The Hidden Efficacy of Interventions: Gene × Environment Experiments from a Differential Susceptibility Perspective

Marian J. Bakermans-Kranenburg and Marinus H. van IJzendoorn

Center for Child and Family Studies, Leiden University, 2333 AK Leiden, The Netherlands; email: bakermans@fsw.leidenuniv.nl, vanijzen@fsw.leidenuniv.nl

Annu. Rev. Psychol. 2015. 66:381–409

First published online as a Review in Advance on August 11, 2014

The *Annual Review of Psychology* is online at psych.annualreviews.org

This article's doi: 10.1146/annurev-psych-010814-015407

Keywords

differential susceptibility, G × E interaction, diathesis stress, serotonin, dopamine, microtrial, randomized controlled trial

Abstract

The efficacy of interventions might be underestimated or even go undetected as a main effect when it is hidden in gene-by-environment (G × E) interactions. This review moves beyond the problems thwarting correlational G × E research to propose genetic differential susceptibility experiments. G × E experiments can test the bright side as well as the dark side of the moderating role of genotypes traditionally considered to represent vulnerability to negative conditions. The differential susceptibility model predicts that carriers of these risk genotypes profit most from interventions changing the environment for the better. The evolutionary background of G × E and differential susceptibility is discussed, and statistical methods for the analysis of differential susceptibility (versus diathesis stress) are reviewed. Then, based on results from 22 randomized G × E experiments, meta-analytic evidence for the differential susceptibility model is presented. Intervention effects are much stronger in the susceptible genotypes than in the nonsusceptible genotypes. The final sections suggest possibilities to broaden the G component in the G × E equation by including genetic pathways, and to broaden the E component by including methylation level and gene expression as promising ways to probe the concept of the environment more deeply and address the perennial issue of what works for whom.

Contents

INTRODUCTION

In the seminal Dunedin study, Caspi, Moffitt, and collaborators found a widely cited interaction between the serotonin transporter (*5-HTT*) gene and childhood maltreatment experiences elevating the risk for depression (Caspi et al. 2003). Individuals with one or two short alleles in the promoter region of the *5-HTT* gene showed more depressive symptoms, diagnosable depression, and suicidality when they had experienced stressful life events than did individuals homozygous for the long allele. The groundbreaking nature of this study is perhaps the best explanation for the mixed blessing of what happened afterward: The Dunedin results were for years at the center of both equivocal replication efforts and unrelenting criticism. This case study, which we continue below, illustrates critical issues in research design and statistical analysis that have plagued research on gene-by-environment (G × E) interactions.

In one of the critical reviews of the research based on the original Caspi et al. (2003) study, Munafò and colleagues (2009) found no evidence of a significant interaction effect, which they attributed partly to the absence of any main effect for the serotonin transporter linked polymorphic region (*5-HTTLPR*) polymorphism on depression. They argued that in the absence of a genetic main effect, the interaction effect size must also be zero, because if the environmental effect is increased for a specific genotype, there must be a main effect of the gene as well (Munafò et al. 2009, p. 216). This line of reasoning reflects the diathesis-stress or cumulative-risk model; the limited, one-sided nature of this model is discussed below.

Echoing the Munafò et al. (2009) perspective, Risch et al. (2009) also considered established genetic main effects a necessary condition for exploring G × E effects. A significant G × E effect in the absence of a main effect would be improbable, as this would require a reversal in the direction of the association between depression and life events, with the risk of depression increasing with the number of life events among those with the short (*ss* or *sl*) *5-HTT* genotype and decreasing with the number of life events among those with the long (*ll*) genotype (Risch et al. 2009, p. 2469). The effects of the two genotypes would then cancel each other out. This scenario is indeed one of the possibilities, but not the only one. Environmental effects can be absent for one genotype but present for the other genotype, with good outcomes under favorable conditions and bad outcomes under unfavorable conditions—"for better and for worse," as proposed by the differential susceptibility model. In that case, significant G × E effects are found in the absence of a genetic main effect (the two directions within one genotype cancel each other out). In a similar vein, we argue that the efficacy of interventions might go undetected as a main effect when it is hidden in G × E interactions.

As a prototypical example, the studies on the interaction between *5-HTT* and adversity highlight the promises of G × E but also point to a number of problems inherent in correlational G × E studies. These problems make their contribution to a thorough test of G × E necessary but not sufficient. In this article, we argue that G × E experiments solve several issues raised by critics of correlational G × E studies.

Weaknesses and Challenges of Correlational G × E Studies

The following issues are to some extent part and parcel of correlational G × E research and have been identified as major problems by many scholars before us (e.g., Duncan & Keller 2011, McClelland & Judd 1993, Munafò et al. 2009, Wachs & Plomin 1991). In essence, all of these issues concern threats to statistical power, which is not a trivial point in a field where precise measurement equals labor-intensive measures and thus implies limits to manageable sample sizes.

Skewed distribution of E. Quality of the environment may not be distributed normally. In virtually all studies testing the interaction between *5-HTTLPR* and adversity, the distributions of adversity are positively skewed, with the minority of participants having experienced significant stressful life events. Although good for the participants, such a distribution of the environment is unwelcome from a statistical point of view because it dramatically lowers the power to find G × E effects (McClelland & Judd 1993).

Skewed distribution of G. The distribution of polymorphisms may be skewed, and this also limits the statistical power. In order to obtain more equal groups, subgroups of genotypes are sometimes combined (e.g., for *5-HTTLPR*, the *ss* and *sl* carriers are sometimes combined and contrasted with the *ll* genotypes). But it is not always clear whether these combinations reflect

differences in gene functionality, and some genes are (in)famous for having been studied through multiple possible groupings of genotypes. An example is the dopamine receptor D4 (*DRD4*) gene, where the variable number of tandem repeats leads to more variants in the grouping of genotypes than the number of its variants. This may result in multiple testing, leading to potentially spurious findings (Munafò et al. 2014).

G and E are correlated. Genetic and environmental factors may not be independent, and when through passive or evocative gene-environment correlation (rGE) specific environmental conditions are more often present for certain genotypes, the supposed G × E effect is in fact a G × G effect. Studies examining the influence of parenting on child behavioral outcomes dependent on child genotype most often include only one child per family (e.g., Sulik et al. 2012), and between-family differences in parenting are considered to be directly reflected in differences between the children. However, parenting is confounded with shared genetic factors in parents as well as in children within a family. Designs with more than one child per family enable an independent estimate of the contributions of parenting and genetics (Van IJzendoorn & Bakermans-Kranenburg 2012). In a similar vein, unmeasured genotypes that increase the chance of exposure to a specific environmental factor may invalidate the conclusion. Confounders such as age or ethnicity are typically used as covariates to control for confounding effects, but taking into account their main effect is not sufficient to control for potential effects on the interaction; what is needed is control for the covariate × gene and covariate × environment effects (Keller 2014). This of course further lowers the power of the analyses.

Measurement errors. If the environment is poorly assessed, the G × E equation contains two components with divergent error variations (relatively small error variance in G and substantial error variance in E), which lowers the power and increases the risks for both type 1 and type 2 errors (Van IJzendoorn et al. 2011a, Wachs & Plomin 1991). Indeed, it has been argued that better measurement would be more crucial than larger samples to detect G × E (Rutter 2006, Wong et al. 2003), and in the debate about the replicability of the Caspi et al. (2003) results, this has been shown to be a key issue. Nonreplications more often relied on weak assessments of stressful life events and depression (Karg et al. 2011; Uher & McGuffin 2008, 2010). Evidence of genetic moderation was stronger for studies using objective measures or in-person interviews to assess stress than for studies using self-report questionnaires (Karg et al. 2011).

These issues point to serious shortcomings of correlational G × E studies. In a critical paper, Duncan & Keller (2011) argued that G × E studies suffered from publication bias, low statistical power, and a high false discovery rate. Large G × E effect sizes (defined as 1% explained phenotypic variance) would require sample sizes of at least 600 subjects to achieve a statistical power of 80%. Considering the modest modal sample size in psychological research, the authors suggest that most if not all G × E findings in the literature might be spurious and nonreplicable. Duncan & Keller (2011) argued that "the primary reason that power to detect interactions tends to be low is that the variance of the product term tends to be low in *nonexperimental* studies" (p. 1044, italics added).

As we argue below, genetically informed randomized controlled trials (RCTs) (G × experimental E or G × eE) address some of the problems inherent to research on the interplay between individuals' genetic make-up and their environment. Experimental manipulation of the environment results in more control of the E component in the G × E equation and greatly enhances the power of G × E analyses. But before we turn to G × eE, let us first delineate two different perspectives on G × E interaction effects and see how these can be distinguished and integrated.

DIATHESIS STRESS AND DIFFERENTIAL SUSCEPTIBILITY

Vulnerability and Susceptibility

For more than three decades, G × E research has been guided by the transactional/dual-risk (Sameroff 1983), cumulative-risk (Rutter 2010), or diathesis-stress (Gottesman & Shields 1967, Monroe & Simons 1991, Zuckerman 1999) models. The diathesis-stress model suggests that children with a vulnerable constitution (risky genes) and poor developmental experiences (e.g., insensitive parenting, low-quality child care, stressful life experiences) deviate from the developmental pathway of their peers. The Dunedin finding discussed above is a prime example of diathesis stress. Individuals carrying certain risk alleles (i.e., the *5-HTT s* allele) were found to be more likely to develop psychopathology when exposed to adversity.

Empirical studies show, however, that individuals vary not only in how much they are negatively affected by environmental stressors and adversity (the dark side) but also in the extent to which they are positively influenced by environmental resources and supports (the bright side). Moreover, the same characteristics that make individuals vulnerable to adversity also make them disproportionately likely to benefit from contextual support (Belsky et al. 2007). The differential susceptibility hypothesis proposes that in positive environments vulnerable children may outperform their peers who turn out to be less susceptible not only to bad environments but also to optimal environments (Bakermans-Kranenburg & Van IJzendoorn 2007, Belsky 1997a, Belsky et al. 2007, Ellis et al. 2011) (see **Figure 1**). The differential susceptibility model may

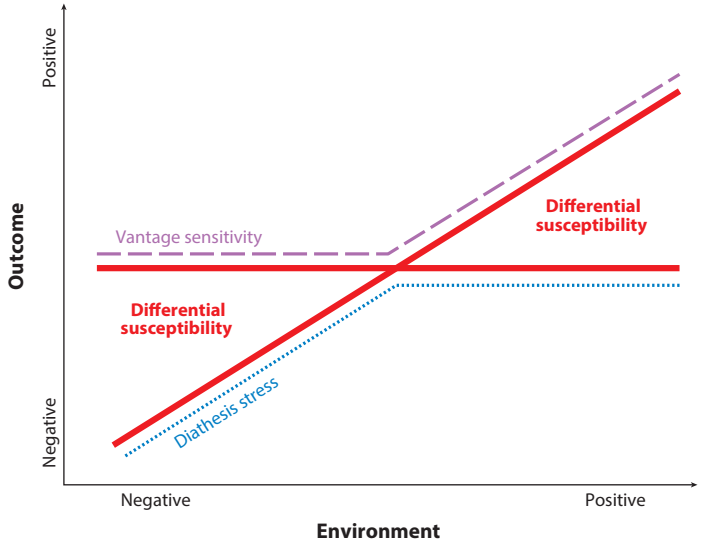

Figure 1

Models of differential susceptibility (*red lines*), diathesis stress, and vantage sensitivity. The differential susceptibility model hypothesizes that susceptible individuals are disproportionately influenced by both negative and positive environments (*diagonal line*), whereas nonsusceptible individuals are not influenced (strong version) or less influenced (weak version) by both negative and positive environments (*horizontal line*). The diathesis-stress or cumulative risk model (*blue dotted line*) contends that vulnerable and resilient individuals function similarly in a positive environment but diverge in negative environments, with vulnerable individuals showing worse outcomes. The vantage sensitivity model (*purple dashed line*) contends that individuals function similarly in a negative environment but diverge in positive environments, with some individuals showing better outcomes.

seem complementary to the diathesis-stress model, but at the deepest level it is fundamentally different, in part because its evolutionary foundation implies that certain genotypes characterizing a substantial percentage of the population must be called susceptibility genes instead of risk genes.

It is also essential to delineate differential susceptibility from a recently launched concept labeled vantage sensitivity. The term vantage sensitivity was meant to describe the notion that some individuals profit more than others from positive environmental factors such as warm parenting or high-quality daycare (Pluess & Belsky 2013); it is derived from differential susceptibility reasoning and relies on the term vantage as used by Manuck and colleagues (Manuck & McCaffery 2014, Sweitzer et al. 2012) to characterize the bright side of differential susceptibility (Bakermans-Kranenburg & Van IJzendoorn 2011). Note that vantage sensitivity is the mirror image of diathesis stress: It focuses on the effects of the positive side of the environmental continuum instead of the negative side. Vantage sensitivity lacks a firm theoretical and evolutionary background (as described below), and although it piggybacks on the differential susceptibility model, it seems to explain less than half as much. Moreover, the flip side of having three models—diathesis stress, differential susceptibility, and vantage sensitivity—is that almost any pattern of variation in response to environmental input can be explained by referring to one of the models. This runs the risk of rendering the set of models immune against empirical scrutiny, and when models are unfalsifiable, they must be considered unscientific (Popper 1979).

Biological Sensitivity and Differential Susceptibility

Two variants of differential susceptibility thinking emerged around the same time, biological sensitivity to context theory (Boyce et al. 1995, Boyce & Ellis 2005, Ellis et al. 2005) and differential susceptibility theory (Belsky 1997a, 2005; Belsky et al. 2007, 2009). Both theories were grounded in evolutionary theory.

Specifically, the first theory suggested that developmental variation in biological sensitivity to context had been maintained by natural selection because it produced different fitness outcomes in different environments encountered over evolutionary history. Biological sensitivity to context, as an endophenotypic property, is indexed by heightened reactivity in one or more of the stress response systems, and it functions to regulate openness or susceptibility to both harmful and supportive environmental influences. From a somewhat different perspective, Belsky (1997a, 2000, 2005) proposed that, as a form of bet hedging against an uncertain future, natural selection maintained genes for both conditional and alternative developmental strategies. Because the future will always be uncertain, parents never know which rearing strategies will prove most successful in terms of enhancing the child's reproductive fitness and thus the parent's inclusive fitness.

Even though biological sensitivity to context theory and differential susceptibility theory emerged independently and do differ in important respects, they share much in common, and in a remarkable step forward they were integrated under the umbrella of differential susceptibility theory (Ellis et al. 2011). The "for better and for worse" perspective led to the metaphor of orchids and dandelions, with orchids representing the highly susceptible individuals who wither away in stressful environments but flourish in nurturing environments, and dandelions representing individuals with the relative ability to function adequately in various circumstances (Boyce & Ellis 2005). The botanical correctness of this metaphor is not indisputable because some orchids do well in mountainous poor soil, and the adaptation of dandelions may be the result of epigenetic variation (see below). The shorthand also inadvertently suggests a categorical rather than a continuous characteristic. Nevertheless, the metaphor worked well with the nonscientific press, as evident from 287,000 results for a Google search for "orchids and dandelions theory" (see also Dobbs 2009, Rockoff 2013).

Markers of Susceptibility

What characteristics render individuals susceptible to environmental influences? Three broad constructs have been proposed and tested as markers of susceptibility: (*a*) reactive temperament, (*b*) biological sensitivity to stress, and (*c*) genetic make-up. These constitutional factors may point to an underlying factor affecting how individuals experience or approach their environment, but little research has addressed the associations among these susceptibility markers or tested whether they operate in an additive or interactive way.

Evidence for genetic moderation of environmental effects according to the differential susceptibility model has been specifically tested for serotonin- and dopamine-related gene polymorphisms. In the future other genotypes may be identified as markers of susceptibility [e.g., monoamine oxidase A (*MAOA*), brain-derived neurotropic factor (*BDNF*)], but so far serotonin- and dopamine-related gene polymorphisms have been most prominently in the spotlights of research on genetic differential susceptibility. Consequently, meta-analyses have focused on dopamine-related genotypes and *5-HTTLPR* as genetic susceptibility markers. For dopamine-related genotypes (*DRD2*, *DAT*, and *DRD4* polymorphisms; 15 studies, $N = 1{,}232$), the combined effect size for the association between adverse rearing influences and behavioral disturbance amounted to $r = 0.37$ for carriers of the risk alleles and only $r = 0.10$ for the comparisons without the risk alleles, so carriers of the risk alleles were indeed more vulnerable to environmental adversity. But the same was true on the bright side: The effect size for associations between support and better adaptation was $r = 0.31$ for carriers of the putatively risk alleles and $r = -0.03$ for those without the risk alleles (Bakermans-Kranenburg & Van IJzendoorn 2011). Thus, genotypes that in adverse contexts put children at risk for behavior problems made them also benefit more from support.

The differential susceptibility meta-analysis on *5-HTTLPR* included 77 studies ($N = 9{,}361$; Van IJzendoorn et al. 2012). In the total set, children with *s* alleles were more negatively affected than *ll* carriers by adverse contexts with regard to negative outcomes, but they did not benefit significantly more from positive environments. Ethnicity was a significant moderator, and in studies with mostly Caucasian participants (52 studies; $N = 6{,}626$), the combined effect size for negative outcomes in adverse environmental conditions was $r = 0.18$ for *ss/sl* carriers and $r = 0.04$ for *ll* carriers; *ss/sl* children also profited significantly more from positive environments ($r = 0.17$) than did *ll* children ($r = 0.05$). For children with the *ll* genotype, the associations between positive or negative environment and positive or negative developmental outcome were absent. Thus, *5-HTTLPR* was a marker of differential susceptibility in Caucasian samples, with carriers of short alleles being more sensitive to environmental influences than carriers of long alleles. There were not enough studies with other ethnicities for separate analyses on these groups. This is unfortunate because the findings demonstrate that ethnicity may make a critical difference in G × E interactions.

EVOLUTIONARY BACKGROUND OF G × E AND DIFFERENTIAL SUSCEPTIBILITY

From an evolutionary perspective, there are positive reasons for expecting gene-environment interactions. As Rutter (2006, p. 192) contended, "...there is the basic underlying evolutionary concept of natural selection which argues that genes are involved in the adaptation of organisms to their environment, that all organisms in a species will not respond to environmental change in the same way, and that this within-species variation in response involves individual differences in genetic endowment. In short, genetic variation in response to the environment is the raw material for natural selection."

The Relative Nature of Good and Bad

It seems implausible that genetic variation in response to the environment would only have to do with vulnerability to negative environments and the concomitant variance in nonoptimal outcomes, as under the diathesis-stress model. First, vulnerability genes (e.g., *5-HTT s* allele, *DRD4* 7-repeat allele) are found in relatively high frequencies in all human populations (Chang et al. 1996, Chiao & Blizinsky 2010, Gelernter et al. 1999), which suggests that they have been maintained by natural selection and must have benefits for fitness and reproduction. Were they only predictive of variance in susceptibility "for the worse," they would not have survived. In a similar vein, it is difficult to imagine how genotypes that only affect vantage susceptibility would be maintained over generations for only part of the population, without drift to a monoculture of such alleles that enhance sensitivity to positive conditions. Second, what is considered a good developmental outcome is prompted by the specific historical and cultural context, and behavior that we tend to regard as maladaptive or problematic may enhance fitness and reproduction in other circumstances. This points to the need to take into account the whole range of environmental contexts, the tails of which carry the connotations of "bad" and "good" that make no scientific sense from an evolutionary perspective.

As an example, the *5-HTTLPR* short allele has been related to greater responsivity of the HPA axis and cardiovascular system to aversive stimuli (Gotlib et al. 2008; Mueller et al. 2010; Way & Taylor 2010, 2011; see also Taylor et al. 2011), selective attention to negative stimuli (Pergamin-Hight et al. 2012), and increased amygdala reactivity in response to emotional stimuli (Caspi et al. 2010, Hariri et al. 2002). Such hypervigilance may be disadvantageous in stable, supportive contexts but is beneficial in unstable or life-threatening conditions (Belsky & Pluess 2013, Homberg & Lesch 2011).

In this regard it is telling that functional polymorphisms in *5-HTTLPR* have been found only for humans, rhesus macaques, and four other macaque species (Dobson & Brent 2013). The short allele would be responsible for the successful adaptation of humans and rhesus macaques to virtually all regions of the world (Suomi 2006). However, Dobson & Brent (2013) note that the other four macaque species live in only a few regions. The driving force may be variance in competition levels within a group over time. Hypervigilance to social threats is beneficial in times of elevated competition but unnecessarily costly in terms of time and energy when things are less threatening. Thus, when levels of intragroup competition are variable but balanced over the lifetime, carriers of the short alleles and carriers of the long alleles have similar levels of reproductive success, and the short allele is maintained over generations (Dobson & Brent 2013).

Variation in Susceptibility as an Evolutionary Outcome

One step further is the hypothesis that individuals from the same population more generally differ in their degree of susceptibility or plasticity and that natural selection favors this form of between-individual variation. Based on animal research, two types of susceptibility can be distinguished—developmental and activational susceptibility or plasticity—and both types may be relevant to differential susceptibility in humans. Developmental plasticity implies the emergence of different behavioral phenotypes in different environments as a result of different developmental trajectories evoked by those environments. This includes changes in the nervous system or in physiology as a result of experience, with long-term consequences resulting from early environmental influences. Activational plasticity refers to adaptation to the immediate context, such that an individual expresses different behaviors in different conditions (Snell-Rood 2013). Activational plasticity requires a more refined nervous system, with accompanying higher costs; moreover, it requires

time and energy to monitor the environment carefully—whether consciously or unconsciously (DeWitt et al. 1998).

Why would such costly phenotypes be maintained? Two evolutionary explanations for differential susceptibility in humans have been suggested (Dingemanse & Wolf 2013, Wolf et al. 2008). First, the benefits of susceptibility are often negatively frequency dependent, that is, when there is competition for resources, the benefits increase with a decreasing number of susceptible individuals in the population. The explanation is as follows: When a majority of nonsusceptible individuals use strategy *A* but it becomes advantageous to use strategy *B*, responsive or susceptible individuals choose strategy *B* while unresponsive individuals stick to strategy *A*. However, when too many individuals use strategy *B* and the resources are limited, the benefits of choosing *B* diminish. Thus, the benefits of responsiveness decrease with the frequency of responsive individuals (Wolf et al. 2008). This explains why both susceptible and nonsusceptible genotypes are maintained within a population.

Second, even small between-individual differences in social behavior lead to the coexistence of susceptible and less susceptible individuals. In Hawk-Dove-like contest situations, predictable differences in aggressiveness between individuals select for susceptible (activationally plastic) individuals who flexibly switch between aggressive and nonaggressive strategies. They are nonaggressive when confronted with opponents who are aggressive and vice versa. In turn, the presence of flexible individuals favors a rigid strategy in other individuals. When confronted with a flexible individual in the Hawk-Dove-like contest situation, rigid adoption of the aggressive strategy triggers nonaggressive behavior in the flexible individuals (Dall et al. 2004, Wolf et al. 2011). Individual differences in social behavior thus elicit flexible responses, and flexible responders offer opportunities to nonflexible individuals. As a result, natural selection results in a mixture of susceptible (adapters) and less susceptible (rigid) types (Dingemanse & Wolf 2013).

Natural selection acts on phenotypes: The relative fitness of a genotype depends on the survival and reproductive success of the associated phenotypes in specific environments, pointing to the importance of G × E effects. Manuck (2010) noted that selection only recognizes reproductive success and that the evolutionary underpinnings of the differential susceptibility model hold only if the behavioral phenotypes resulting from the G × E interaction also predict reproductive outcomes. If not, he argues, the variance in behavioral outcomes in crossover interaction would more likely be a by-product of other regulatory processes. We tend to disagree with this point. Given the myriad of behaviors that are directly or indirectly relevant to reproductive success or inclusive fitness (Hamilton 1964a,b), including aggression, social competence, and depression, it is difficult to think of a behavioral phenotype that would be totally unrelated to reproductive success.

Variation in Susceptibility Within the Family

In Belsky's evolutionary reasoning, the prediction of variation in susceptibility is pushed to the level of the family (Belsky 1997a,b, 2000, 2005). He argues that parents could never know for certain what rearing strategies would prove most successful in terms of promoting a child's reproductive success. Therefore, children should vary in their susceptibility to the rearing environment, especially within families. The inclusive fitness of parents would be optimal if some of their children were affected by their parenting efforts and were well prepared for a future correctly envisioned by their parents, and if others were not susceptible to their parents' routines and pursued alternative strategies (e.g., Sulloway 1996).

Moreover, not only would parents increase their reproductive fitness through variation in their offspring's susceptibility to rearing influences, but the same would also be true of siblings in these families. After all, just like parents and children, siblings share 50% of their genetic layout. Thus, variation in susceptibility enhances siblings' and parents' inclusive fitness to the same extent (Belsky 2005, Ellis et al. 2011). Note that this line of reasoning assumes susceptibility of the type denoted as developmental plasticity: Once adaptation has taken place, the individual follows his or her developmental trajectory, which may promote or decrease reproductive success. The prediction of variation in susceptibility within families is original and bold; unfortunately, the possibilities of testing this hypothesis in modern times are elusive, at least in humans. Animal research may take up this challenge.

STATISTICAL APPROACHES

In the first years of documenting empirical evidence for differential susceptibility, the criteria for demonstrating differential susceptibility and distinguishing it from support for other models of development were outlined rather imprecisely. The criteria included a formal test for a crossover (or disordinal) interaction and visual inspection of the figure representing the regression lines for the two groups distinguished by the moderator variable, e.g., high or low on temperamental reactivity, or with or without the *DRD4* 7-repeat allele. Support for the differential susceptibility model would be indicated by similarity of the figure to a prototypical plot displaying two regression lines crossing at some point in the middle of the distribution of the predictor variable (as in **Figure 1**), with one regression line (for the nonsusceptible group) showing a slope close to zero and the other regression line (for the susceptible group) showing a slope clearly different from zero (Belsky et al. 2007).

Evaluation of Interaction Shape

Over the years, a number of scholars have developed statistical approaches to closely examine the precise shape of G × E interactions and to decide more formally whether interactions fit diathesis-stress or differential susceptibility models. The central aim of these approaches is to distinguish ordinal or removable interactions (indicating diathesis stress) from disordinal or crossover interactions (indicating differential susceptibility).

Kochanska and colleagues (2011) suggested an approach using the analysis of regions of significance (Aiken & West 1991, Preacher et al. 2006; see **http://quantpsy.org/interact/mlr2.htm**); that is, the specific values of the predictor below which and above which the regression lines for the two groups (different on the moderator) differ significantly in terms of the outcome. When these values both fall within the range of 2 SD below and above the mean of the predictor (representing 95% of the sample), the data would support the differential susceptibility model; indeed, in these cases the "for better and for worse" effects are found within the range of empirically observed values of the predictor.

Roisman and colleagues (2012) suggested two additional metrics for quantifying evidence for differential susceptibility (see **http://www.yourpersonality.net/interaction/**). The first is the ratio of the area between the regression lines on the right-hand side of the interaction crossover point and the total area between the regression lines (left and right from the crossover point), with the −2SD and +2SD lines as boundaries. In the prototypical case of differential susceptibility this ratio, referred to as the proportion of interaction, amounts to 0.50; in the case of diathesis stress it is close to zero. The other metric concerns the proportion of the sample that is differentially affected by the moderator; that is, the proportion of the sample with predictor values higher than the crossover point of the regression lines. The proportion affected index approaches 0.50 in cases

of differential susceptibility and 0.00 or 1.00 in cases of diathesis stress, depending on the direction of the predictor.

Model Fitting

A further method for distinguishing ordinal from disordinal interactions, and thus distinguishing between diathesis stress and differential susceptibility models, was developed by Widaman et al. (2012). A crucial distinction between the two types of interaction is the location of the crossover point, which in the case of an ordinal interaction is at the boundary or outside the range of observed predictor values. In the Widaman et al. (2012) approach, the predictor is centered at the crossover point, and a confidence interval for the crossover point is estimated. When both the crossover point and its confidence interval fall within the range of observed predictor values, the interaction represents differential susceptibility; when both fall outside the range of observed predictor values, the interaction represents diathesis stress. Moreover, the fit of the regression equation to models representing strong and weak versions of diathesis stress and differential susceptibility can be tested (Belsky et al. 2013). The strong version of differential susceptibility implies that those who are not susceptible are not at all affected by the environmental predictor; the weak version implies that some are less affected than others. In a similar vein, strong and weak versions of diathesis stress pertain to the extent that nonvulnerable individuals are affected by the predictor to a lesser extent than vulnerable individuals (weak version) or not at all (strong version).

The advantage of this approach is the formal testing of nested diathesis-stress and differential susceptibility models. However, the strong differential susceptibility model and the weak diathesis-stress model have the same number of parameter estimates, and as a result these two cannot be tested against one another. Belsky et al. (2013) suggest that in this case the model with the highest proportion of variance explained should be preferred, but under these circumstances, the benefits of formal testing are lost. Importantly, the model-testing approach is more liberal than the standard approach for testing interaction effects, so in some cases it can be concluded that the data support differential susceptibility even when the interaction term (susceptibility factor and predictor, G × E) is not statistically significant. The standard approach of first testing the statistical significance of G × E is conservative, and perhaps too conservative, when the aim is to test a hypothesis specified a priori (Belsky et al. 2013).

Restriction of range. Both in the Widaman et al. (2012) and in the Roisman et al. (2012) approach, support for one of the models is dependent on the range of observed predictors. Thus, when the crossover point and its confidence interval fall outside the range of observed predictor values, support for the diathesis-stress (or vantage) model is specific to the pertinent sample. If the sample is not representative of a population with more extreme positive or negative environments, the differential susceptibility model is falsely rejected. Furthermore, limitations of correlational G × E studies such as correlated G and E, substantial measurement error in E, and lack of statistical power also limit the validity of the conclusions based on the statistical methods to test differential susceptibility, whatever their complexity.

G × E EXPERIMENTS: POWERFUL TESTING OF G × E

In order to overcome the limitations of correlational and longitudinal research on differential susceptibility, this emerging field needs experimental designs to test differential susceptibility (Van IJzendoorn et al. 2011a). Differential susceptibility experiments have at least three distinct advantages compared to correlational studies.

1. G and E are uncorrelated. In RCTs, the environment is manipulated in standard ways, and randomization breaks the potential rGE. For example, in experimental interventions that elevate the level of parental sensitivity, changes in child outcomes are causally related to the manipulated environment (Bakermans-Kranenburg et al. 2003). Correlations between genotype and environment cannot play a contaminating role because genes and other constitutional factors (e.g., temperament, stress reactivity) may only moderate the effectiveness of the intervention, that is, constitute markers for differential susceptibility. When random assignment to the intervention and control groups or—in the case of within-subject designs—randomized order of the manipulations (the environmental change) is stratified according to the differential susceptibility marker, the independence of marker and outcome is also guaranteed.

2. Measurement error is reduced. Differential susceptibility experiments avoid the issue of unequal measurement errors in the differential susceptibility equation, that is, the varying error components in the interaction equation of the individual constitution and the environment (in genetic terms, G × E). If genotyping is done in a careful way but the environment is assessed poorly, the G × E equation contains two components with highly divergent error components (smaller for G than for E), creating high risks for type 1 and type 2 errors. This may be the most important reason for Plomin's paradox (Wachs & Plomin 1991) that gene-environment interactions may be omnipresent (as the raw material for evolutionary variation and selection) (Rutter 2006) but appear difficult to find and to replicate. Replication of G × E findings is critically dependent on accurate assessments of both the genotype and the environment (McGuffin et al. 2011). Focused and standardized experimental manipulation of the environment (with certified fidelity of implementation) circumvents the problem of large measurement errors in the assessment of environments, in particular when the intervention has been proven to be efficacious.

3. Power is enhanced. Randomized controlled G × E experiments require considerably fewer subjects to obtain the same statistical power as correlational G × E studies. Experimental studies create more variance in the product term because interventions stimulate experimental participants to become maximally different from controls. As a result, the power of experimental G × E studies is much larger than that of correlational studies. In a series of simulations with two factors (e.g., treatment and genotype) on a sample size of 100 subjects, McClelland & Judd (1993) demonstrate that the power of a common correlational study with truncated distributions at the extremes and many observations toward the center of the distributions (e.g., due to selective nonresponse) decreases to 6% of that of a proper two-factorial experiment. Compared to a somewhat less optimal G × E design with skewed distribution of genotypes, the power of a correlational G × E study would still be 13 times smaller. In this scenario 1,300 subjects are needed to achieve the same power in a correlational study as in an experiment with 100 subjects—and this does not depend on the effect size of the moderator (McClelland & Judd 1993).

Genetic Differential Susceptibility Experiments

Genetic differential susceptibility experiments constitute a subclass of experimental G × E studies, testing the bright side of the moderating role of genotypes shown to be related to vulnerability to negative conditions. Under the diathesis-stress model, these vulnerable genotypes are associated with poor outcomes when exposed to stress or unfavorable environments. The differential susceptibility model predicts that carriers of these very same genotypes profit most from experimental manipulation of the environment for the better. In this section we present the most important G × E experiments, namely RCTs that fulfill the requirement of randomized assignment of participants to conditions, in this case to variations in the environment. Thus, genetic variation can be a fixed factor, but randomized E is a necessary condition.

It has been argued that "G × E research is inherently nonexperimental" (Keller 2014, p. 22) because the genetic variable cannot be experimentally changed. This argumentation is incorrect. In cancer research, designs are common in which patients with different types of tumors are randomly assigned to, for example, an established drug treatment (care as usual) and a new drug regime. In this primordial RCT example the environment (the treatment) is experimentally manipulated, but the type of cancer (the organismic characteristic) is not manipulated. In fact, Thomas (2010a) argues that clinical trials provide unique opportunities to study G × E interactions. Notably, the Widaman et al. (2012) approach to distinguish between differential susceptibility and diathesis stress models can also be applied to randomized controlled G × E experiments (see Plak et al. 2015, Van IJzendoorn & Bakermans-Kranenburg 2015).

A G × E experiment should not be confused with the use of laboratory test settings to assess behavioral or developmental outcomes and to contrast the outcomes between genotypes (as in, e.g., Bakermans-Kranenburg & Van IJzendoorn 2011, Gallardo-Pujol et al. 2013, Israel et al. 2009, Verschoor & Markus 2012). An experiment is defined by the manipulation of a predictor rather than by the assessment method. Studies examining genetic moderation of the response to psychopharmacological treatment (e.g., Cheon et al. 2007, Markus et al. 2012) are not included here to avoid an overly heterogeneous domain of review.

Nanotrials, Microtrials, and Field Trials

The environment can be manipulated, and the outcome can be assessed, at various levels. At the level of the nanotrial, pertinent studies examine the immediate neural or behavioral responses to a small range of positive and negative stimuli, to minor manipulations of stress levels, or to subtle priming. Such manipulations cannot be considered stand-alone components of larger field trials; rather, nanotrials are meant to elucidate mechanisms of change underlying differential susceptibility that may include differences in attention, state regulation, orienting responses, or thresholds for punishment and reward. Microtrials use a focused manipulation of a somewhat broader component of the environment that could easily be included in a field trial as one of its effective ingredients, e.g., computerized early literacy instruction with and without personalized feedback, or high and low levels of structure in a parenting situation. Finally, field trials test the variance in response to interventions with high ecological value, such as parent training or broad educational or social programs. We next present the available G × E experiments categorized at the nano-, micro-, and field-trial levels.

Differential susceptibility has been defined as a process leading to more-or-less enduring developmental changes, not to short-term, fleeting effects (Ellis et al. 2011). Such long-term processes are usually not addressed in nano- or microtrials. Nevertheless, such trials are important for elucidating intermediate steps in the cascade of changes in field trials, for example, the neural or hormonal changes that accompany developmental changes.

Field Trials

Externalizing behavior problems. An early G × E experiment supporting differential susceptibility theory was an RCT ($N = 157$) with video-feedback training in families with a toddler at risk for externalizing behavior problems (Video-feedback Intervention to promote Positive Parenting and Sensitive Discipline; VIPP-SD) (Bakermans-Kranenburg et al. 2008b). The intervention families participated in six home visits with video feedback. The intervention was effective in enhancing parental sensitive discipline and reducing child externalizing symptoms, but the latter only when the children were carriers of the *DRD4* 7-repeat allele. In fact, when parents showed

a more-than-average increase in the use of positive discipline, the decline in externalizing behavior was strongest in children with the *DRD4* 7-repeat, an illustration of a dose-response relation that confirms the causal association between experimentally induced change of environment and child behavior problems. Moreover, the impact of the intervention went "under the skin" of the children as evidenced by the lower levels of daily cortisol production in the susceptible group (Bakermans-Kranenburg et al. 2008a).

In an RCT with elementary school–age children ($N = 50$), Van den Hoofdakker and colleagues (2012) explored whether a dopamine transporter gene (*SCL6A3/DAT1)* moderated the effectiveness on ADHD symptoms and behavior problems of a behavioral parent training program in addition to routine clinical care (in the control group) that included family counseling, psychoeducation, and advice. The parent training consisted of 12 two-hour group sessions across five months; during the sessions parents were instructed in behavior management techniques, with an emphasis on the frequent praise of and prompt reward for children's prosocial behavior. Children carrying one or no *DAT1* 10-repeat allele profited strongly from the training-related positive change in their caregiving environment, whereas children with two *DAT1* 10-repeat alleles remained unaffected (Van den Hoofdakker et al. 2012).

The Fast Track Randomized Control Trial was a 10-year-long intervention to prevent high-risk kindergarteners from developing persistent externalizing psychopathology. The intervention was broad, including parent training groups and home visits to promote the development of positive family-school relationships and parental management skills, self-efficacy, and life management; child social skill training groups and tutoring in reading during the early years were also provided. During adolescence, curriculum-based parent and youth group meetings were offered. The outcome was externalizing psychopathology at age 25 years (Albert et al. 2015). Genetic moderation of response to Fast Track was examined for 10 single-nucleotide polymorphisms (SNPs) of the glucocorticoid receptor gene *NR3C1*. Among European-American children ($N = 242$), the intervention was more effective among carriers of the rs10482672 *A* allele, after correcting for multiple testing. Among carriers of the *A* allele, 18% of the intervention participants had at least one symptom of externalizing psychopathology at age 25, compared to 75% of the control participants. In contrast, for participants without the *A* allele, the percentages were 56% for intervention children and 57% for control children.

Alcohol use. The Strong African American Families (SAAF) program focused on African American adolescents and their (often single-parent) families. The SAAF program consisted of seven weekly meetings. Parents were taught the consistent use of nurturant-involved parenting practices along with high levels of monitoring and control, with clear expectations about alcohol use. Children learned about the consequences of alcohol use and were taught resistance strategies, and families engaged in activities designed to increase family cohesion (Brody et al. 2006). The control families received leaflets on various aspects of development in early adolescence. Past-month substance use was assessed when subjects were 11 (pretest), 12 (posttest), 13 (follow-up), and 14 (long-term follow-up) years old ($N = 337$). The SAAF intervention had a positive effect on parenting practices. *DRD4*—but not *5-HTTLPR*—of the children moderated the effect on children's substance use. SAAF appeared to be more effective for teenagers with a *DRD4* 7-repeat allele than for teenagers with two *DRD4* 4-repeat alleles (Beach et al. 2010).

The Strong African American Families-Teen (SAAF-T) program was implemented in 502 families with 16-year-old adolescents who were followed for 22 months (Brody et al. 2014). SAAF-T consisted of five meetings, again with separate caregiver and adolescent training followed by joint caregiver-adolescent sessions during which families practiced the skills they learned in the separate sessions. The control group received a family-centered intervention designed to promote

healthy behaviors among adolescents. In the control group, male carriers of the *DRD4* 7-repeat allele were at risk for increased substance use (cigarettes, alcohol, marijuana), whereas after SAAF-T they showed significantly decreased substance use, at the same level as that of carriers of shorter alleles in both the control and intervention groups. This moderating role of *DRD4* was not found in female adolescents (Brody et al. 2014).

The SAAF and SAAF-T data were combined in an analysis of a larger number of genotypes related to dopaminergic (*DRD2*, *DRD4*, *ANKK1*) and GABAergic (*GABRG1*, *GABRA2*) systems (Brody et al. 2013). More than 900 youths were included in the combined study. Control subjects carrying a risk variant of the *GABRG1*, *GABRA2*, and *DRD2* genes showed larger increases in alcohol use than did their counterparts in the experimental condition or youth with the nonrisk variants in both the control and experimental groups. Risk genes were also combined into a cumulative or multilocus risk index, but this left 21% of the sample without a genetic risk profile, which makes tests inconclusive.

The PROSPER (Promoting School-Community-University Partnerships to Enhance Resilience) preventive intervention trial was another large-scale G × E intervention. PROSPER aimed at reducing alcohol abuse in sixth- to ninth-grade youngsters through family-focused and school-based interventions with community-level random assignment. Participants (545 adolescents, mostly non-Hispanic white) were genotyped for *DRD4*. Only part of the sample participated in the family intervention; thus, primarily the effectiveness of the school-based intervention was tested. The focus was on increasing the awareness of consequences of substance use, resisting pressures from others, establishing future aspirations, and promoting positive interactions with parents. The program consisted of 11 to 15 sessions with interactive teaching methods and small-group activities (Spoth et al. 2007). At the ninth-grade assessment, the adolescents reported their initiation of alcohol use in a three-item questionnaire (Cleveland et al. 2015). No main effect of intervention on alcohol use was found, and a significant two-way G × E interaction was also absent. A significant three-way interaction effect with parental involvement was found. The intervention decreased alcohol use in carriers of the *DRD4* 7-repeat allele, but only when these adolescents had reported a moderate to high involvement of the primary caregiver in their lives at baseline. Participants without the 7-repeat allele seemed not to profit from the intervention.

The G × E experiments on substance use are exemplary in the use of large samples and long-term assessment of outcomes. However, the assessment of alcohol or substance use was restricted to a few self-report questions that might be liable to response biases, in particular if "cool" but illegal practices are to be reported, and response biases may differ in size and direction in the intervention and control groups. Interpretations of disparities between SAAF and PROSPER should take differences in ethnicity into account. Not only frequency but also functionality of genotypes might differ between ethnicities (see Van IJzendoorn et al. 2012, Williams et al. 2003; but see Vijayendran et al. 2012).

Internalizing problems. Following an initial quasi-experimental study on differential effectiveness of cognitive behavioral therapy (CBT) in children with anxiety disorders (Eley et al. 2012), Bockting and colleagues (2013) tested *5-HTTLPR* as a moderator of the response to CBT in recurrently depressed adults ($N = 180$ Caucasian patients). It was expected that the *ss* genotype would be associated with a better response to CBT in preventing recurrence as compared with the *sl/ll* genotypes. The control group received care as usual, and the outcome was time to recurrence assessed prospectively over 5.5 years using the Structured Clinical Interview for DSM-IV. *5-HTTLPR* did not predict CBT response; the effectiveness of the treatment was similar in both genotype groups.

In a similar vein, the Living Well With Stroke study tested the effectiveness of a brief psychosocial treatment in reducing depressive symptoms after stroke for participants with the various *5-HTTLPR* genotypes (Kohen et al. 2011). Clinically depressed patients with ischemic stroke ($N = 61$) were randomly assigned to a 9-session brief pleasant events, problem-solving intervention plus antidepressant (intervention), or to usual care plus antidepressants (control). The Hamilton Rating Scale for Depression was used at baseline and after treatment. The *s* allele was associated with better treatment outcome: Among patients with the *ss* genotype, behavioral treatment had a large effect, but among *ll* carriers there was no evidence of an intervention effect. The effect for *sl* carriers was intermediate.

Child abuse and neglect. Parenting interventions in maltreating families or families at high risk for neglect or abuse are critically important to prevent (further) child maltreatment. Cicchetti et al. (2011) conducted an RCT in 106 families (majority of black ethnicity) with a 1-year-old infant, using two different intervention modalities, child-parent psychotherapy and psychoeducational parenting intervention. Individual parenting interventions were implemented in weekly sessions for one year, constituting major opportunities for change. Attachment security and disorganization measured at baseline and follow-up (26 months of age) were the most important outcomes. Substantial increases in attachment security and decreases in attachment disorganization were observed in maltreated children whose families were involved in one of the intervention modalities. However, genetic moderation (*5-HTTLPR*, *DRD4*) was absent.

One might speculate that a long-term, intensive, corrective therapeutic intervention dwindles the power of genotype in moderating its impact on children's attachment development. The study may therefore not be considered a disconfirmation of the differential susceptibility theory but rather an indicator of its boundaries. Analogous to behavior genetics where environmental effects are stronger in more heterogeneous environments, one might submit that with drastic changes in previously extremely depriving environments, the influence of genetics and of G × E may be negligible.

The Bucharest Early Intervention Project (Brett et al. 2015, Nelson et al. 2014), however, seems to refute this idea. This intervention was a unique field experiment with random assignment of 136 abandoned children suffering from neglect in orphanages in Bucharest, Romania, who were 6 to 30 months of age at baseline. Sixty-eight children were randomly assigned to institutional care as usual and 68 to a newly created foster care arrangement, with regular professional support of the foster parents. Ethnicity of the children was Romanian and Roma (allelic frequencies were not different, and ethnicity did not influence the results, but the power to find such differences was limited). The interaction of *5-HTTLPR* genotype with intervention group status predicted change in externalizing behavior from baseline to 42 and 54 months. In the group remaining in institutional care, children with the *ss* genotype showed the most externalizing behaviors, whereas in foster care, the *ss* genotype was associated with the lowest externalizing scores. In contrast, for *sl/ll* carriers, intervention group status did not predict externalizing behavior.

Microtrials

Aggression. Several G × E experiments have tested the moderating role of the *MAOA* gene promoter polymorphism in the influence of a negative (maltreating) environment on the development of antisocial and aggressive behaviors (Caspi et al. 2002). In one of the first G × E experiments on *MAOA* and aggression, McDermott and colleagues (2009) provoked 70 male participants to forced administration of an unpleasant tasting hot sauce to a confederate as retaliation for being bereft of a larger (80%) or smaller (20%) amount of money earned in completing a vocabulary

task. Both carriers of the low-activity variant of *MAOA* and their peers with the other genotype administered more hot sauce in the 80% condition than in the milder 20% condition, but the difference was somewhat larger in carriers of the low-activity variant of *MAOA*.

Cyberball is another way to trigger (reactive) aggression in a laboratory setting (Williams & Jarvis 2006). In Cyberball, participants play ball with two imaginary individuals over the Internet, but after a promising start with an equal distribution of tosses, the excluded participant does not receive any more tosses from the other participants, which results in rather severe negative mood and feelings of rejection. The computer game was used in a microtrial on 57 Spanish male students to mimic accepting (i.e., inclusion) and rejecting (i.e., exclusion) social environments (Gallardo-Pujol et al. 2013). The Point Subtraction Aggression Paradigm was used to measure aggressive responses (stealing points from an opponent). Social exclusion was related to more aggressive acts, as was the low-activity variant of *MAOA*. But in the ostracized group, carriers of the low-activity *MAOA* allele showed significantly higher aggression than their peers with the high-activity polymorphism. In the social inclusion group (not excluded during Cyberball), lower levels of aggression and no difference in aggression between the two *MAOA* polymorphisms were found (Gallardo-Pujol et al. 2013), indicating more positive effects of the inclusion for carriers of the low-activity *MAOA* allele.

Although *5-HTTLPR* is one of the usual genes investigated in studies on the emergence of depression and anxiety (see below), Verona et al. (2006) examined this genotype in a G × E experiment on aggression. In a study on 111 students (mixed ethnicities), the authors examined the moderating role of *5-HTTLPR* in the association between experimentally induced stress and resulting aggression. Participants were randomly assigned to a nonstress condition or a physical stress condition consisting of a small harness placed around the chest from which air blasts were directed at the throat. Aggression was assessed through the intensity and duration of electrical shocks delivered by the participants to a confederate who made mistakes in a series of digit span tasks. In the stress condition, male *ss* carriers administered higher shock levels than did male *sl* or *ll* carriers. In the no-stress condition, low levels of aggression were displayed by all genotypes. In females, genotype did not make a difference (Verona et al. 2006).

Internalizing problems. Andersson and colleagues (2013) examined the moderating effect of both *COMT*val/met and *5-HTTLPR* on the effectiveness of a nine-week Internet-delivered CBT treatment for social anxiety disorder (without personal contact with the therapist); the waitlist control group received delayed treatment after nine weeks. Neither genotype influenced CBT outcome, nor was any G × G × E effect found. The treatment was equally effective in all four genotype groups (total $N = 202$).

Cognitive development. A prime example of a cognitive microtrial is the early literacy instruction with the intelligent tutoring system Living Letters, a series of 40 brief computer games aimed at phonemic awareness (Kegel et al. 2011). In one of the intervention modes, a talking tutor was built into the games to reward any move in the right direction of a solution. The same games without a tutor constituted a second intervention mode, and the control group played hide-and-seek computer games. Children were randomly assigned to the three groups. Information on *DRD4* genotype was available for 182 participants. At posttest, the children with *DRD4* 7-repeat alleles performed worst when they were in the control group (the difference was not significant but was in the expected direction), whereas they outperformed all other groups when they had received the personalized feedback from a one-to-one sensitive computerized tutor—unavailable in regular schools. The program was ineffective in children without the *DRD4* 7-repeat allele. Only children

with the *DRD4* 7-repeat allele were responsive to the quality of instruction, and thus they might be called the "orchids" in the classroom.

In a replication and extension of this study, 257 children with delayed literacy skills were randomly assigned to one of three conditions: Living Letters, Living Books, or a control condition with computer games not focused on early literacy skills. Similar to Living Letters (which focuses on alphabetic knowledge), Living Books (which emphasizes text comprehension) included a computerized tutor that coached the learning process by providing personalized feedback. Children read digital storybooks and answered questions about story events and difficult words in the text. Living Books was effective in children with the *DRD4* 7-repeat allele but ineffective in children without the *DRD4* 7-repeat allele. Living Letters was effective in both groups, but more so in children with the *DRD4* 7-repeat allele (Plak et al. 2015).

Computerized training programs have also been used to enhance children's fluid intelligence and working memory capacity. Söderqvist et al. (2012) developed three programs consisting of 25 home sessions that were 15 minutes in length, with the level of difficulty automatically adjusted according to performance. They included 96 4-year-olds in a G × E experiment that focused on 11 SNPs related to the dopamine system. One program consisted of nonverbal reasoning exercises to enhance fluid intelligence, another program focused on working memory, and a third intervention was a combination of these two programs. Participants were randomly assigned to these three programs or a placebo training, stratified for gender. As expected, training enhanced cognitive development, but after correction for multiple comparisons, no significant genetic moderation was demonstrated.

Nanotrials

Priming can be used to enhance individuals' attention to or feelings about specific features of the environment and to test whether a subtle change in the perception of the environment affects participants' behavior, for example, their prosocial behavior. Sasaki and colleagues (2013) used priming for religion (intervention condition) or neutral priming (control condition) with 178 college students (mixed Caucasian and Asian ethnicity). The outcome was the participants' motivation to volunteer for organizations with green aims such as energy efficiency on campus. The *DRD4* 7-repeat and *DRD4* 2-repeat alleles were considered the susceptibility variants in this mixed-ethnicity group. Religion priming enhanced prosociality, moderated by genotype. Only carriers of the *DRD4* 7-repeat or *DRD4* 2-repeat alleles were susceptible to the influence of priming; for their peers with other genotypes, the priming condition did not affect their willingness to volunteer.

Attention bias modification (MacLeod et al. 2002) has become increasingly popular as a very brief but possibly effective therapy for anxiety disorders (Hakamata et al. 2010). The basic idea is to change attentional biases to threat-related words or pictures by forcing participants in dot-probe or Stroop-like computer tasks to pay more attention to positive or neutral instead of negative stimuli. In their study on 116 healthy adults, Fox and her colleagues (2011) used a standard attention bias modification procedure with positive and negative pictures matched for arousal. The participants had to choose whether two dots, presented immediately after the display of a positive and a negative picture, were in a horizontal or vertical position. The dots appeared at the spot of the negative picture (training negative bias) or the positive picture (training positive bias). Participants with a low-expression variant of *5-HTTLPR* changed their attention more than did carriers of the high-expression variant not only in the negative condition but also in the positive condition, showing the "for better and for worse" pattern that is characteristic of activational differential susceptibility. They were more vulnerable to a negative bias induction, but they did

also profit more from attention bias modification in a positive direction; they are thus potentially more open to therapeutic efforts related to threat biases (Fox et al. 2011).

Meta-Analysis of Randomized G × E Experiments

The narrative review shows that some G × E experiments seem to support genetic moderation of manipulation of the environment for the better, whereas others only show a main effect of the intervention. Obviously, a number of studies lacked the statistical power to find any significant moderation due to small sample size. A quantitative or meta-analysis, however, allows for a more powerful estimate of the overall trend in the data. Meta-analysis can thus be used to examine whether the randomized G × E experiments support or refute the differential susceptibility model.

The 22 studies discussed in the previous sections included 3,257 participants, 1,228 of whom were carriers of susceptibility genes. The combined effect size of the intervention effects in this susceptible group amounted to a Pearson $r = 0.33$ (95% CI 0.23, 0.42; $p < 0.01$). The nonsusceptible group consisted of 2,029 cases, and the combined size of the intervention effects in this group was not significant, $r = 0.08$ (95% CI −0.02, 0.17; $p = 0.12$). The contrast between the two combined effect sizes was significant ($p < 0.01$), with much stronger effects for the susceptible group (for details, see Van IJzendoorn & Bakermans-Kranenburg 2015). In the 14 studies with more than 80% Caucasian participants ($N = 689$ susceptible, $N = 1{,}371$ nonsusceptible), we found basically the same results, with significantly larger intervention effects for the susceptible genotypes. The eight studies with less than 80% Caucasian participants were too heterogeneous in ethnicity to be combined in separate analyses.

Twelve interventions targeted externalizing behaviors (including alcohol use). Carriers of susceptible genotypes were significantly more affected by the interventions ($r = 0.31$, $p < 0.01$) than their nonsusceptible peers ($r = 0.01$, $p = 0.87$, contrast $p < 0.01$). Eleven studies were field trials; they were significantly more effective in carriers of susceptible genotypes ($r = 0.34$, $p < 0.01$) than in carriers of nonsusceptible genotypes ($r = 0.04$, $p = 0.60$, contrast $p < 0.01$). The difference among the nine microtrials was not significant (susceptible groups $r = 0.30$, $p < 0.01$, nonsusceptible groups $r = 0.17$, $p < 0.01$, contrast $p = 0.09$), but the two nanotrials showed larger effects in susceptible genotype groups ($r = 0.38$, $p < 0.01$) than in nonsusceptible genotypes ($r = -0.09$, $p = 0.32$). Dopamine-related genes were indeed markers of susceptibility; the 11 studies with dopamine-related genotypes as moderator showed larger intervention effects in susceptible genotype groups ($r = 0.35$, $p < 0.01$) than in nonsusceptible genotypes ($r = -0.00$, $p = 0.96$, contrast $p < 0.01$). Seven studies with *5-HTTLPR* as moderator showed significant combined effects in the susceptible genotype group ($r = 0.30$, $p < 0.01$) but also in the nonsusceptible genotype ($r = 0.16$, $p = 0.04$); the contrast was not significant ($p = 0.15$).

As a final step, we computed the difference between the Fisher Z-transformed effect sizes for the susceptible and nonsusceptible groups within each study. The combined effect size was *Fisher* $Z = 0.23$ (95% CI 0.09, 0.37; $p < 0.01$), showing a significant combined effect for the difference between susceptible and nonsusceptible genotypes. The funnel plot of these effect sizes did not show publication bias, thus trim-and-fill was not necessary, and the Eggers test was not significant.

In sum, the meta-analytic results indicate that randomized G × E experiments testing the bright side of moderation by genotypes related to vulnerability to negative experiences support the differential susceptibility model. Dopamine-related genes emerged clearly as susceptibility markers. The effects of experimental manipulation of the environment for the better were much stronger in the susceptible genotypes than in the nonsusceptible genotypes (see **Figure 2**).

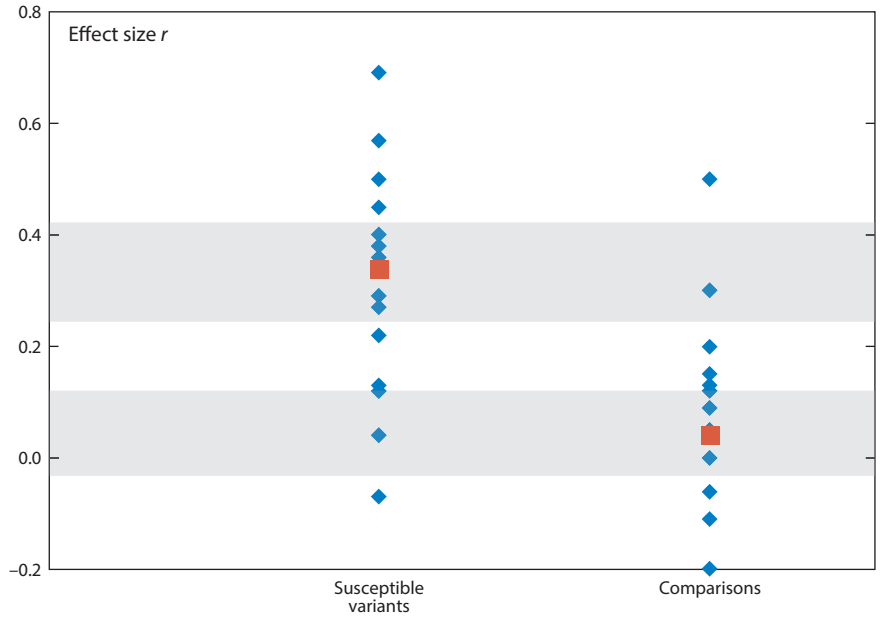

Figure 2

A meta-analysis of randomized G × E experiments testing the bright side of the moderating role of genotypes shown to be related to vulnerability to negative conditions. The intervention effects in subgroups carrying the susceptible variants are depicted in the left side of the figure, and the effect in the comparisons is depicted in the right side. The combined effect size of the intervention effects in susceptible variants was $r = 0.33$, $p < 0.01$ (*red square*) with a 95% confidence interval ranging from 0.23 to 0.42 (*grey area*). The combined effect size of the intervention effects in the comparisons was $r = 0.08$, $p = 0.12$ (*red square*) with a 95% confidence interval ranging from −0.02 to 0.17 (*grey area*). The combined effect size for carriers of the susceptible variants was significantly larger than that for the comparisons, $p < 0.01$.

FUTURE DIRECTIONS

What are the next steps in G × E experiments? Extending the G component of the G × E equation by including genetic pathways seems a logical follow-up to the work done so far with candidate genes, while including methylation level and gene expression may open up new horizons for broadening the E component of the G × E equation.

Genetic Pathways

In most G × E experiments, single genes or only a few candidate genes are selected as markers of differential susceptibility. Of course, candidate genes are never supposed to carry the whole weight of genetic influences, but they are considered sensitive indicators of underlying genetic pathways. With more efficient genotyping methods becoming available, the G component in the G × E equation may be broadened to include genetic pathways. Genetic pathways are biologically based sets of functional variants of genes that together regulate the modulation of specific neurotransmitter systems such as the serotonin or dopamine systems. This approach is different from a cumulative-genetic or plasticity gradient (Belsky & Beaver 2011, Brody et al. 2013) that counts the number of risk or plasticity polymorphisms and assumes that more of those polymorphisms lead to more susceptibility regardless of their functional biological cohesiveness.

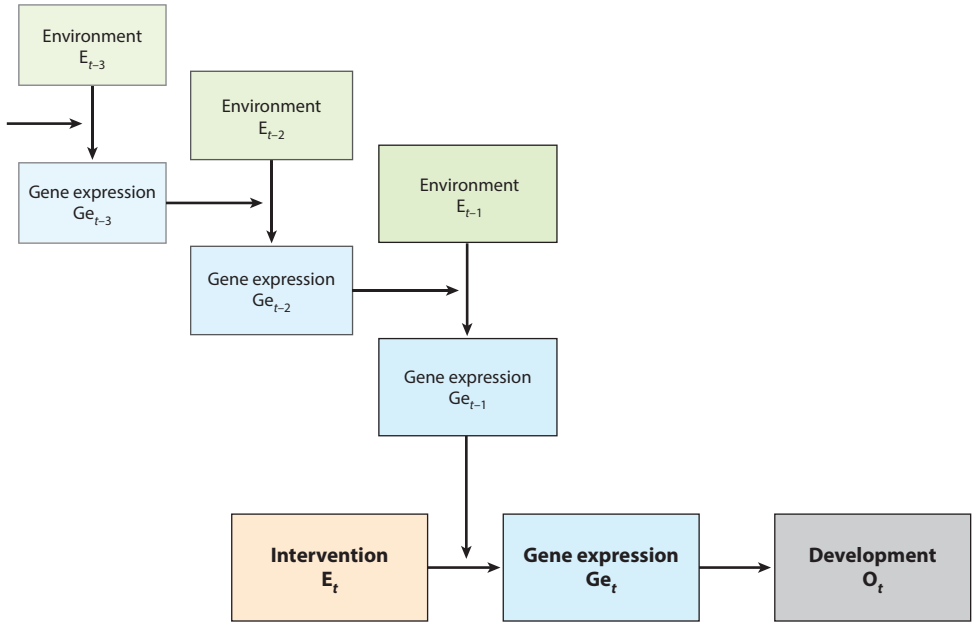

Figure 3

Model of gene-expression-by-environment experiments (Ge × eE). In this model the effects of experimental interventions at a specific point in time (E_t) on development (O_t) are mediated by differentially expressed genes (Ge_t), and moderated by expressed genes (Ge_{t-1}) that in turn result from genetically (Ge_{t-2}) moderated environmental influences (E_{t-1}), and so on. Adapted with permission from Van IJzendoorn et al. (2011b).

Epigenetics: Methylation and Gene Expression

In the end, gene expression is crucial because epigenetic changes through methylation or acetylation might affect the functional significance of structurally identical genotypes (Fraga et al. 2005, Meaney 2010, Van IJzendoorn et al. 2011b). For example, methylated *5-HTTl* alleles might become functionally more similar to nonmethylated *s* alleles (Van IJzendoorn et al. 2010). Ironically, dandelions are a prime example of the power of epigenetics because they show flexible adaptation despite asexual reproduction. The plants that grow from the mother plant's seeds are structurally genetic clones, identical to the mother plant, but their methylation patterns vary strongly in response to environmental conditions (Verhoeven et al. 2010). The implication is that G × E experiments should evolve into gene-expression-by-environment experiments, or Ge × eE.

A possible mechanism might be that prenatal methylation causes some individuals to be prenatally programmed in a way that makes them postnatally more liable to respond negatively to adversity (e.g., see Oberlander et al. 2008). But when these individuals are postnatally exposed to positive environments or interventions, changes in methylation might occur rather quickly and lead to more optimal development in response to the enhanced quality of the child-rearing environment. This would lead to an epigenetically enriched model of differential susceptibility (see **Figure 3**). In this model, the effects of experimental interventions (E_2) on human development (O) are hypothesized to be mediated by epigenetic changes in specific genes (Ge_2) and moderated by differentially expressed genes (Ge_1), which in turn result from genetically (G) moderated prenatal influences (E_1) (see **Figure 3**).

The Shifting Balance Between G and E

Children's genetic make-up affects their rearing environment, which in its turn canalizes children's development. Passive or evocative G × E correlations have been argued to turn G × E into G × G (Manuck & McCaffery 2014). However, the environment also leaves its imprint on the genome through epigenetic processes to determine the onset or offset of gene expression, regulating the production of proteins and enzymes crucial for development (Meaney 2010). Epigenetics, conceptualized as environment-dependent genetic effects, might thus be interpreted as turning G × E into Ge × E, or more radically, E × E. This illustrates the persistent but unproductive diathesis of genes and environments. Instead of thinking in terms of independent components of the G × E equation, it is time for an altered focus on the inherent interdependence of G and E. Replying (after a public lecture) to the question of which factor contributed more to the development of personality, nature or nurture, psychologist Donald Hebb argued aptly that G and E are so tightly interwoven that they determine development not unlike length and width determine the surface of a rectangle (see Meaney 2001). Environmentally impregnated genes and genetically shaped environments are involved in continuous interplay, and the outcome of this interaction is development.

CONCLUSION

In the sections above, we have reviewed G × E research from various angles. Here we point to gaps in genetic research that illustrate why G × E experiments are needed, and we identify gaps in G × E experiments inspired by the differential susceptibility model. We conclude that G × E experiments are vital to further unravel the issue of what intervention works for whom.

Missing Heritability or Missing Environment?

Genetic research is immensely complex, and the various pieces of the puzzle have not added up to a coherent overall picture. Twin studies have shown that heritability accounts for 40–60% of the variance in almost any human trait. However, genome-wide association studies (GWAS) have failed to find combinations of SNPs that explain more than a few percentages of this variance (Plomin 2013, Thomas 2010b). Genome-wide complex trait analysis (Yang et al. 2011) is a promising tool for estimating heritability by pairwise comparisons of biologically unrelated individuals' overall genetic make-up and their complex phenotype, but it still leaves a considerable gap in comparison to behavioral genetics estimates of the same traits (Trzaskowski et al. 2013). Common to these approaches is the neglect of the environment. In addition to ingenious quasi-experiments (e.g., Conley et al. 2013), G × E experiments are important tools for examining the interplay between genes and environment (Van IJzendoorn et al. 2011a). The environment is unpacked in micro- and nanotrials that provide evidence with regard to what manipulation of which specific environmental dimension results in measurable change in individuals with predefined characteristics.

Testing the Two Sides of Differential Susceptibility

G × E trials also allow for testing the most daring and fundamental proposition of the differential susceptibility hypothesis, namely that the very same individuals who show the worst outcomes in untoward environments would also profit most from enriched or supportive environments. Note that in none of the studies conducted so far has this idea been tested experimentally. All studies have compared individuals exposed to different environments using between-subject designs. Experimental within-subject designs enable proper testing of the "for better and for worse"

aspect of the hypothesis. The only experiment coming close is the between-subjects nanotrial using attention bias modification with both positive and negative pictures (Fox et al. 2011). In this seminal study, effects were stronger for negative and positive conditions in individuals with the hypothesized susceptible genotype than in individuals with the other genotype. This pattern of results is clearly consistent with differential susceptibility.

Ethics of G × E Applications

The search for optimal fit between individual characteristics and type of intervention may be critical for the efficient use of limited resources (Ellis et al. 2011). Personalized interventions based on genetics may come to be as desirable as personalized medicine (Schleidgen et al. 2013). Differential susceptibility theory, however, adds an ethical complication in implying that environmental changes might be effective only for some individuals (orchids) and not for others (dandelions). If and only if susceptibility is categorical and generic, i.e., there is no gradient of susceptibility, and nonsusceptibility is not specific to one domain or type of intervention, this may lead to the ethical dilemma of having to choose between costly universal interventions or discriminatory selective interventions. G × E experiments are essential to sort out whether differential susceptibility is a categorical trait or a dimensional characteristic, and whether it is generic or domain specific.

The Hidden Efficacy of Interventions

Behavioral, educational, therapeutic, and social interventions seem to suffer from very modest efficacy in reaching their goals (Conti & Heckman 2010, Van IJzendoorn et al. 2011a). By aggregating across more-susceptible and less-susceptible individuals, effect estimates may be strongly underestimated. Our meta-analysis of the first wave of G × E experiments documents the large gap between the minimal intervention effects in nonsusceptible groups and the substantial intervention effects in susceptible individuals. Differential susceptibility is a promising new way to address the perennial issue of what works for whom (Halford et al. 2008, Roth & Fonagy 2005), with a priori expectations about crucial genetic moderators. G × E experiments may uncover the hidden efficacy of interventions.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

The authors were supported by awards from the Netherlands Organization for Scientific Research (M.J. Bakermans-Kranenburg: Vici grant; M.H. van IJzendoorn: Spinoza Prize) and the Gravitation program of the Dutch Ministry of Education, Culture, and Science and the Netherlands Organization for Scientific Research (NWO grant number 024.001.003).

LITERATURE CITED

Aiken LS, West SG. 1991. *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage

Albert WD, Belsky DW, Crowley DM, Latendresse SJ, Aliev F, et al. 2015. Can genetics predict response to complex behavioral interventions? Evidence from a genetic analysis of the Fast Track Randomized Control Trial. *Dev. Psychopathol.* In press

Andersson E, Ruck C, Lavebratt C, Hedman E, Schalling M, et al. 2013. Genetic polymorphisms in monoamine systems and outcome of cognitive behavior therapy for social anxiety disorder. *PLOS ONE* 8:e79015

Bakermans-Kranenburg MJ, Van IJzendoorn MH. 2007. Genetic vulnerability or differential susceptibility in child development: the case of attachment. *J. Child Psychol. Psychiatry* 48:1160–73

Bakermans-Kranenburg MJ, Van IJzendoorn MH. 2011. Differential susceptibility to rearing environment depending on dopamine-related genes: new evidence and a meta-analysis. *Dev. Psychopathol.* 23:39–52

Bakermans-Kranenburg MJ, Van IJzendoorn MH, Juffer F. 2003. Less is more: meta-analyses of sensitivity and attachment interventions in early childhood. *Psychol. Bull.* 129:195–215

Bakermans-Kranenburg MJ, Van IJzendoorn MH, Mesman J, Alink LRA, Juffer F. 2008a. Effects of an attachment-based intervention on daily cortisol moderated by *DRD4*: a randomized control trial on 1–3-year-olds screened for externalizing behavior. *Dev. Psychopathol.* 20:805–20

Bakermans-Kranenburg MJ, Van IJzendoorn MH, Pijlman FT, Mesman J, Juffer F. 2008b. Experimental evidence for differential susceptibility: Dopamine D4 receptor polymorphism (DRD4 VNTR) moderates intervention effects on toddlers' externalizing behavior in a randomized controlled trial. *Dev. Psychol.* 44:293–300

Beach SRH, Brody GH, Lei MK, Philibert RA. 2010. Differential susceptibility to parenting among African American youths: testing the DRD4 hypothesis. *J. Fam. Psychol.* 24:513–21

Belsky J. 1997a. Variation in susceptibility to rearing influences: an evolutionary argument. *Psychol. Inq.* 8:182–86

Belsky J. 1997b. Theory testing, effect-size evaluation, and differential susceptibility to rearing influence: the case of mothering and attachment. *Child Dev.* 68:598–600

Belsky J. 2000. Conditional and alternative reproductive strategies: individual differences in susceptibility to rearing experience. In *Genetic Influences on Human Fertility and Sexuality: Theoretical and Empirical Contributions from the Biological and Behavioral Sciences*, ed. J Rodgers, D Rowe, W Miller, pp. 127–46. Boston: Kluwer

Belsky J. 2005. Differential susceptibility to rearing influences: an evolutionary hypothesis and some evidence. In *Origins of the Social Mind: Evolutionary Psychology and Child Development*, ed. B Ellis, D Bjorklund, pp. 139–63. New York: Guilford

Belsky J, Bakermans-Kranenburg MJ, Van IJzendoorn MH. 2007. For better and for worse: differential susceptibility to environmental influences. *Curr. Dir. Psychol. Sci.* 16:300–4

Belsky J, Beaver KM. 2011. Cumulative-genetic plasticity, parenting and adolescent self-regulation. *J. Child Psychol. Psychiatry* 52:619–26

Belsky J, Jonassaint C, Pluess M, Stanton M, Brummett B, Williams R. 2009. Vulnerability genes or plasticity genes? *Mol. Psychiatry* 14:746–54

Belsky J, Pluess M. 2013. Genetic moderation of early child-care effects on social functioning across childhood: a developmental analysis. *Child Dev.* 84:1209–25

Belsky J, Pluess M, Widaman KF. 2013. Confirmatory and competitive evaluation of alternative gene–environment interaction hypotheses. *J. Child Psychol. Psychiatry* 54:1135–43

Bockting CL, Mocking RJ, Lok A, Koeter MW, Schene AH. 2013. Therapygenetics: the *5HTTLPR* as a biomarker for response to psychological therapy? *Mol. Psychiatry* 18:744–45

Boyce WT, Chesney M, Alkon A, Tschann JM, Adams S, et al. 1995. Psychobiologic reactivity to stress and childhood respiratory illnesses: results of two prospective studies. *Psychosom. Med.* 57:411–22

Boyce WT, Ellis BJ. 2005. Biological sensitivity to context: I. An evolutionary-developmental theory of the origins and functions of stress reactivity. *Dev. Psychopathol.* 17:271–301

Brett ZH, Humphreys KL, Smyke AT, Gleason MM, Nelson CA, et al. 2015. *5HTTLPR* genotype moderates the longitudinal impact of early caregiving and attachment on externalizing behavior. *Dev. Psychopathol.* In press

Brody GH, Chen YF, Beach SRH. 2013. Differential susceptibility to prevention: GABAergic, dopaminergic, and multilocus effects. *J. Child Psychol. Psychiatry* 54:863–71

Brody GH, Chen YF, Beach SRH, Kogan SM, Yu T, et al. 2014. Differential sensitivity to prevention programming: a dopaminergic polymorphism-enhanced prevention effect on protective parenting and adolescent substance use. *Health Psychol.* 33:182–91

Brody GH, Murry VM, Kogan SM, Gerrard M, Gibbons FX, et al. 2006. The Strong African American Families program: a cluster randomized prevention trial of long-term effects and a mediational model. *J. Consult. Clin. Psychol.* 74:356–66

Caspi A, Hariri AR, Holmes A, Uher R, Moffitt TE. 2010. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *Am. J. Psychiatry* 167:509–27

Caspi A, McClay J, Moffitt TE, Mill J, Martin J, et al. 2002. Role of genotype in the cycle of violence in maltreated children. *Science* 297:851–54

Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, et al. 2003. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* 301:386–89

Chang F-M, Kidd JR, Livak KJ, Pakstis AJ, Kidd KK. 1996. The worldwide distribution of allele frequencies at the human dopamine D4 receptor locus. *Hum. Genet.* 98:91–101

Cheon K, Kim B, Cho S. 2007. Association of 4-repeat allele of the dopamine D4 receptor gene exon III polymorphism and response to methylphenidate treatment in Korean ADHD children. *Neuropsychopharmacology* 32:1377–83

Chiao JY, Blizinsky KD. 2010. Culture-gene coevolution of individualism-collectivism and the serotonin transporter gene. *Proc. R. Soc. B* 277:529–37

Cicchetti D, Rogosch FA, Toth SL. 2011. The effects of child maltreatment and polymorphisms of the serotonin transporter and dopamine D4 receptor genes on infant attachment and intervention efficacy. *Dev. Psychopathol.* 23:357–72

Cleveland HH, Schlomer GL, Van den Bergh DJ, Feinberg M, Greenberg M, et al. 2015. Influences of interventions, parenting, and the *DRD4* gene on alcohol use in early adolescence. *Dev. Psychopathol.* In press

Conley D, Rauscher E, Siegal ML. 2013. Beyond orchids and dandelions: testing the *5-HTT* "risky" allele for evidence of phenotypic capacitance and frequency-dependent selection. *Biodemogr. Soc. Biol.* 59:37–56

Conti G, Heckman JJ. 2010. Understanding the early origins of the education-health gradient: a framework that can also be applied to analyze gene-environment interactions. *Perspect. Psychol. Sci.* 5:585–605

Dall SRX, Houston AI, McNamara JM. 2004. The behavioral ecology of personality: consistent individual differences from an adaptive perspective. *Ecol. Lett.* 7:734–39

DeWitt TJ, Sih A, Wilson DS. 1998. Costs and limits of phenotypic plasticity. *Trends Ecol. Evol.* 13:77–81

Dingemanse NJ, Wolf M. 2013. Between-individual differences in behavioral plasticity within populations: causes and consequences. *Anim. Behav.* 85:1031–39

Dobbs D. 2009. The science of success. *Atlantic* 12:1–15

Dobson SD, Brent LJN. 2013. On the evolution of the serotonin transporter linked polymorphic region (5-HTTLPR) in primates. *Front. Hum. Neurosci.* 7:1–9

Duncan LE, Keller MC. 2011. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *Am. J. Psychiatry* 168:1041–49

Eley TC, Hudson JL, Creswell C, Tropeano M, Lester KJ, et al. 2012. Therapygenetics: the *5HTTLPR* and response to psychological therapy. *Mol. Psychiatry* 17:236–37

Ellis BJ, Boyce WT, Belsky J, Bakermans-Kranenburg MJ, Van IJzendoorn MH. 2011. Differential susceptibility to the environment: a neurodevelopmental theory. *Dev. Psychopathol.* 23:7–28

Ellis BJ, Essex MJ, Boyce WT. 2005. Biological sensitivity to context: II. Empirical explorations of an evolutionary-developmental theory. *Dev. Psychopathol.* 17:303–28

Fox E, Zougkou K, Ridgewell A, Garner K. 2011. The serotonin transporter gene alters sensitivity to attention bias modification: evidence for a plasticity gene. *Biol. Psychiatry* 70:1049–54

Fraga MF, Ballestar E, Paz MF, Ropero S, Setien F, et al. 2005. Epigenetic differences arise during the lifetime of monozygotic twins. *Proc. Natl. Acad. Sci. USA* 102:10604–9

Gallardo-Pujol D, Andrés-Pueyo A, Maydeu-Olivares A. 2013. *MAOA* genotype, social exclusion and aggression: an experimental test of a gene-environment interaction. *Genes Brain Behav.* 12:140–45

Gelernter J, Cubells JF, Kidd JR, Pakstis AJ, Kidd KK. 1999. Population studies of polymorphisms of the serotonin transporter protein gene. *Am. J. Med. Genet.* 88:61–66

Gotlib IH, Joormann J, Minor KL, Hallmayer J. 2008. HPA-axis reactivity: a mechanism underlying the associations among *5-HTTLPR*, stress, and depression. *Biol. Psychiatry* 63:847–51

Gottesman II, Shields J. 1967. A polygenic theory of schizophrenia. *Proc. Natl. Acad. Sci. USA* 58:199–205
Hakamata Y, Lissek S, Bar-Haim Y, Britton JC, Fox NA, et al. 2010. Attention bias modification treatment: a meta-analysis toward the establishment of novel treatment for anxiety. *Biol. Psychiatry* 68:982–90
Halford WK, Markman HJ, Stanley S. 2008. Strengthening couples' relationships with education: social policy and public health perspectives. *J. Fam. Psychol.* 22:497–505
Hamilton WD. 1964a. The genetical evolution of social behavior: I. *J. Theor. Biol.* 7:1–16
Hamilton WD. 1964b. The genetical evolution of social behavior: II. *J. Theor. Biol.* 7:17–52
Hariri AR, Mattay VS, Tessitore A, Kolachana B, Fera F, et al. 2002. Serotonin transporter genetic variation and the response of the human amygdala. *Science* 297:400–3
Homberg JR, Lesch KP. 2011. Looking on the bright side of serotonin transporter gene variation. *Biol. Psychiatry* 69:513–19
Israel S, Lerer E, Shalev I, Uzefovsky F, Riebold M, et al. 2009. The oxytocin receptor (OXTR) contributes to prosocial fund allocations in the dictator game and the social value orientations task. *PLOS ONE* 4:e5535
Karg K, Burmeister M, Shedden K, Sen S. 2011. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. *Arch. Gen. Psychiatry* 68:444–54
Kegel CAT, Bus AG, Van IJzendoorn MH. 2011. Differential susceptibility in early literacy instruction through computer games: the role of the dopamine D4 receptor gene (*DRD4*). *Mind Brain Educ.* 5:71–79
Keller MC. 2014. Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biol. Psychiatry* 75:18–24
Kochanska G, Kim S, Barry RA, Philibert RA. 2011. Children's genotypes interact with maternal responsive care in predicting children's competence: diathesis-stress or differential susceptibility? *Dev. Psychopathol.* 22:605–16
Kohen R, Cain KC, Buzaitis A, Johnson V, Becker KJ, et al. 2011. Response to psychosocial treatment in post-stroke depression is associated with serotonin transporter polymorphisms. *Stroke* 42:2068–70
MacLeod C, Rutherford E, Campbell L, Ebsworthy G, Holker L. 2002. Selective attention and emotional vulnerability: assessing the causal basis of their association through the experimental manipulation of attentional bias. *J. Abnorm. Psychol.* 111:107–23
Manuck SB. 2010. The reaction norm in gene × environment interaction. *Mol. Psychiatry* 15:881–82
Manuck SB, McCaffery JM. 2014. Gene-environment interaction. *Annu. Rev. Psychol.* 65:41–70
Markus CR, Verschoor E, Smeets T. 2012. Differential effect of the *5-HTT* gene-linked polymorphic region on emotional eating during stress exposure following tryptophan challenge. *J. Nutr. Biochem.* 23:410–16
McClelland GH, Judd CM. 1993. Statistical difficulties of detecting interactions and moderator effects. *Psychol. Bull.* 114:376–90
McDermott R, Tingley D, Cowden J, Frazzetto G, Johnson DDP. 2009. Monoamine oxidase A gene (*MAOA*) predicts behavioral aggression following provocation. *Proc. Natl. Acad. Sci. USA* 106:2118–23
McGuffin P, Alsabban S, Uher R. 2011. The truth about genetic variation in the serotonin transporter gene and response to stress and medication. *Br. J. Psychiatry* 198:424–27
Meaney MJ. 2001. Nature, nurture, and the disunity of knowledge. *Ann. N. Y. Acad. Sci.* 935:50–61
Meaney MJ. 2010. Epigenetics and the biological definition of gene × environment interactions. *Child Dev.* 81:41–79
Monroe SM, Simons AD. 1991. Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. *Psychol. Bull.* 110:406–25
Mueller A, Brocke B, Fries E, Lesch KP, Kirschbaum C. 2010. The role of the serotonin transporter polymorphism for the endocrine stress response in newborns. *Psychoneuroendocrinology* 35:289–96
Munafò MR, Durrant C, Lewis G, Flint J. 2009. Gene-environment interactions at the serotonin transporter locus. *Biol. Psychiatry* 65:211–19
Munafò MR, Zammit S, Flint J. 2014. A critical perspective on gene-environment interaction models: What impact should they have on clinical perceptions and practice? *J. Child Psychol. Psychiatry* 55:1092–101
Nelson CA, Fox NA, Zeanah CH. 2014. *Romania's Abandoned Children: Deprivation, Brain Development, and the Struggle for Recovery*. Cambridge, MA: Harvard Univ. Press

Oberlander TF, Weinberg J, Papsdorf M, Grunau R, Misri S, Devlin AM. 2008. Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (*NR3C1*) and infant cortisol stress responses. *Epigenetics* 3:97–106

Pergamin-Hight L, Bakermans-Kranenburg MJ, Van IJzendoorn MH, Bar-Haim Y. 2012. Variations in the promoter region of the serotonin-transporter gene (*5HTTLPR*) and biased attention for emotional information: a meta-analysis. *Biol. Psychol.* 71:373–79

Plak R, Kegel C, Bus AG. 2015. Genetic differential susceptibility in literacy delayed children: a randomized controlled trial on emergent literacy in kindergarten. *Dev. Psychopathol.* In press

Plomin R. 2013. Child development and molecular genetics: 14 years later. *Child Dev.* 84:104–20

Pluess M, Belsky J. 2013. Vantage sensitivity: individual differences in response to positive experiences. *Psychol. Bull.* 139:901–16

Popper KR. 1979. *Objective Knowledge: An Evolutionary Approach*. Oxford, UK: Oxford Univ. Press

Preacher KJ, Curran PJ, Bauer DJ. 2006. Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *J. Educ. Behav. Stat.* 31:437–48

Risch N, Herrell R, Lehner T, Liang KY, Eaves L, et al. 2009. Interaction between the serotonin transporter gene (*5-HTTLPR*), stressful life events, and risk of depression: a meta-analysis. *JAMA* 301:2462–71

Rockoff JD. 2013. Nature vs. nurture: New science stirs debate. How behavior is shaped; who's an orchid, who's a dandelion. *Wall Street J.*, Sept. 16

Roisman GI, Newman DA, Fraley C, Haltigan JD, Groh AM, et al. 2012. Distinguishing differential susceptibility from diathesis-stress: recommendations for evaluating interaction effects. *Dev. Psychopathol.* 24:389–409

Roth A, Fonagy P. 2005. *What Works for Whom? A Critical Review of Psychotherapy Research.* New York: Guilford. 2nd ed.

Rutter M. 2006. *Genes and Behavior. Nature-Nurture Interplay Explained*. Oxford, UK: Blackwell

Rutter M. 2010. Gene-environment interplay. *Depress. Anxiety* 27:1–4

Sameroff AJ. 1983. Developmental systems: contexts and evolution. In *Handbook of Child Psychology*, ed. P Mussen, pp. 237–94. New York: Wiley

Sasaki JY, Kim HS, Mojaverian T, Kelley LDS, Park IY, Janušonis S. 2013. Religion priming differentially increases prosocial behavior among variants of the dopamine D4 receptor (*DRD4*) gene. *Soc. Cogn. Aff. Neurosci.* 8:209–15

Schleidgen S, Klingler C, Bertram T, Rogowski WH, Marckmann G. 2013. What is personalized medicine: sharpening a vague term based on systematic literature review. *BMC Med. Ethics* 14:55

Snell-Rood EC. 2013. An overview of the evolutionary causes and consequences of behavioural plasticity. *Anim. Behav.* 85:1004–11

Söderqvist S, Bergman Nutley S, Peyrard-Janvid M, Matsson H, Humphreys K, et al. 2012. Dopamine, working memory, and training induced plasticity: implications for developmental research. *Dev. Psychol.* 48:836–43

Spoth R, Guyll M, Lillehoj CJ, Redmond C, Greenberg M. 2007. PROSPER study of evidence-based intervention implementation quality by community-university partnerships. *J. Community Psychol.* 35:981–99

Sulik MJ, Eisenberg N, Lemery-Chalfant K, Spinrad TL, Silva KM, et al. 2012. Interactions between serotonin transporter gene haplotypes and quality of mothers' parenting predict the development of children's noncompliance. *Dev. Psychol.* 48:740–54

Sulloway FJ. 1996. *Born to Rebel: Birth Order, Family Dynamics, and Creative Lives*. New York: Pantheon

Suomi SJ. 2006. Risk, resilience, and gene × environment interactions in rhesus monkeys. *Res. Child.* 1094:52–62

Sweitzer MM, Halder I, Flory JD, Craig AE, Gianaros PJ, et al. 2012. Polymorphic variation in the dopamine D4 receptor predicts delay discounting as a function of childhood socioeconomic status: evidence for differential susceptibility. *Soc. Cogn. Affect. Neurosci.* 8:499–508

Taylor SE, Way BM, Seeman TE. 2011. Early adversity and adult health outcomes. *Dev. Psychopathol.* 23:939–54

Thomas D. 2010a. Methods for investigating gene-environment interactions in candidate pathway and genome-wide association studies. *Annu. Rev. Public Health* 31:21–36

Thomas D. 2010b. Gene-environment-wide association studies: emerging approaches. *Nat. Rev. Gen.* 11:259–72

Trzaskowski M, Dale PS, Plomin R. 2013. No genetic influence for childhood behavior problems from DNA analysis. *J. Am. Acad. Child Adolesc. Psychiatry* 52:1048–56

Uher R, McGuffin P. 2008. The moderation by the serotonin transporter gene of environmental adversity in the aetiology of mental illness: review and methodological analysis. *Mol. Psychiatry* 13:131–46

Uher R, McGuffin P. 2010. The moderation by the serotonin transporter gene of environmental adversity in the etiology of depression: 2009 update. *Mol. Psychiatry* 15:18–22

Van den Hoofdakker BJ, Nauta MH, Dijck-Brouwer DAJ, van der Veen-Mulders L, Sytema S, et al. 2012. Dopamine transporter gene moderates response to behavioral parent training in children with ADHD: a pilot study. *Dev. Psychol.* 48:567–74

Van IJzendoorn MH, Bakermans-Kranenburg MJ. 2012. Differential susceptibility experiments: going beyond correlational evidence: comment on beyond mental health, differential susceptibility articles. *Dev. Psychol.* 48:769–74

Van IJzendoorn MH, Bakermans-Kranenburg MJ. 2015. Genetic differential susceptibility on trial: meta-analytic support from randomized controlled experiments. *Dev. Psychopathol.* In press

Van IJzendoorn MH, Bakermans-Kranenburg MJ, Belsky J, Beach S, Brody G, et al. 2011a. Gene-by-environment experiments: a new approach to finding the missing heritability. *Nat. Rev. Genet.* 12:881

Van IJzendoorn MH, Bakermans-Kranenburg MJ, Ebstein RP. 2011b. Methylation matters in child development: toward developmental behavioral epigenetics. *Child Dev. Perspect.* 5:305–10

Van IJzendoorn MH, Belsky J, Bakermans-Kranenburg MJ. 2012. Serotonin transporter genotype *5HTTLPR* as a marker of differential susceptibility? A meta-analysis of child and adolescent gene-by-environment studies. *Transl. Psychiatry* 2:1–6

Van IJzendoorn MH, Caspers K, Bakermans-Kranenburg MJ, Beach SRH, Philibert R. 2010. Methylation matters: Interaction between methylation density and *5HTT* genotype predicts unresolved loss or trauma. *Biol. Psychiatry* 68:405–7

Verhoeven JF, Jansen JJ, Van Dijk PJ, Biere A. 2010. Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytol.* 185:1108–18

Verona E, Joiner TE, Johnson F, Bender TW. 2006. Gender specific gene-environment interactions on laboratory-assessed aggression. *Biol. Psychol.* 71:33–41

Verschoor E, Markus CR. 2012. Physiological and affective reactivity to a 35% CO_2 inhalation challenge in individuals differing in the 5-HTTLPR genotype and trait neuroticism. *Eur. Neuropsychopharmacol.* 22:546–54

Vijayendran M, Cutrona C, Beach SRH, Brody GH, Russell D, Philibert RA. 2012. The relationship of the serotonin transporter (SLC6A4) extra long variant to gene expression in an African American sample. *Am. J. Med. Genet. Part B* 159B:611–12

Wachs TD, Plomin R. 1991. *Conceptualization and Measurement of Organism-Environment Interaction*. Washington, DC: Am. Psychol. Assoc.

Way BM, Taylor SE. 2010. The serotonin transporter promoter polymorphism is associated with cortisol response to psychosocial stress. *Biol. Psychiatry* 67:487–92

Way BM, Taylor SE. 2011. A polymorphism in the serotonin transporter gene moderates cardiovascular reactivity to psychosocial stress. *Psychosom. Med.* 73:310–17

Widaman KF, Helm JL, Castro-Schilo L, Pluess M, Stallings MC, Belsky J. 2012. Distinguishing ordinal and disordinal interactions. *Psychol. Methods* 17:615–22

Williams KD, Jarvis B. 2006. Cyberball: a program for use in research on interpersonal ostracism and acceptance. *Behav. Res. Methods* 38:174–80

Williams RB, Marchuk DA, Gadde KM, Barefoot JC, Grichnik K, et al. 2003. Serotonin-related gene polymorphisms and central nervous system serotonin functions. *Neuropsychopharmacology* 28:533–41

Wolf M, Van Doorn GS, Weissing FJ. 2008. Evolutionary emergence of responsive and unresponsive personalities. *Proc. Natl. Acad. Sci. USA* 105:15825–30

Wolf M, Van Doorn GS, Weissing FJ. 2011. On the coevolution of social responsiveness and behavioural consistency. *Proc. R. Soc. B* 278:440–48

Wong MY, Day NE, Luan JA, Chan KP, Wareham NJ. 2003. The detection of gene-environment interaction for continuous traits: Should we deal with measurement error by bigger studies or better measurement? *Int. J. Epidemiol.* 32:51–57

Yang J, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Gen.* 88:76–82

Zuckerman M. 1999. *Vulnerability to Psychopathology: A Biosocial Model*. Washington, DC: Am. Psychol. Assoc.

Annual Review of Psychology

Volume 66, 2015

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Psychology* articles may be found at http://www.annualreviews.org/errata/psych