all the research cited herein has been undertaken—is no doubt quite different in many respects from our ancestral one, including just several hundred years ago.1 Most notably, the range of environments that have been studied vis-à-vis G×E interaction is more limited than many that once existed—and indeed were rather common. In populations studied today, powerful adverse conditions—such as starvation, epidemic disease, war, and predation on humans—are often absent, even if they remain all-too-common in other populations that are not (yet) investigated. This raises the very real possibility that available research captures a rather truncated end of the stress continuum. This is a challenge for both risk and susceptibility perspectives, even raising the possibility that tests of these alternative G×E hypotheses in more challenging environments could have greater power than the work cited herein and in the general G×E literature.

**Statistical implications of contrasting G×E models**

From a statistical perspective, diathesis–stress models delineate G×E interactions in an ordinal manner (Widaman et al., 2012). That is, some individuals function more poorly than others as a result of an adverse experience/exposure, but no better than others in the absence of the adversity under investigation. In consequence, graphical plots of such results—see Figure 1a—will be ordinal in character, characterized by two (or more) simple regression lines that will not cross within the observed range of environmental measurement or will cross at the extreme end of the observed environment variable (which may just reflect the absence of the adversity under investigation and thus reflect benign as well as supportive conditions).

In contrast, differential susceptibility models presume G×E interactions will be disorderal in form (Widaman et al., 2012). In this case, the power of the environmental parameter to predict

---

1Indeed, we were remiss until a helpful reviewer brought our attention to our omission.
that would make it consistent with the diathesis
the chance the interaction detected will be ordinal in form.
Environmetal variable (i.e., from
crossover point, especially when paired with limited range of envi-
unchanged. And the larger the estimated absolute value of the
value of the crossover point, with the interaction effect
the tested G×E interaction, the larger the estimated absolute
(main) effect of genotype,
1991; Cohen, Cohen, West, & Aiken, 2003, pp. 288
main effect of the moderator to the inte
the crossover point is determined by the
predicted value of crossover point and the range of environmental
Environment interaction. In addition, the location of the crossover
thesis
—
distinction of G×E findings consistent, respectively, with the dia-
dinal and ordinal forms of interaction and, thereby, the empirical
expectations of G×E findings consistent, respectively, with the dia-
thesis–stress and differential susceptibility models of Person ×
vironment variable. In addition, the location of the crossover
point with respect to the environment variable is determined by the
predicted value of crossover point and the range of environmental
predictor. It is important to appreciate that the predicted value of
of the crossover point is determined by the relative magnitude of the
main effect of the moderator to the interaction (Aiken & West, 1991; Cohen, Cohen, West, & Aiken, 2003, pp. 288–289).
As made mathematically explicit in Appendix A, the larger the
(main) effect of genotype—the would-be moderator—relative to
the tested G×E interaction, the larger the estimated absolute
value of the crossover point, with the interaction effect
unchanged. And the larger the estimated absolute value of the
crossover point, especially when paired with limited range of envi-
ronmental variable (i.e., from “adverse” to “not adverse” rather
than from “adverse” to “supportive/enriching”), the greater
the chance the interaction detected will be ordinal in form.
That would make it consistent with the diathesis–stress model
of Person × Environment interaction.

Therefore, two factors need to be considered when it comes to
distinguishing ordinal (i.e., diathesis stress) and disordinal (i.e.,
differential susceptibility) forms of G×E interaction. One is the
predicted crossover point, as already noted, calculated by the
main effect of the moderator relative to the interaction effect;
the other is the location of the crossover point with respect to the
environmental variable, determined by the estimated value of
the crossover point and the range of the environment variable.
G×E work, then, which documents a significant main effect of
genotype, is most likely to prove consistent with the diathesis–
stress model, especially when the environmental predictor covers
only a limited range of the possible environment. In contrast,
an insignificant main effect of genotype, which is more likely to be
characteristic of a disordinal interaction, is more likely to prove
consistent with differential susceptibility theory, especially when
the environmental predictor covers full range of environment
(i.e., from “adverse” to “supportive/enriching”). This observation,
which does not hold 100% of the time (i.e., “more likely” rather
than “always”), will be returned to when GWAS derived polygenic
scores are discussed in the third major phase of G×E inquiry to be
considered below.

**Phases of G×E Inquiry**

Returning to the previously delineated phases of G×E inquiry
based on the selection of moderating genes, we begin by consider-
ing work using single candidate genes (Phase 1) before turning to
the typically later phases which rely on multiple genes, sometimes
based on prior G×E research and sometimes based on GWAS
findings. To repeat, our purpose is to highlight assumptions,
often unstated, central to the selection of genes in each of these
phases and, as a result, their influence on which G×E model
under consideration is most likely to receive empirical support.

**Phase 1: Single candidate genes**

Here we consider two distinct conceptualizations of single candidate
genes, the first emphasizing the risk they pose to wellbeing (i.e.,
“risk alleles”) and the second based on for-better-and-for-worse dif-
ferential susceptibility thinking (i.e., “plasticity alleles”).

**Risk alleles**

In the first (and continuing) phase of G×E research, studies relied
on a single candidate gene as a genetic moderator of an environ-
mental effect in predicting, typically, a psychiatric disorder or
symptoms of a disorder, including but not limited to antisocial
behavior, substance use, and depression. The earliest work of
this kind was based exclusively on diathesis stress thinking, as it
focused on contextural adversity (rather than the full range of envi-
ronment, from adverse to supportive) in predicting some form of
problematic functioning (rather than some index of competent
functioning, e.g., educational achievement), while relying on
genes thought to directly influence the outcome in question. In
such work, then, both the contextual condition and the genetic
factor are regarded as risk factors. Consider in this regard Caspi
and associates’ (2002) pioneering work testing the interplay of
childhood maltreatment and the monoamine oxidase A
(MAOA) polymorphism in predicting later antisocial behavior.
The MAOA gene, located on the X chromosome, encodes the
MAOA enzyme, which metabolizes neurotransmitters such as
norepinephrine, serotonin, and dopamine, rendering them inac-
tive (Shih, Chen, & Ridd, 1999). This study focused on a

---

**Figure 1.** Graphic representation of prototypic diathesis stress (a) and differential susceptibility (b) models of Organism × Environment interaction.
functional polymorphism in the promoter of the MAOA gene to characterize genetic vulnerability to maltreatment because the neurotransmitters just mentioned had all been linked to aggressive behavior in humans and mice (e.g., Cases et al., 1995; Manuck, Flory, Ferrell, Mann, & Muldoon, 2000; Rowe, 2001; Shih & Thompson, 1999), thereby making this gene a "biologically plausible" candidate when investigating G×E interaction and antisocial behavior.

Notably, Caspi and colleagues (2003) adopted the same biological-plausibility approach in selecting a different polymorphism in their second G×E study informed by diathesis stress thinking. These investigators hypothesized that the serotonin transporter (5-HTTLPR) gene—and the short allele of the specific gene linked polymorphic region (serotonin transporter linked polymorphic region (5-HTTLPR))—would operate as a second risk factor, amplifying the effect of life stress, the first risk factor, on depression. 5-HTTLPR was targeted based on its previously documented influence on the reuptake of serotonin at brain synapses (Lesch, Greenberg, Higley, Bennett, & Murphy, 2002) and the role that this neurotransmitter plays in the etiology of depression (Tamminga, 2002). Specifically, Caspi and colleagues (2003) focused on the short allele of 5-HTTLPR gene because it had been linked to a decreased level of serotonin which, in turn, was thought to undermine psychological wellbeing.

**Plasticity alleles**

Consideration of the findings emerging from the Caspi et al. (2002; 2003; 2005) research and other diathesis-stress-informed G×E, single candidate gene work (e.g., Eley et al., 2004; Grabe et al., 2005; Kaufman et al., 2004; Sheese, Voelker, Rothbart, & Posner, 2007) stimulated Belsky and Pluess (2009) to reconsider—and re-interpret—the nature of the G×E interactions being discerned. They noted that in many cases it appeared that those individuals carrying putative "vulnerability genes" (Rutter, 2006) or "risk alleles" (Burmeister et al., 2008) not only proved more susceptible to the negative effects of adversity, but also seemed to do better than others when contextual adversity was absent (e.g., not maltreated, few negative life events). In consequence, multiple candidate genes that had been regarded as vulnerability factors or risk genes based on arguments of biological plausibility and diathesis stress thinking appeared to operate as more general plasticity factors, in line with differential susceptibility theorizing. Indeed, this possibility led Taylor and associates (2006) to wonder what G×E findings would look like if environmental measurement ranged from negative to positive rather than, as typical of most diathesis-stress-informed G×E work, from adverse to not adverse. When these investigators examined the same G×E interaction predicting depression that Caspi and associates (2003) had, but included positive not just negative life events in their environmental measurement, results revealed for-better-and-for-worse responsiveness in the case of 5-HTTLPR short-allele carriers.

Consistent with these differential-susceptibility-related results are those of other candidate G×E investigations focused on 5-HTTLPR, which sought to predict anxiety (Gunnther et al., 2007; Stein, Schork, & Gelernter, 2008) and attention-deficit/hyperactivity disorder (ADHD) (Retz et al., 2008) and/or evaluate effects of negative life events (Gunnther et al., 2007), childhood emotional abuse (Stein et al., 2008), and a generally adverse child-rearing environment (Retz et al., 2008)—even if originally conducted to test diathesis stress hypotheses. Notably, a priori tests of differential susceptibility thinking using 5-HTTLPR as the genetic moderator have also yielded results consistent with this conceptual framework, though once again the for-better-and-for-worse predictions were based on the results of prior G×E studies, including ones based originally on diathesis stress thinking, not primarily or exclusively on ideas about biological plausibility. Consider in this regard studies focused on the effects of prenatal maternal anxiety on infant negative emotionality (Pluess et al., 2011); of positive parenting on children's positive mood (Hankin et al., 2011); and of prenatal depression on temperament dysregulation of infants (Babineau et al., 2014). Perhaps even more noteworthy evidence in line with differential susceptibility thinking emerged from Brody, Beach, Philibert, Chen, and Murry's (2009) randomized control trial (RCT) which found that a multi-faceted family-based intervention program for rural African-American teenagers disproportionately benefited 5-HTTLPR short-allele carriers when it came to prevention of risky behavior. Once again, then, a long-regarded risk gene was found to operate as an "opportunity" factor. (For a review of such studies investigating genetic moderation of RCTs of parenting interventions, see Belsky & van IJzendoorn, 2017.)

The 7-repeat variant of dopamine receptor D4 (DRD4) gene provides another example of candidate G×E research based initially on biological plausibility and diathesis stress thinking that served to stimulate tests of differential susceptibility theorizing based on prior empirical findings (i.e., not biological plausibility claims). The dopamine system is involved in attentional, motivational, and reward mechanisms (Robbins & Everitt, 1999). The 7-repeat DRD4 was originally considered a vulnerability factor due to its statistical association with lower dopamine reception efficiency (Robbins & Everitt, 1999) and psychological maladjustment (e.g., ADHD: Faraone, Doyle, Mick, & Biederman, 2001; high novelty-seeking behavior: Kluger, Siegfried, & Ebstein, 2002). Such evidence provided the basis for, initially, the biologically plausible, diathesis stress hypothesis that maternal insensitivity would predict greater externalizing behavior, especially in the case of those carrying the 7-repeat allele (Bakermans-Kranenburg & van IJzendoorn, 2006). Results revealed, however, that this genetic variant functioned more as a general plasticity than vulnerability factor; and the same proved true when the same investigators turned their attention to the effects of maternal unresolved loss or trauma on children's attachment disorganization (Van IJzendoorn & Bakermans-Kranenburg, 2006). In fact, these results served as the basis for Bakermans–Kranenburg and associates' (2008) RCT, which found that carriers of the 7-repeat allele of DRD4 disproportionately benefited from a parenting intervention designed to reduce child externalizing behavior. Related thinking about this allelic variant led Beach, Brody, Lei, and Philibert (2010) to discover that the same was true when they investigated the efficacy of a substance use intervention for rural African-American adolescents.

Notably, not all G×E studies testing the differential susceptibility hypothesis selected candidate genes based on the re-interpretation of findings from diathesis-stress-based studies, but rather ones often based on biological plausibility arguments. That is, some candidate G×E research testing differential susceptibility thinking also relied on the claim of biological plausibility. Consider in this regard Simons and colleagues' (2013a, 2013b) investigation of Gamma-Aminobutyric Acid Type A Receptor Subunit Gamma1 gene (GABRG1) and social environmental factors in predicting substance use. These researchers selected the GABRG1 polymorphism based on the observation that it is expressed primarily in the amygdala and areas receiving innervation from the striatum such as the substantia nigra (Pirker,
Schwarzer, Wieselthaler, Sieghart, & Sperk, 2000; Schwarzer et al., 2001), with the former implicated in response to punishment (Margules, 1971) and the latter in response to reward (Nazzaro, Seeger, & Gardner, 1981), as well as to addiction (Hurd & Herkenham, 1993). In a second candidate G×E study by Simons and associates (2013a, 2013b) testing differential-susceptibility thinking, the same biological-plausibility logic led to a focus on GABA receptor subunit alpha-2 (GABRA2) and its interaction with harsh parenting in predicting hostility toward romantic partners.

To summarize, then, candidate gene work investigating G×E interaction was initially guided by diathesis stress thinking, with the genetic moderator selected based on arguments of biological plausibility, including associations with the phenotype to be explained, and conceptualized as a second risk factor (in addition to the contextual one). Evolutionary reasoning coupled with evidence from this initial body of diathesis-stress-informed G×E work raised the prospect that individuals did not just vary in terms of their susceptibility to the negative effects of adversity, but to both negative and positive environmental exposures. Such thinking stimulated both observational and experimental research testing differential susceptibility theorizing, with the genes selected sometimes based on biological plausibility claims and sometimes just on results of prior G×E research.

Phase 2: Initial multiple gene strategies

Developmental scholars studying effects of the environment on development have long appreciated the empirical utility of looking at the collective effect of multiple environmental parameters. Indeed, the “cumulative risk” strategy is based on the view that it is not so much a single adversity that undermines developmental wellbeing, but the accumulation of risk factors. This led to the creation of cumulative risk scores, which simply reflect the number of putative contextual risks to which a person is exposed. Thus, investigators might have a list of 10 risk conditions (e.g., poverty, single parent, maternal depression, child abuse) and score children in terms of the total number which applies to them. This score is then used to predict later development, with extensive work showing that the more risks to which a child has been exposed, the poorer the child’s development (e.g., Evans, Li, & Whipple, 2013; Rutter, 1979, 1981; Sameroff, Seifer, Barocas, Zax, & Greenspan, 1987).

It did not take long for G×E researchers to implement a related approach involving combinations of individual candidate genes. As we will see, this was initially done based simply on prior candidate-gene findings rather than explicit “biological plausibility” claims, but soon thereafter would-be plasticity genes were composited based on such biological-plausibility arguments.

Cumulative gene scores based on candidate-gene findings

Belsky and Beaver (2011) were the first to appreciate that the cumulative-predictor approach could be applied to the study of genetic moderation of environmental influences, making them the first G×E investigators to rely on what they termed “cumulative genetic plasticity” (based on five putative “plasticity alleles”: the 10R of dopamine active transporter 1 (DAT1), the A1 of DRD2, the 7-repeat of DRD4, the short allele of 5-HTTLPR, and the 2R/3R of MAOA). Explicitly testing differential susceptibility theory using what they considered a genetic plasticity score—not a genetic risk score—these researchers observed that the more putative plasticity alleles teenage boys carried, the more and less self-regulated they were under, respectively, supportive and unsupportive parenting conditions. Rauscher (2017) also found evidence consistent with differential susceptibility thinking upon relying on the exact same genetic plasticity score as Belsky and Beaver (2011) when studying the financial success of siblings. Just as notable, perhaps, was Stocker and colleagues’ (2016) success in documenting differential susceptibility to environmental influence when creating a similar multiple-gene index using four putative plasticity genes each implicated in prior G×E work yielding differential susceptibility like results (5-HTT, DRD2, DRD4, and catechol-O-methyltransferase (COMT)).

Others have followed this same strategy of creating cumulative genetic plasticity scores—based on prior candidate G×E findings—to test the differential susceptibility hypothesis. Consider in this regard Masarik and colleagues’ (2014) work using a five-gene index (serotonergic-related genes: 5-HTTLPR; dopaminergic-related genes: DRD2, DRD4, dopamine active transporter gene, DAT; catechol-O-methyltransferase gene, COMT) to test the effect of parenting quality in childhood on romantic relations in adulthood; Simons and associates’ (2011) study, which combined just two putative plasticity genes, DRD4 and 5-HTTLPR, to serve as moderators when investigating the effect of favorable/adverse social environments on aggression among African-American adolescents; and, as final examples, work which combined the same two polymorphisms to test the effect of life stress on life history strategies (e.g., future orientation, risky sexual behavior) of adolescents (Gibbons et al., 2012) and prenatal maternal depression on infant negative emotionality (Green et al., 2017).

Important to note is that in all cases just highlighted, there was no explicit “biological plausibility” basis for combining genes. Composited genes were chosen simply because they had been found to operate, as single candidate genes, in differential susceptibility fashion in prior G×E work and were available in the data sets being used. To be clear, however, in this prior, foundational work, the genes were selected for study on the basis of “biological plausibility” claims.

Bio-plausibility-based cumulative gene strategies

As it turns out, biological-plausibility thinking also led to combining candidate genes for use in G×E research. Indeed, theory and evidence suggested that, whereas some genes were principally considered serotonergic in character, others were dopaminergic, meaning that these genes played important roles in the functioning, respectively, of these two different, even if related, neurotransmitter subsystems. Thus, whereas serotonergic genes were considered to play an especially important role in shaping sensitivity to punishment and displeasure, dopaminergic genes were thought to be particularly influential in the case of reward sensitivity and sensation seeking (Carver, Johnson, & Joormann, 2008; Frank, Moustafa, Haughey, Curran, & Hutchison, 2007).

Certain alleles of the relevant genes were thus hypothesized to predispose some individuals to be more responsive to their environment than others because they influence thresholds for sensing pleasure or displeasure on the basis of environmental cues. That is, due to genetic endowment, some individuals were considered to be more readily shaped than others by environmental rewards or/and punishments. This line of thinking led to the creation of distinct serotonergic and/or dopaminergic composite scores.2

2It is not unusual to see investigators use the terminology of “risk” even when advancing general-plasticity rather than risk-factor hypotheses (e.g., Cook & Fletcher, 2015; Sun
Investigators guided by diathesis stress thinking have relied on biological plausibility arguments when using cumulative genetic scores reflective of neurotransmitter processes, treating their genetic composites as a risk factor amplifying the effects of contextual adversity. An example of such an approach can be found in Vrshke–Schallhorn and associates’ (2015) attempt to predict depression. These investigators selected five single nucleotide polymorphisms (SNPs) located in serotoninergic genes not only because the serotoninergic system is implicated in depression (Booij, Van der Does, & Riedel, 2003; Fournier et al., 2010), but because each of the five SNPs had been linked, individually, to depression in genotype–phenotype research (Anttila et al., 2007; Brummett et al., 2014; Drevets et al., 2007; Gao et al., 2012; Li, Duan, & He, 2006). As such, the five composited SNPs were conceptualized as a second risk factor in this dual-risk-related investigation.

In diathesis-stress-informed work, although serotoninergic genes are often the focus when the outcomes being predicted are internalizing spectrum disorders (e.g., depression, anxiety), it is dopaminergic genes that are often selected when the outcomes to be explained reflect externalizing problems (e.g., aggression, antisocial behavior, substance abuse). And once more, this is due to their hypothesized, biologically plausible role in shaping reward sensitivity and attentional processes (Janssens et al., 2015).

In view of what has just been observed, it seems noteworthy that meta-analysis indicates that dopaminergic polymorphisms moderate environmental effects in a differential-susceptibility-related manner (Bakermans-Kranenburg & van IJzendoorn, 2011). This underscores the need to reconsider whether a dopaminergic polygenic score should be regarded as a general plasticity factor (polygenic plasticity score, PPS) rather than a risk factor (polygenic risk score,PRS). One investigation underscoring this point was carried out by Thibodeau, Cicchetti, and Rogosch (2015) who composited five genetic variants in four dopaminergic genes (i.e., DRD4, DRD2, DAT1, COMT) and evaluated the effect of number of maltreatment subtypes (i.e., neglect, emotional maltreatment, physical abuse, and sexual abuse) on antisocial behavior (i.e., aggression, delinquency, and disruptive peer behavior). It turned out that the putative genetic risk index operated more as a plasticity index. A similar, for-better-and-for-worse moderating role of dopaminergic genes also emerged when predicting (a) adolescent delinquency using attachment to school (DRD4, DRD2, DAT1; Fine et al., 2016), (b) adolescent externalizing problems using constructive and destructive interparental conflict (DRD4, COMT, DAT1; Davies, Pearson, Cicchetti, Martin, & Cummings, 2019), and (c) 9-year-old children’s telomere length using social disadvantage (Mitchell et al., 2014).

Most recently, multiple genes known to be involved in the functioning of the hypothalamus–pituitary–adrenal axis (HPA) also have been composited in GxEx research, this time based on the role this system is considered to play in stress and depression. Using three HPA axis genes (corticotropin releasing hormone receptor 1 (CRHRI), FK506 binding protein 5 (FKBP5), and nuclear receptor subfamily 3, group C, member 2 (NR3C2)), Feurer and colleagues (2017) investigated the association between interpersonal stress and depresssive symptoms in early childhood.

Similarly, McKenna, Hammens, and Brennan (2020) created a somewhat similar three HPA-gene index which was expected to amplify the risk of prenatal stress on offspring depression at age 20.

In sum, composited genetic scores used in GxEx research have been created on the basis of prior candidate GxEx findings, mostly in the absence of any arguments about biological plausibility, as well as scores founded on biological plausibility arguments. Notably, when such research was undertaken to test diathesis stress hypotheses, genes were selected based on biological plausibility arguments. When, however, candidate genes were combined to test differential susceptibility thinking, it was sometimes based on biological plausibility claims and other times simply based on compositing genes found in prior candidate-gene work to moderate some environmental effect in a for-better-and-for-worse manner.

### Phase 3: GWAS-derived polygenic scores

As failures to replicate both genotype–phenotype and GxEx findings emerged, the limits of focusing on single and even several candidate genes—when creating genetic risk or plasticity scores—came to be widely appreciated. GWAS studies reliant on many thousands of cases played a critical role in highlighting the potential benefit of polygenic scores based on hundreds, even thousands—and sometimes millions—of SNPs that proved to be associated, even if very weakly, with a particular phenotype at a statistically significant level after accounting for multiple testing. Obviously, this raised questions about the practice of focusing on single—or even just a few—candidate genes in both genotype–phenotype and GxEx research. Here we consider the use, then, of GWAS-derived polygenic scores in testing both diathesis stress and differential susceptibility hypotheses.

Two critical points must be underscored—and are often under-appreciated—when considering polygenic scores derived from GWAS in GxEx research. The first is that these are based exclusively on SNPs and thus do not consider another form of genetic variation, indeed, one that has proven popular—and productive—in GxEx research. Here we are referring to tandem repeats (TRs), such as in the case of DRD4 in which “7-repeats” have been distinguished from other variants of this polymorphism. This raises the question of the adequacy of GWAS results for informing GxEx inquiry.

The second critical point is that the identification of SNPs that correlate with a psychological, behavioral or other “trait” or personal attribute is that they, collectively, reflect the cumulative main effects of numerous individual genes vis-à-vis a particular phenotype (after accounting for multiple testing). Consider in this regard, the distinctive GWAS-derived PRS for obesity and smoking created by Belsky, Moffitt, and Caspi (2013) which predicted, respectively, rapid growth in early life and accelerated progress from smoking initiation to heavy smoking. Critically important to appreciate in all such work is that no concerns regarding biological plausibility play a role in the polygenic scores derived from such genome-wide inquiry. Genes considered to influence a phenotype are identified—and compositied—based simply on the extent to which they correlate with the phenotype in question.

---

3Problems related to ethnic variation in SNP frequencies may challenge the development of SNP-based PRS/PPS because GWAS-identified SNPs are usually not variants that...
**Diathesis stress**

Although we will eventually question the wisdom of using these kinds of GWAS-derived polygenic scores to test differential-susceptibility hypotheses, this is clearly an appropriate way to proceed when testing predictions based on diathesis stress thinking. To make clear this latter point, it helps to re-call that diathesis stress thinking reflects a focus on “dual risk.” Thus, it is when two risks or presumptive main effects, one contextual (e.g., child maltreatment) and the other organismic (e.g., genetic), co-occur that the anticipated adverse effect of each is most likely to be realized; one alone is much less likely to predict some problematic manner of functioning (e.g., depression, substance abuse). And this is because, again, the effect of each is presumed to amplify the effect of the other. Were this not the case, only additive main effects would be discerned, not a G×E interaction.

One good example of this dual-risk, polygenic approach to testing diathesis stress thinking is found in the work of Peyrot and associates (2014). These investigators first built on a GWAS “discovery” study of more than 15,000 individuals which resulted in the creation of a PRS for major depressive disorder. They then used the resultant PRS algorithm with another sample, finding that exposure to childhood trauma (i.e., Risk 1) coupled with higher PRS (i.e., Risk 2) predicted depression in a study of Dutch adults. Notably, these G×E results were extended in yet another investigation of depression, this one of more than 150,000 individuals in the “discovery” sample, which also showed that the same PRS interacted in a diathesis-stress manner with personal life events (Colodro-Conde et al., 2018).

GWAS-derived polygenic scores also have proven useful in testing diathesis–stress hypotheses concerning externalizing problems. Salvatore and Dick (2015) first identified 176,562 SNPs that survived multiple testing in their association with externalizing disorders in what, from a GWAS perspective, must be considered a small sample of adults (n = 1,249). These SNPs, weighted by the strength of their association with the phenotype in question in the discovery sample, were then found to predict externalizing disorders in a sample of 455 adolescents and young adults. Most important for our purposes was the diathesis-stress finding indicating that the link between PRS and externalizing disorders was stronger when peers were substance users and/or parents failed to monitor their offspring than when these contextual conditions did not present. These results were replicated, in part, by Sadeh and colleagues (2016) in their work with military veterans. As hypothesized, a higher externalizing PRS predicted more externalizing psychopathology and negative inhibitory control in interaction with lifetime number of trauma experiences, though, intriguingly, results proved more in line with differential susceptibility than diathesis-stress theorizing.

**Differential susceptibility**

In recent years, investigators have adopted, perhaps unwisely, the dual-risk approach to creating GWAS-derived polygenic scores (i.e., SNPs directly associated with the outcome to be predicted in G×E work) even when explicitly testing differential susceptibility predictions. Consider, as an example, Sun et al.’s (2018) work on childhood obesity, which relied on an 11-SNP PRS—not a PPS—that had been found to predict child obesity in GWAS discovery work with more than 35,000 individuals of European ancestry (Felix et al., 2016). Even if the PRS in question moderated the effect of cumulative stress (as indexed by cortisol in hair) on child obesity in this study of 1,000 Chinese children, one can wonder why a PRS based on genotype–phenotype associations (i.e., main effects) would be expected to operate in differential susceptibility fashion, as it, in fact, did. Recall from our earlier summary of statistical issues (along with Appendix A) that using a moderator that exerts a main effect predisposes an interaction to be ordinal in character, consistent with diathesis-stress theorizing, even if that is not always the case. So the question becomes, are PRS scores based on GWAS for the very phenotype being predicted in G×E work the most appropriate way to test differential susceptibility thinking? We think not.

While the hypothesis-free GWAS approach—which in no way considers claims of biological plausibility when compositing genes—may be very well suited for detecting genetic variants to be used in testing explicit diathesis–stress/dual-risk hypotheses, one can question, as we just have, its utility for testing predictions based on differential susceptibility thinking. And this is because there is no reason to presume that GWAS-derived polygenic scores—based as they are on main-effect associations linking individual SNPs with the very phenotype to be explained in G×E inquiry—should moderate an environmental predictor in a for-better-and-for-worse manner. After all, interactions can occur between two factors when one, or even both, do not directly predict the outcome in question (Aiken & West, 1991). Thus, one is forced to wonder why a moderator not presumed to function in a dual-risk fashion, should even be selected for use in a G×E based on genotype–phenotype associations in GWAS. Recall in this regard comments made earlier about the statistical properties of disorderal interactions characteristic of differential-susceptibility-related results: they typically, even if not always, require that at least one of the predictors does not exert a main effect!

These observations, as it turns out, do not imply that GWAS-derived polygenic scores cannot be used to test differential-susceptibility-related predictions in G×E work. But, in order to do so, it would seem that the PPS—not PRS—should be based on SNPs that are related to some index of plasticity rather than, as in diathesis-stress work, the phenotype to be explained. In light of this observation, recent work by Keers and associates (2016) seems especially notable, as it implemented a most original and creative approach to identifying genes related to plasticity in order to test differential susceptibility theorizing; we would be remiss if we did not make clear that the work to be summarized was itself “theory-free” in design. These investigators first identified SNPs associated with phenotypic differences in anxiety within 1,026 monozygotic (MZ) twin pairs, based on the idea that identical twins who differed across this (or any other) phenotype would be genetically different from twins who did not differ from each other. Indeed, the thinking was that the former twins would prove phenotypically different because of their sensitivity to the environment and their differential experiences and exposures while growing up, whereas the latter would not because of their limited sensitivity to the effects of any such differences in developmental experiences or contextual exposures. In other words, the GWAS-identifed SNPs were presumed to be plasticity, not risk factors, because genetic variants that increase environmental sensitivity should increase discordance within
MZ twin pairs on outcomes as they render each member of the pair more responsive to the nonshared environment. Of note in this regard is that behavior–genetic research consistently indicates that nonshared environmental effects prove more powerful than shared ones (Hughes & Plomin, 2000), though some of this difference is surely due to the (all-too-often unstated) fact that the estimate of the nonshared environment parameter includes measurement error.

With a PPS algorithm in hand—based on SNPs associated in the discovery sample with greater versus lesser within-twin differences in anxiety—Keers and associates (2016) proceeded, using two new samples, to address two G×E questions. The first was whether this PPS for plasticity moderated the effect of parenting on emotional symptoms in a manner consistent with differential susceptibility. Once the results of this observational study yielded evidence consistent with this expectation, these same investigators turned to a treatment–response sample comprising children and adolescents with anxiety disorders in order to address their second plasticity-related question: Would the PPS help to explain the differential efficacy of cognitive-behavioral therapy (CBT)? It turned out that it did! Although Keers and associates did not detect a moderating effect of PPS on overall response to treatment, they did discern an interaction between PPS and treatment intensity: children with high PPS benefited the most from individual CBT than less intensive forms of group or brief parent-let CBT, whereas those with low PPS responded equally to each treatment type.

Perhaps due to the relative recency of Keers et al.’s (2016) approach to deriving a PPS based on GWAS with MZ twins, we are aware of only a single other inquiry that has followed up on their work. Lemery-Chalfant, Clifford, Dishion, Shaw, and Wilson (2018) formed a PPS based on Keers et al.’s (2016) GWAS and used it to evaluate genetic moderation of the efficacy of a parenting intervention with high-risk 10-year-old suffering from internalizing problems. Consistent with differential susceptibility theorizing, children scoring highest on the PPS benefited the most from the treatment, whereas their high-PPS-scoring counterparts in the control group manifest the most internalizing symptoms. Clearly, more work of this kind is called for—in an attempt to identify SNPs related to susceptibility to environmental effects—for better and for worse.

**Conclusion**

As noted in the Introduction, many criticisms have been wielded against molecular–genetic studies, be they of the genotype–phenotype or G×E variety. Here we have sought to illuminate an issue that we regard as under-discussed, if considered at all, in so much G×E research and commentary, namely, the selection of genes presumed to moderate environmental effects. We have relied upon two contrasting models of Organism × Environment interaction to illuminate this issue: diathesis stress/dual risk and differential susceptibility. Toward these ends, we have distinguished three major phases inquiry (see Table 1) when it comes to selecting genes for inclusion in G×E studies: (a) single candidate genes and (b) multiple candidate genes, sometimes based on the “biological plausibility” that they will be involved in G×E interaction; and (c) polygenic scores based on GWAS.

The first two phases of inquiry, in which one or a few candidate genes are selected for reasons of biological plausibility, prior genotype–phenotype findings, and/or the results of prior G×E results, seem unlikely to illuminate the nature of G×E interaction going forward, if only because of the now well-appreciated fact that virtually no phenotype of interest to psychological and behavioral scientists is influenced by such a limited number of genes. GWAS convincingly shows that thousands of genes appear important for understanding most phenotypes of interest. This insight, then, would seem to apply not only to typical phenotypes like antisocial behavior or educational achievement, but to developmental plasticity itself.

Based on our theoretical analysis, there is no reason to question reliance on GWAS-derived PRSs to test diathesis-stress-motivated G×E inquiry. This is because a PRS is based on the very phenotype that is to be predicted in G×E research, thus making it a very good organismic risk factor, in addition to the environmental predictor, to use when testing a dual-risk hypothesis. So, according to this way of thinking about G×E interplay, it is when risk factors amplify the effect of one another that evidence consistent with diathesis stress emerges.

But if the hypothesis to be evaluated, or at least entertained, is derived from differential-susceptibility thinking, there seems to be a fundamental problem with relying on GWAS findings that yield polygenic scores based on the fact that their components individually predict the phenotype of interest. What is needed is a strategy for identifying numerous genes that (likely) reflect susceptibility to environmental influence. This could be done in a theory-heavy way, based on ideas about the genetics of plasticity, or a theory-free way using GWAS. Clearly, Keers et al. (2016) have made an important contribution to this effort. One of the most interesting questions still to be addressed will be whether any identified plasticity SNPs will reflect sensitivity to environmental effects on specific outcomes or across diverse ones. One is forced to wonder, for example, whether the SNPs Keers et al. (2016) identified might—or might not—prove useful in testing differential susceptibility hypotheses with outcomes other than anxiety.

---

**Table 1. Features of gene selection for three phases (and subphases) of Gene × Environment interaction (G×E) studies**

| Phase (and subphase) | Number of genes | Gene selection approach | Underlying theoretical assumption |
|----------------------|----------------|-------------------------|----------------------------------|
| 1a                   | Single gene    | Biological-plausibility-based | Diathesis stress                 |
| 1b                   | Single gene    | Re-interpretations of identified risk genes | Differential susceptibility |
| 2a                   | Multiple genes | Identified plasticity genes | Differential susceptibility |
| 2b                   | Multiple genes | Biological-plausibility-based | Diathesis stress |
| 3a                   | Multiple (GWAS) Genes | Data-driven GWAS-derived | Diathesis stress |
| 3b                   | Multiple (GWAS) genes | GWAS-derived plasticity genes | Differential susceptibility |

GWAS: genome-wide association studies

---

https://doi.org/10.1017/S0954579420000966 Published online by Cambridge University Press
To advance the field it will be critical to move beyond the kind of GxE inquiry that has relied on exploratory tests of a GxE interaction, followed by inspection of slopes reflective of the predictor–outcome relation for those who vary on the would-be PPS index. Indeed, one must question why GxE studies based on diathesis stress thinking were carried out in such an exploratory fashion. If the prediction is that certain individuals, based on their genetic makeup, will and will not be affected by a particular experience or exposure, why conduct omnibus—and thus exploratory—tests of whether two factors interact. Why not go directly to testing the fit of the theoretically predicted results? By analogy, why conduct a two-tailed test when a one-tailed test based on a directional hypothesis is the statistical gage that better fits the theoretical hand?

Indeed, the common practice of running exploratory analyses even when strong theoretical ideas are being tested led to the development recently of two statistical methods designed to evaluate explicit predictions; and each has proven useful in some GxE research (e.g., diathesis stress: Bakermans & Pluess, 2013a, 2013b; differential susceptibility: Sun et al., 2018). Even though the method of Roisman and associates (2012) is to be used following an exploratory GxE result, it directly evaluates whether the GxE effect is in line with differential susceptibility thinking based on regions of significance (RoS) calculated through Johnson–Neuman (J-N) technique (Hayes & Matthes, 2009) and the proportion of interaction (POI) or proportion affected (PA) index.

The second approach eschews any exploratory test of GxE interaction, being designed to comparatively compare alternative conceptual models, most notably, the two central to this essay and which have informed so much GxE inquiry: diathesis stress and differential susceptibility (Belsky, Pluess, & Widaman, 2013; Belsky & Widaman, 2018; Widaman et al., 2012). This competitive and confirmatory modeling-test approach is based on a re-parameterized model comparison that evaluates, among other things, the critically important crossover point of regression lines. Ultimately, though, when a specific form of a GxE interaction—ordinal or disordinal—is to be tested, derived as it is from the underlying theoretical model guiding the work, the selection of genes should also be guided by whether it is a diathesis stress/ dual risk hypothesis being tested or a differential susceptibility one.

Conflict of Interest. We have no conflict of interest to disclose.

Acknowledgment. This paper has benefited from the helpful suggestions of two anonymous reviewers.

References
Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and interpreting interactions. Newbury Park, CA: Sage.
Anttila, S., Huhkka, K., Huhkka, M., Rontu, R., Hurme, M., Leinonen, E., & Lehtimäki, T. (2007). Interaction between 5-HT1A and BDNF genotypes increases the risk of treatment-resistant depression. Journal of Neural Transmission, 114, 1065–1068. doi:10.1007/s00702-007-0705-9
Babineau, V., Green, C. G., Jolicoeur-Martineau, A., Bouvette-Turcot, A. A., Minde, K., Sassi, R., ... Lydon, J. (2014). Prenatal depression and 5-HTTLPR interact to predict dysregulation from 3 to 36 months—A differential susceptibility model. Journal of Child Psychology and Psychiatry, 56, 21–29. https://doi.org/10.1111/jcpp.12246.
Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2006). Gene–environment interaction of the dopamine D4 receptor (DRD4) and observed maternal insensitivity predicting externalizing behavior in preschoolers. Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology, 48, 406–409. doi:10.1002/dev.20152
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. Development and Psychopathology, 23, 39–52. https://doi.org/10.1017/S0954579410000635.
Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2015). The hidden efficacy of interventions: Gene x environment experiments from a differential susceptibility perspective. Annual Review of Psychology, 66, 381–409. doi:10.1146/annurev-psych-010814-154047.
Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., Mesman, J., Alink, L. R., & Juffer, F. (2008). Effects of an attachment-based intervention on daily cortisol moderated by dopamine receptor D4: A randomized control trial on 1-to 3-year-olds screened for externalizing behavior. Development and Psychopathology, 20, 805–820. https://doi.org/10.1017/S0954579408000382.
Beach, S. R. H., Brody, G. H., Lei, M.-K., & Philibert, R. A. (2010). Differential susceptibility to parenting among African American youths: Testing the DRD4 hypothesis. Journal of Family Psychology, 24, 513–521. doi:10.1037/a0020835.
Belsky, J. (1999). Modern evolutionary theory and patterns of attachment. In Cassidy, J., Shaver, P. R. (Eds.), Handbook of attachment: Theory, research, and clinical applications. (pp. 141–161). New York, NY: Guilford Press.
Belsky, J. (1997a). Theory testing, effect-size evaluation, and differential susceptibility to rearing influence: The case of mothering and attachment. Child Development, 68, 598–600. doi:10.1111/1467-8624.1997.tb04221.x
Belsky, J. (1997b). Variation in susceptibility to rearing influences: An evolutionary argument. Psychological Inquiry, 8, 182–186. doi:10.1207/s15327958pli0803_3
Belsky, J. (2005). Differential susceptibility to rearing influences: An evolutionary hypothesis and some evidence. In B. Ellis & D. Bjorklund (Eds.), Origins of the social mind: Evolutionary psychology and child development (pp. 139–163). New York, NY: Guildford.
Belsky, J., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). For better and for worse: Differential susceptibility to environmental influences. Current Directions in Psychological Science, 16, 300–304. https://doi.org/10.1111/j.1467-8721.2007.00525.x.
Belsky, J., & Beaver, K. M. (2011). Cumulative-genetic plasticity, parenting and adolescent self-regulation. Journal of Child Psychology and Psychiatry, 52, 619–626. https://doi.org/10.1111/j.1467-7610.2010.02327.x.
Belsky, D. W., Moffitt, T. E., & Caspi, A. (2013a). Genetics in population health science: Strategies and opportunities. American Journal of Public Health, 103, S73–S83. https://doi.org/10.2105/AJPH.2012.301139.
Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. Psychological Bulletin, 135, 885–908. doi:10.1037/a0017376.
Belsky, J., & Pluess, M. (2013a). Beyond risk, resilience, and dysregulation: Phenotypic plasticity and human development. Development and Psychopathology, 25, 1243–1261. https://doi.org/10.1017/S095457941300059X.
Belsky, J., & Pluess, M. (2013b). Genetic moderation of early child-care effects on social functioning across childhood: A developmental analysis. Child Development, 84, 1209–1225. https://doi.org/10.1111/cdev.12058.
Belsky, J., Pluess, M., & Widaman, K. F. (2013). Confirmatory and competitive evaluation of alternative gene-environment interaction hypotheses. Journal of Child Psychology and Psychiatry, 54, 1135–1143. https://doi.org/10.1111/jcpp.12075.
Belsky, J., & van IJzendoorn, M. H. (2017). Genetic differential susceptibility to the effects of parenting. Current Opinion in Psychology, 15, 125–130. doi:10.1016/j.copsyc.2017.02.021.
Belsky, J., & Widaman, K. (2018). Editorial Perspective: Integrating exploratory and competitive–confirmatory approaches to testing person-x-environment interactions. Journal of Psychology and Psychiatry, 59, 296–298. doi:10.1111/jcpp.12824.
Booij, L., Van der Does, A. J., & Riedel, W. J. (2003). Monoamine depletion in psychiatric and healthy populations. Molecular Psychiatry, 8, 951–973. http://doi.org/10.1038/sj.mp.4001423.
Brodsky, G. H., Beach, S. R., Philibert, R. A., Chen, Y. F., & Murry, V. M. (2009). Prevention effects moderate the association of 5-HTTLPR and youth risk behavior initiation: Genes environment hypotheses tested via a randomized prevention design. Child Development, 80, 645–661. https://doi.org/10.1111/j.1467-8624.2009.01288.x.
parenting on hostility toward romantic partners. *Journal of Marriage and Family*, 75, 325–341. doi:10.1111/jomf.12101

Stein, M. B., Schork, N. J., & Gelernter, J. (2008). Gene-by-environment (serotonin transporter and childhood maltreatment) interaction for anxiety sensitivity, an intermediate phenotype for anxiety disorders. *Neuropsychopharmacology*, 33, 312. https://doi.org/10.1038/sj.npp.1301422.

Stocker, C. M., Masarik, A. S., Widaman, K. F., Reeb, B. T., Boardman, J. D., Smolen, A., … Conger, K. J. (2017). Parenting and adolescents’ psychological adjustment: Longitudinal moderation by adolescents’ genetic sensitivity. *Development and Psychopathology*, 29, 1289–1304. doi:10.1017/S0954579416001310

Sun, Y., Fang, J., Wan, Y., Hu, J., Xu, Y., & Tao, F. (2018). Polygenic differential susceptibility to cumulative stress exposure and childhood obesity. *International Journal of Obesity*, 42, 1177. https://doi.org/10.1038/s41366-018-0116-z.

Tamminga, C. A. (2002). Partial dopamine agonists in the treatment of psychosis. *Journal of Neural Transmission*, 109, 411–420. https://doi.org/10.1007/s007020200033.

Taylor, S. E., Way, B. M., Welch, W. T., Hilmert, C. J., Lehman, B. J., & Eisenberger, N. I. (2006). Early family environment, current adversity, the serotonin transporter promoter polymorphism, and depressive symptomatology. *Biological Psychiatry*, 60, 671–676. doi:10.1016/j.biopsych.2006.04.019

Thibodeau, E. L., Cicchetti, D., & Rogosch, F. A. (2015). Child maltreatment, impulsivity, and antisocial behavior in African American children: Moderation effects from a cumulative dopaminergic gene index. *Development and Psychopathology*, 27, 1621–1636. doi:10.1017/S095457941500099X

Turkheimer, E., Haley, A., Waldron, M., O’Donofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science*, 14, 623–628. https://doi.org/10.1046/j.0956-7976.2003.psci_1475.x.

Van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2006). DRD4 7-repeat polymorphism moderates the association between maternal unresolved loss or trauma and infant disorganization. *Attachment and Human Development*, 8, 291–307. https://doi.org/10.1080/14616730601048159.

Vrshek-Schallhorn, S., Stroud, C. B., Mineka, S., Zinbarg, R. E., Adam, E. K., Redei, E. E., … Craske, M. G. (2015). Additive genetic risk from five serotonin system polymorphisms interacts with interpersonal stress to predict depression. *Journal of Abnormal Psychology*, 124, 776. http://dx.doi.org/10.1037/abn0000098.

Widaman, K. F., Helm, J. L., Castro-schilo, L., Pluess, M., Stallings, M. C., & Belsky, J. (2012). Distinguishing ordinal and disordinal interactions. *Psychological Methods*, 17, 615–622. https://doi.org/10.1037/a0030003.

Zuckerman, M. (1999). *Vulnerability to psychopathology: A biosocial model*. Washington, DC: American Psychological Association.

### Appendix A

A multiple regression model depicting a G × E interaction can be written as

\[ Y_i = B_0 + B_1X_{1i} + B_2X_{2i} + B_3(X_{1i} \cdot X_{2i}) + \epsilon_i \]

where \( Y_i \) represents scores of person \( i \) on the outcome variable, \( B_0 \) is the intercept, \( B_1 \) represents marginal environment effects, \( B_2 \) is the marginal genetic effects, \( B_3 \) is the \( G \times E \) interaction effects above and beyond any additive combination of their marginal effects, \( \epsilon_i \) is stochastic error score. We can derive the estimated crossover point on \( X_1 \) (i.e., environment variable) by equating the outcome scores (\( Y \)) for two values of the genetic variable (i.e., \( X_{21}, X_{22} \)). This will lead to

\[ B_0 + B_1X_{1i} + B_2X_{21} + B_3(X_{1i} \cdot X_{21}) + \epsilon_i = B_0 + B_1X_{1i} + B_2X_{22} + B_3(X_{1i} \cdot X_{22}) + \epsilon_i \]

\[ B_2X_{21} + B_3(X_{1i} \cdot X_{21}) = B_2X_{22} + B_3(X_{1i} \cdot X_{22}) \]

\[ X_1 = \frac{B_2 (X_{22} - X_{21})}{B_3 (X_{22} - X_{21})} \]

\[ X_1 = - \frac{B_2}{B_3} = C \]

where \( C \) represents the crossover point (Aiken & West, 1991; Cohen et al., 2003, pp. 288–289).

As can be seen, the estimated value of crossover point is determined by the relative magnitude of the main effect of the moderator to the interaction (Aiken & West, 1991; Cohen et al., 2003, pp. 288–289). The estimated crossover point coupled with the range of environmental predictor, in turn, determines the location of the crossover point with regard to the environmental variable, thus distinguishing disordinal and ordinal forms of statistical interaction.