Overview of Behavioral Genetics Research for Family Researchers

This article provides an overview of the methods, assumptions, and key findings of behavioral genetics methodology for family researchers with a limited background. We discuss how family researchers can utilize and contribute to the behavioral genetics field, particularly in terms of conducting research that seeks to explain shared environmental effects. This can be done, in part, by theoretically controlling for genetic confounds in research that seeks to determine cause-and-effect relationships among family variables and individual outcomes. Gene–environment correlation and interaction are especially promising areas for the family researcher to address. Given the methodological advancements in the field, we also briefly comment on new methods in molecular genetics for studying psychological mental health disorders.

OVERVIEW OF BEHAVIORAL GENETICS RESEARCH FOR FAMILY RESEARCHERS

The debate between proponents of nature (Scarr & McCartney, 1983) versus nurture (Collins, Maccoby, Steinberg, Hetherington, & Bornstein, 2000) as an explanation for social behavior is long standing, well documented, and continues today (see Dar-Nimrod & Heine, 2011; Rende & Slomkowski, 2009; Turkheimer, 2011). In family social science, we typically focus on nurture effects and can sometimes overlook possible biological and genetic factors when examining family processes (D’Onofrio & Lahey, 2010). This notion has been challenged by the broad conclusions that have come out of the behavioral genetics field claiming that additive genetic effects, rather than shared environmental effects, explain the most variance in traits and outcomes that we typically study (e.g., Plomin & Daniels, 1987). Those new to behavioral genetics research may not understand what the definitions of shared environmental effects are, how assumptions underlying behavioral genetics methods affect the interpretation of those findings, or how behavioral genetics findings challenge many of the original broad reports of little family influence on behavior (Burt, 2009; Horwitz & Neiderhiser, 2011; Legrand, Keyes, McGue, Iacono, & Krueger, 2008).

Consistent with a recent call for family social science to control for genes and other biological influences (D’Onofrio & Lahey, 2010; Horwitz & Neiderhiser, 2011), we believe that behavioral genetics findings have profound implications for family researchers. We address these implications by reviewing the methods, assumptions, and findings from the behavioral genetics field. This article is written by family researchers with a limited background.
researchers for family researchers who are not familiar with this line of work.

The majority of this article focuses on quantitative behavioral genetics research in terms of theory, methods, and key findings. We focus on future research questions that we believe are particularly appropriate for family researchers to address, including the identification of specific shared environmental variants in the development of adolescent substance use and related adjustment outcomes. We cover research on main effects, gene–environment correlation (rGE), and gene–environment interaction (G \times E). The work of behavioral genetics has led to research in molecular genetics that aims to identify specific genetic variants (main and G \times E effects) in relation to psychopathology. Because of this advancement, we believe it is necessary to also briefly review recent molecular genetic methods and then comment on the relevance of this research to family scholars. Finally, we should mention that this article is not intended to provide a comprehensive review of the field; rather, it is a theoretical review of areas in behavior genetics that seem especially promising for family researchers to address. For a more comprehensive review, readers should refer to Bazzett (2008); Dick, Latendresse, and Riley (2011); Kim (2009); and Plomin, DeFries, McClearn, and McGuffin (2008), which helped guide much of this article.

**WHAT IS BEHAVIORAL GENETICS?**

Behavioral genetics is a branch of psychology that attempts to allocate and explain genetic and environmental contributions to human (and animal) behavior (Behavior Genetics Association, n.d.). Pursuing such knowledge involves defining behavioral outcomes (called phenotypes) and measuring the corresponding genetic influences (called genotypes).

Quantitative behavioral genetics methodology theoretically and statistically controls for genetic effects. “Theoretical control” refers to simply comparing study results (e.g., the association between parenting and child externalizing behaviors) across different, genetically informed family subsamples (e.g., monozygotic versus dizygotic twins, full biological siblings versus adopted or step siblings). Statistically, this methodology controls for genetic effects through a simple variance–covariance decomposition (called biometric modeling). In the next section, we more fully describe biometric modeling and the assumptions used to test behavioral genetic models.

**A LOOK AT THE METHODS IN QUANTITATIVE BEHAVIORAL GENETICS**

Typically, one sees something like Figure 1 in quantitative behavioral genetics papers. Behavioral geneticists would describe this figure as a phenotype being “decomposed”

---

**FIGURE 1. DIAGRAM OF UNIVARIATE DECOMPOSITION.** Variance components are represented by capital letters (additive genetic: A; shared environmental: C; nonshared environmental effects: E) using a twin design. Path coefficients are represented by lower case letters (a, c, e). Circles represent latent factors, squares represent observed variables.
into three basic categories: additive genetic (A), shared environmental (C), and nonshared environmental (E) influences. This means that some behavioral outcome of interest (i.e., the phenotype) will be partialed out in via structural equation modeling to describe basic genetic and environmental effects in aggregate. To understand the basics of how these effects are calculated, it helps to first describe the sampling strategy used in such designs, as well as to give some formal definitions.

**Sampling Strategy**

A quantitative behavioral genetic study typically uses a genetically informed sibling-pair design (Plomin et al., 2008). This type of study recruits a sample of families with at least two siblings who share varying proportions of genes (either within or across families). A twin study is most often used. In the twin design, researchers compare results across monozygotic twins (who share 100% of their DNA) and dizygotic twins (who share 50% of their DNA, on average; e.g., Iacono & McGue, 2002). Another type of behavioral genetics study design is an adoption study (e.g., McGue et al., 2007). In this type of study, researchers compare siblings who are the biological offspring of the same parents (i.e., they are full siblings and share 50% of their DNA) to adopted siblings who are not biologically related to their parents or to one another (i.e., they share 0% of their DNA). In a full adoption study, the biological parents of the adopted offspring are also included (Leve, Neiderhiser, Scaramella, & Reiss, 2008). Correlation comparisons can be made across stepsiblings (who share 0% of their DNA), or half siblings (who share 25% of their DNA). Other designs are used, such as the Children of Twins (CoT) design (D’Onofrio, 2005); however, they are somewhat less common.

**Formal Definitions**

*Additive genetic influences* (A) provide a ballpark estimate of a trait’s heritability (Visscher, Hill, & Wray, 2008) and refer to the causes of two sibling’s similarity as a result of additive genetic influence. The *shared (or common rearing) environmental influence* (C) accounts for similarities between siblings that are not due to genetics (Burt, 2009). Typically, this refers to the family environment, but it is not necessary or sufficient for family variance to only be determined by shared environmental influence. This also includes any other environments siblings share, such as school, peers, and neighborhoods. Finally, *nonshared environmental influences* (E) are defined as unique environments, and as situational or contextual factors that make siblings different. Theoretically, this refers to anything in the environment that makes siblings different, including parental differential treatment or unique peer groups. Empirically, this also includes any measurement error.

**Behavioral Genetics Statistical Strategy: Biometric Modeling**

The variance components defined earlier (A, C, E) and path coefficients (a, c, e) are statistically estimated using biometric modeling techniques (Neale, 2008; Plomin et al., 2008). Squaring the standardized path coefficients gives the percentage of variance explained on the specific trait examined (i.e., \( a^2 = \) heritability, \( c^2 = \) shared environment contribution, \( e^2 = \) nonshared environment contribution). As Figure 1 describes, biometric modeling is based on structural equation modeling, which utilizes the concept of latent factors. Latent factors are variables that are “hidden,” or difficult to measure. In biometric modeling, the latent factors refer to genetic and environmental contributions to a phenotype of interest (A, C, and E, which are assumed to be 1.0).

Figure 1 shows two squares that depict an observed phenotype, one for “Sibling 1” and the other for “Sibling 2.” The circles refer to the latent factors (which are the variance components representing effects of additive genetic, A; shared environmental, C; and nonshared environmental, E). The arrows pointing from the circles to the phenotypes represent the estimation of the extent to which the variance component explains variance in the phenotype. The double-headed arrows connecting the two A variance components across siblings denotes a correlation (r), which is set a priori within sibling subsamples depending on their proportion of shared genes. This model is illustrated for a study of a single trait (phenotype) observed in monozygotic twins (rMZ) and dizygotic twins (rDZ). The A-to-A correlation for monozygotic twins is fixed to 1.0 (because they share 100% of their genes), and for dizygotic twins, it is fixed.
to .50 (because they share 50% of their genes). The correlation between $C$ and $C$ denotes the proportion of the shared environment the siblings have in common; it is set to 1.0 in both subsamples because biometric modeling assumes that monozygotic and dizygotic twins have the same environment (we return to this assumption in a moment).

Finally, there is no a priori expectation for the magnitude of the correlation between the nonshared environment components (i.e., $E$). Statistically, $E$ is anything left over after $A$ and $C$ have been partialed out, including measurement error; in fact, $E$ is akin to the residual (or error term) in a simple regression equation. Theoretically, and consistent with our earlier definition, $E$ refers to anything in the environment that makes siblings different (including unique environments and contextual and situational factors).

Given the gap between what is statistically measured in a biometric model and what is theoretically implied or inferred, care should be taken when interpreting $E$. In fact, studies have produced little consistent evidence of what is specifically accounted for by nonshared environmental influence (for a meta-analysis, see Turkheimer & Waldron, 2000). Not being able to identify specific nonshared environmental variants has led some to believe that nonshared environmental mediates for the most extreme difference in parental differential treatment. If there was a mismatch in concordance (if one twin had greater parent–child conflict but both twins had a similar level of externalizing problems), this would suggest common genetic or environmental factors explain this relationship.

### Biometric Modeling: The Assumptions

The interpretation of most statistical analyses relies on several assumptions. Assumptions can sometimes be easily glossed over or ignored. However, accurate interpretation of statistical findings is best done within an understanding of the analyses’ assumptions, and interpretation of biometric modeling results is no exception. In this section, we review three major biometric modeling assumptions.

The first is the assumption of equality in environments between subsample comparisons. This is a critical assumption in both twin and adoption designs that has sometimes been challenged. Regarding twin designs, monozygotic twins report being treated more similarly than do dizygotic twins during childhood (Borkenau, Riemann, Angleitner, & Spinath, 2002). Monozygotic twins also tend to spend more time together and have more similar peer groups compared to dizygotic twins (Kendler & Gardner, 1998). Even still, monozygotic-twin environmental similarity may be due to genetic rather than environmental influences (Loehlin & Nichols, 1976). In other words, monozygotic twins may spend more time together and share more friends than dizygotic twins as a result of their genetic similarity rather than a result of having a different environment.

The equality assumption is also challenged in adoption designs. For example, adoptive parents tend to have higher socioeconomic status (SES), more marital stability, and better mental health than parents with biological offspring, which are used as comparisons (McGue et al., 2007). Given this restriction of range, shared environmental effects can be deflated (Stoolmiller, 1999). Yet McGue et al. (2007) report that the range restriction in externalizing problems and SES had little effect on adoptive-sibling correlations for drug use, IQ, or delinquency.

The second major behavioral genetics assumption is that biometric modeling accounts only for additive genetic effects (Plomin, Haworth, & Davis, 2009). In general, the additive genetic effect refers to the effect of what varies in our genome, added up over multiple locations across the genome. However, not all genetic influence is additive (e.g., dominance,
epistasis). Whereas additive genetic influence refers to the affect of alleles, one by one, from all over the genome, dominance refers to the effect of one allele depending on the other in the same location in the genome. For example, phenylketonuria (PKU) operates with a dominant-recessive influence because affected offspring must have two copies of the PKU allele (if they have only one copy, they can be carriers and pass down that influence to their own children). Much like dominance, epitasis refers to the affect of one allele on another but at multiple locations across the genome. It has been argued that most psychological traits appear to have an additive, polygenic effect (Plomin et al., 2009). It is important to note though, that when we speak about heritability within an additive genetic framework, we are referring only to narrow, and not broad-sense heritability (broad-sense accounts for all genotypic variance; Hamilton, 2009).

Main-effects biometric modeling also assumes that there is no correlation or interaction between genetic and environmental effects. In fact, there has been an increasing amount of work done on both of these concepts, referred to gene–environment interplay (Horwitz & Neiderhiser, 2011; Moffitt, Caspi, & Rutter, 2005; Neiderhiser, 2011). In their landmark article, Scarr and McCartney (1983) describe three basic types of gene–environment correlations (rGE): (1) evocative rGE (e.g., parents respond to children in the same family differently as a result of their different genotypes); (2) active rGE (e.g., children actively seek out environments that fit their phenotypes), and (3) passive rGE (e.g., parents’ genotypes influence both their own and their child’s behaviors). These explanations have been utilized when trying to make sense of why some traits are heritable, such as parent-child relationship quality (e.g., McGue, Elkins, Walden, & Iacono, 2005). Moreover, there are statistical tests for gene–environment interaction (Purcell, 2002), and research in this area is extant (we cover some key points in this area of research later, after we finish covering assumptions and main effects findings).

Finally, biometric modeling assumes that dizygotic twins (or full siblings) share exactly 50% of their DNA, but this is actually an average. Full siblings can share about 40–60% of their DNA (Visscher et al., 2006), and this variation can affect the accuracy of biometric estimates. Also, recall that human DNA is 99.5% identical across individuals. Thus, only a small proportion of DNA varies across family members, and it is those small variations that are being estimated when accounting for genetic differences across siblings.

In general, the assumptions used in behavioral genetics designs do not appear to be grossly violated; however, it is important to remember these assumptions in reading any behavioral genetics article. The best way to understand whether a violation of assumptions greatly influences study findings is (1) to find out if results have been replicated across a variety of behavioral genetics designs (e.g., twin and adoption studies) and (2) to be mindful of the concept of gene–environment interplay. Interpretations of the heritability of traits can be interpreted in a gene–environment correlation framework; if gene–environment interactions are not specifically tested, they may be relevant in the development of that trait as well. Also, (3) it is useful to find out whether the equal-environments assumption has been examined in each sample to see how much it matters to the phenotypes of interest (e.g., Kendler & Gardner, 1998; McGue et al., 2007). Finally, (4) recall that while most psychological traits are additive in nature, other genetic effects may be relevant as well.

Quantitative Behavioral Genetics: Major Findings

Now that we have laid the basic groundwork of behavioral genetics methods, definitions, and assumptions, we review key behavioral genetics research findings and explore their implications for family researchers. The goal of this discussion is to further evaluate the sources of $a^2$, $c^2$, and $e^2$ on the particular traits (observed phenotypes, such as personality and antisocial behavior) based on the current literature. Our primary focus is on the shared environment ($c^2$), but we also cover additive genetic ($a^2$) and nonshared environmental ($e^2$) effects. We focus on the shared environment because, given that we are considered “experts” on the family, we believe that family researchers could greatly contribute to this topic by conducting research that explains specific shared environmental variants.

Traits with negligible shared environmental effects. Many behavioral genetics studies have yielded additive genetic effects in a variety of
traits, including aggression (DiLalla, 2002), personality (Bouchard & McGue, 2003), eating disorders (Bulik, Sullivan, Wade, & Kendler, 2000), feelings and attitudes (McGue, Sharma, & Benson, 1996), political attitudes and behaviors (Hatemi, Dawes, Frost-Keller, Settle, & Verhulst, 2011), cognitive abilities and psychopathology (Plomin & Rende, 1990), and even family relationship quality (Elkins, McGue, & Iacono, 1997; McGue et al., 2005). Here we focus on two phenotypes whose small shared environmental effects have been broadly and exhaustively replicated: personality and mental health disorders.

The shared environment appears to account for almost no variance in personality. Additive genetic effects explain between 39% and 58%, and nonshared environmental effects account for 40% to 56% of the univariate variance in personality (Bouchard & McGue, 2003; Tellegen et al., 1988). Confidence in these estimates is established through their replication across samples (e.g., McGue, Bouchard, Iacono, & Lykken, 1993; Plomin, Owen, & McGuffin, 1994) using various personality structures, including a three-factor model (Tellegen et al., 1988), a five-factor model (Jang, Livesley, & Vernon, 1996; Loehlin, McCrae, Costa, & John, 1998; Yamagata et al., 2006), and a general factor of personality (Rushton, Bons, & Hur, 2008). Overall, the existing literature suggests that genetic effects can explain a substantial portion of personality traits. The presence of little to no shared environment effects and substantial nonshared effects, which include measurement error, suggests that either the shared environment has little effect on personality development or that relevant shared environment components of personality are not identified due to things like gene–environment correlation or interaction.

Similar patterns of genetic and environmental estimates have been found for mental disorders (for a review, see Kendler & Prescott, 2006). For example, using a pooled sample of national registries, Lichtenstein et al. (2009) found that less than 5% of variance in schizophrenia was attributed to the shared environment, and nearly 64% was attributed to additive genetic effects. A similar pattern was found for bipolar disorder (59% additive genetic, 3% shared environment). In fact, the comorbidity between these disorders is predominately explained by additive genetic effects, which suggests a common genetic cause for both disorders (Lichtenstein et al., 2009).

Taken together, these findings imply that if we are interested in the etiology of personality and serious psychiatric disorders, then examining shared environmental effects in isolation may not be a promising approach. An alternative approach may be to examine family effects that contribute to nonshared environmental variance, because nonshared effects on personality and psychiatric disorders appear to be as substantial as additive genetic effects. However, this notion has led to few significant findings. For example, Pike, McGuire, Hetherington, Reiss, and Plomin (1996) aimed to detect nonshared environment covariance in the relationship between parenting and adolescent antisocial behavior and depression, but they found negligible effects for the environment and greater effects for genetic influence in this relationship. In their well-known meta-analysis, Turkheimer and Waldron (2000) found little evidence for replicated nonshared environmental findings. More recent approaches to studying the environment, including attention to appropriate measurement tools, the use of advanced statistical procedures, and consideration of gene–environment interaction seem especially promising (Horwitz & Neiderhiser, 2011; Loehlin, 2010; Turkheimer & Waldron, 2000). We return to the concept and methods after our discussion of main effects.

**Traits with moderate to substantial shared environmental effects.** In addition to traits with large additive genetic effects (e.g., personality), there are also traits with moderate to substantial shared environmental effects. One of the first articles to show moderate shared environmental effects on a variety of traits was by McGue et al. (1996). Significant correlations between adoptive siblings were found for the behavioral traits (e.g., externalizing, antisocial behavior, prosocial behavior), but consistent with Tellegen et al.’s (1988) findings, not for personality traits and affective traits (e.g., internalizing, negative emotionality). Remember, according to behavioral genetics assumptions, because adopted siblings are not genetically related (they are selected as such for a behavioral genetics design), any correlation between them refers to similarity that is attributed to the shared environment. The shared environment explained roughly 10% of the variance in the behavioral traits, which implies a modest shared
environmental effect. These findings were consistently found across several studies. A recent meta-analysis of shared environmental effects (Burt, 2009) found that shared environment effects ranged from 10% to 16% on behavioral disinhibition traits (e.g., conduct disorder, oppositional defiant disorder, anxiety, depression). Additive genetic and nonshared environmental contributions to these childhood disorders were still substantial, ranging from 26% to 59%. In contrast, shared environmental effects increased to nearly 30% when considering important covariates such as age and gender.

There are also moderate shared environmental effects on IQ. For example, Buchanan, McGue, Keyes, and Iacono (2009) reported estimates of 19% for shared environmental, 65% for additive genetic, and 16% for nonshared environmental influences. Demonstrating a potential specific influence of SES on this shared environmental effect, Duyme, Dumaret, and Tomkiewicz (1999) reviewed 65 files of deprived adopted children (who are not genetically related to their adoptive parents) who were given IQ tests pre- and post-placement. These adopted children were determined to be deprived because of their experience of abuse or neglect, were adopted at older ages, and had IQs of less than 86. They documented an increase in IQ of more than 10 points after children were placed with their adoptive families. This suggests that merely being placed with their new adoptive families helped explain the increase in IQ points (which suggests shared environmental influences). Moreover, IQ increased more in higher-SES families, which suggests a specific shared environmental influence of SES on IQ in these adoptive families (because SES is a family-level variable). On average, IQ increased 7 points in low-SES adoptive families, 15 points in middle-SES adoptive families, and nearly 20 points in high-SES adoptive families. These results imply that SES may at least partially explain the shared environmental effect on IQ.

Shared environment effects are in fact quite substantial for behaviors such as adolescent substance use, generally explaining 30–60% of variance in adolescent substance use (McGue, Elkins, & Iacono, 2000). In McGue et al.’s (2000) study, the additive genetic contribution to substance use in adolescence was approximately 13%, and the nonshared environmental contribution was approximately 26%, thus indicating important contributions for both. Substance-use results of this magnitude have been replicated (see Hopfer, Crowley, & Hewitt, 2003; Lynskey, Agrawal, & Health, 2010; Rende & Slomkowski, 2009) and call attention to the importance of furthering our understanding of shared environmental effects on adolescent substance use, specifically.

The measurement of phenotypes is essential to consider. For example, observer-rated behavior may be a better measure of shared environmental effects than self-reported parent-child relationship quality. Using a twin sample of toddlers, Deater-Deckard (2000) found that shared environment effects were better explained by observer ratings of children’s difficult behavior than by parent ratings. In this study, parents’ ratings of positive and negative child behavior showed genetic and nonshared environment mediation in child externalizing outcomes. In contrast, observer ratings showed shared environment mediation to those outcomes. Burt, Klahr, Rueter, McGue, and Iacono (2011) also found that observer ratings accounted for more shared environmental variance (31%) than other informant ratings, including maternal report (23%) and self-report (20%).

More Complex Analyses in Quantitative Behavioral Genetics

So far, we have determined that shared environmental effects are important, particularly for children and adolescents, and especially for outcomes and attributes such as IQ, externalizing behaviors, and substance use. Discovering the specific attributes of the shared environment that influence phenotypes is a necessary next step to further our knowledge of the etiology of these phenotypes. In other words, these findings point to something in the shared environment that may affect children who act out or explain why adolescents use substances. If we can identify the antecedent(s), these findings could help inform prevention and intervention research.

To move beyond this basic descriptive research, it is necessary to use more complex methods that examine gene–environment interplay (Horwitz & Neiderhiser, 2011; Moffitt et al., 2005; Neiderhiser, 2011). There are at least two approaches to doing this. The first decomposes the correlations between family processes and child outcomes (rGE), and the second models gene–environment interactions (G × E). Here
we review and comment on literature that has examined both approaches.

Behavioral genetics: Decomposing correlations between family processes and child outcomes. Much like the methods used in a basic biometric model of a phenotype (also known as a univariate decomposition), a bivariate or multivariate biometric decomposition can also be used via behavioral genetic designs. In this type of model, in addition to the variance of each phenotype getting decomposed into genetic and environmental contributions (described in Figure 1), the covariance between phenotypes is also decomposed into genetic and environmental contributions. Thus, these models can estimate environmental versus genetic influence in the association between two phenotypes.

Previous research using these multivariate biometric models has found that the relationship between parenting and child externalizing behavior is at least partly environmental in nature (Burt, McGue, Krueger, & Iacono, 2007; Klahr, McGue, Iacono, & Burt, 2011; Neiderhiser, Pike, Hetherington, & Reiss, 1998). For example, Burt et al. (2007) determined that roughly 19% of the total variance in delinquency was explained by the shared environment (a fairly substantial amount of variance). Burt et al. additionally tested whether parent-child relationship quality (e.g., conflict, involvement) influenced child delinquency through the shared environment using multivariate decomposition. Indeed, they found that 2.8% of the variance in delinquency was accounted for by parenting variables and that those parenting variables explained 15% of the total shared environmental effect on delinquency.

Later research has shown additional shared environmental influences on associations between the family environment and child outcomes. For example, Shelton et al. (2008) also found shared environmental influences on the association between maternal warmth and adolescent conduct problems, further replicating Burt et al.’s (2007) report. Yet the proportion of shared environmental influence explained is still rather low in these reports, which suggests that further research is needed.

It has also been argued that family environment variables are too broadly defined (i.e., the measures need more theoretical guidance), which may be one reason we see less of an environmental influence on these family environment variables in behavioral genetics research. To challenge this, Latendresse, Rose, Viken, Pulkinen, Kaprio, and Dick (2010) analyzed several measures of parenting and found some measures had greater shared environmental effects (parental discipline, relational tension) than others (autonomy granting, parental knowledge). Moreover, they found that parental knowledge and warmth alone explained 6–32% of the total shared environmental effects on adolescent drinking. These findings suggest that parent-child relationship quality may be an important contributor to shared environmental effects. However, more research is needed to understand what other aspects of the environment may contribute to shared environmental effects.

Family researchers can contribute to this unfolding of shared environmental effects by asking questions similar to the following: What are the specific characteristics of the shared environment that contribute to traits and outcomes? What is a theory of the shared environment as it relates to sibling similarity in traits such as substance use? What is it about the shared environment that matters and why?

For example, following a growing body of research, it appears (and not surprisingly so) that much more than parenting influences child and adolescent outcomes within the family framework. Siblings, in particular, appear to be key influences on adolescents. In fact, while some research has found shared environmental influences of parenting (e.g., Burt et al., 2007; Latendresse et al., 2010), the magnitude of shared environmental effects appears to be greater for characteristics of sibling relationship quality than for parent-child relationship quality (Neiderhiser, Reiss, & Hetherington, 2007). Conversely, parent-child relationship quality influences children more (than sibling relationship quality) through genetic influences (Bussell et al., 1999; Fagan & Najman, 2005; Feinberg, Neiderhiser, Howe, & Hetherington, 2001; McGue & Iacono, 2009; Pike et al., 1996).

This “genetic influence” of parenting on children’s outcomes is usually interpreted through a gene–environment correlation framework (Burt, 2011; Horwitz & Neiderhiser, 2011; Scarr & McCartney, 1983). Specifically, an adolescent’s genotype may, in some way, be associated with parental behavior. This is consistent with the bidirectional parent-child relationship often employed in family theories, such that parents
respond to their children on the basis of their unique, genetically influenced personality and behavior. This example is defined as evocative rGE. Another type of rGE is active; in this, adolescents are seeking out their own environments to match their genetic predispositions (e.g., antisocial adolescents wish to be less involved with their parents and more involved with antisocial peers). A final type of rGE is passive; in this, parents pass down their own genetic predispositions while simultaneously providing environmental influence (e.g., parents with histories of antisocial behavior transmit genes to their children and are less involved with their children).

Given the proximity of siblings’ ages, it has been argued that siblings may be important socializing influences on outcomes related specifically to substance use (Lynskey et al., 2010). For example, using a genetically informed sample (full biological siblings and adopted siblings), Samek and Rueter (2011) found that sibling similarity in substance use was not significantly different across genetically related versus unrelated pairs and that feeling close to an elder sibling reduced the younger sibling’s overall substance-use behavior 3.5 years later. This, along with other research on sibling relationship quality using quantitative behavioral genetics designs (Slomkowski, Rende, Novak, Lloyd-Richardson, & Niaura, 2005), suggests that the shared environment may largely pertain to the sibling relationship context.

**Behavioral genetics: Gene–environment interactions.** In an interaction, the strength of the correlation between the independent variable (X) and the dependent variable (Y) changes when considering the influence of a third variable, a moderator. For example, if the association between X and Y is stronger among fathers than among mothers, parent gender would be said to moderate the association between X and Y. Gene–environment interactions (G × E) are defined as genetic sensitivity to the environment (Purcell, 2002); they help explain why people who share the same environmental experience have different consequences to events. G × E modeling detects whether the contributions of genetic, shared, and nonshared environmental components on some phenotypes change as a result of including a moderator in the analysis. Although we cannot comment on all existing gene–environment interactions in this article, we discuss a few of the most replicated effects to better illustrate this definition (for an overview, see Dick, 2011; Spinath & Johnson, 2011).

The most widely known moderator of the association between genetic and environmental contributions to various traits is age. In general, genetic effects on a trait increase with time, and shared environmental effects decrease; therefore, age is considered to moderate genetic and environmental contributions by “activating” genetic risk as people get older (Burt, 2011). For example, shared environmental effects explained more variance in depression among 8- to 10-year-olds (76%) than among 11- to 17-year-olds (47%). Conversely, additive genetic effects explained less variance in depression among 8- to 10-year-olds (0%) and more among 11- to 17-year-olds (29%; Rice, Harold, & Thaper, 2002). Rice et al. (2002) found nonshared environmental influence to be the same across 8- to 10-year-olds and 11- to 17-year-olds (24%). Similarly, age moderates the relationship between genetic and environmental effects and general cognitive ability. Using six twin studies from four countries, Haworth et al. (2009) found that additive genetic effects account for 41% of the variance in general cognitive ability in childhood, 55% in adolescence, and 66% in young adulthood. A similar reversal was found for shared environment estimates: 33% in childhood, 18% in adolescence, and 16% in young adulthood. Similar results have also been found for substance use (Kendler, Schmitt, Aggen, & Prescott, 2008) and peer-group deviance (Kendler et al., 2007).

Evidence of the moderating effect of gender in gene–environment interactions is less clear. For example, regarding tobacco, alcohol, and drug use, Han, McGue, and Iacono (1999) reported higher estimates of additive genetic effects and lower estimates of shared environment for males than for females. However, these differences were not statistically significant. In general, analyzing samples with reported low substance use (females typically use fewer substances than males; younger adolescents typically use fewer substances than older adolescents and adults) usually results in higher estimates of shared environmental effects. Therefore, it seems likely that the shared environment is particularly important in explaining low levels of drug use. It also indicates that genetic effects are more important when the risk of using drugs increases (known as activator effects; Burt, 2011).
In terms of depression, Silberg, Rutter, D’Onofrio, and Eaves (2003) reported significantly stronger additive genetic effects for females and stronger shared environment effects for males. Yet Rice et al. (2002) found the opposite pattern. Kendler, Prescott, Myers, and Neale (2003) did not find any significant difference in shared environment and additive genetic effects between men and women across latent factors representing behavioral disinhibition and internalizing problems. This, along with other evidence, further suggests a limited moderating effect of gender on the relationships between $a^2$, $c^2$, and $e^2$ and internalizing and externalizing outcomes (Eley, Lichtenstein, & Stevenson, 1999; Taylor, McGue, & Iacono, 2000).

An additional moderator on male adolescent externalizing appears to be residency in urban versus rural environments (Rose, Dick, Viken, & Kaprio, 2001; Legrand, Keyes, McGue, Iacono, & Krueger, 2008). In their replication of Rose et al. (2001), Legrand et al. (2008) found that genetic effects were stronger for substance use and antisocial behavior among male adolescents living in urban environments ($a^2$ ranging from .49 to .57; $c^2$ ranging from .05 to .22), and that shared environmental effects were stronger for those in rural environments ($a^2$ ranging from .02 to .04; $c^2$ ranging from .35 to .62). In this way, living in a rural environment acts as a deactivator of genetic risk for antisocial behavior, at least for males.

Potentially relating to the moderating effect of urban versus rural environments, SES also appears to moderate the association among variance components and antisocial behavior. For example, Tuvblad, Grann, and Lichtenstein (2006) found a gene–environment interaction for adolescent antisocial behavior as a result of neighborhood socioeconomic conditions. Genetic effects were stronger for male adolescents from higher-SES neighborhoods, and shared environment effects were stronger for male adolescents from lower-SES neighborhoods. Again, however, this was the case only for male adolescents. Specifically, in less advantaged neighborhoods, additive genetic effects were estimated at 1% and shared environment effects at 69%. In more advantaged neighborhoods, additive effects were estimated at 37% and shared environment effects at 13%. It is unclear what other kinds of variables might explain shared environmental effects for female adolescents exhibiting antisocial behavior.

Peer deviance has also been shown to be an activator of genetic risk in several studies (e.g., Agrawal et al., 2010; Beaver, Gibson et al., 2009; Button, Stallings, Hyun Rhee, Boardman, & Hewitt, 2009; Dick et al., 2007; Guo, Elder, Cai, & Hamilton, 2009; Harden, Hill, Turkheimer, & Emery, 2008). For example, using a twin sample of young adult women, Agrawal et al. (2010) found that women who had more friends who used substances had a greater heritability for regular substance involvement. This effect can in part be interpreted within a gene–environment correlation framework: Young women may select friends on the basis of their similar dispositional traits (e.g., being open to using alcohol and drugs); the social environment of those friends then modifies the genetic architecture of substance-use involvement.

Finally, perceptions of parenting and parenting behaviors have acted as moderators of genetic versus environmental risk in several studies on outcomes such as personality and parent-child relationships (Krueger, South, Johnson, & Iacono, 2008; South, Krueger, Johnson, & Iacono 2008), parenting and antisocial behavior (Button et al., 2007; Feinberg, Button, Neiderhiser, Reiss, & Hetherington, 2007), parental divorce and disordered eating (Suisman, Burt, McGue, Iacono, & Klump, 2011), divorce and child problem behavior (Robbers et al., 2011), and so on. Although we cannot cover this topic extensively in this review, we provide a few examples to illustrate the importance of continued research on parenting using G × E methods.

Recall our earlier discussion on the negligible shared environmental effects on personality. Krueger et al. (2008) found that shared environmental factors became increasingly influential when adolescents perceived greater levels of parent–child conflict for the personality trait of negative emotionality, specifically. Moreover, South et al. (2008) found that the genetic versus environmental contributions to parent-child relationship quality varied as a function of their child’s personality. For example, South et al. found that genetic risk in parent involvement increased as a function of children’s general negative emotionality but that shared environmental influences on children’s positive regard for parents increased as a function of children’s
negative emotionality. Compared to the main-effects research on personality we discussed earlier, these examples illustrate the complex etiology of parenting in adolescent development and the need for continued examinations of gene–environment interplay.

Conclusion on Quantitative Behavioral Genetics Research

On the basis of these findings of moderators of genetic and environmental effects, we can make several conclusions. First, it is likely that age moderates genetic versus environmental influences on most traits; shared environmental effects diminish with age, whereas genetic effects appear to increase with age. This makes sense given that children spend less time in the shared family environment as they get older and are more likely to seek out environments that match their inherent likes and dislikes (gene–environment correlation; Scarr & McCartney, 1983). We can also conclude that some things about SES and rural versus urban living environments appear to have been identified as particularly influential shared environmental effects. This further justifies the need to better understand the relationship between SES and rural versus urban environments in an effort to identify those who are at risk for adjustment problems and to inform interventions. Moreover, we can conclude that family relationship quality, particularly among siblings, may bring an important shared environment contribution to traits such as childhood conduct problems and adolescent substance use. More research is needed overall that seeks to decompose correlations between family and child behavioral outcomes (particularly those with moderate to substantial shared environmental effects) to better understand what the shared environment really is and how it works. Continued research is needed overall in gene–environment interactions to understand when the shared environment is particularly influential. In general, we need more research that seeks to understand gene–environment interplay in the context of development (Burt, 2011; Spinath & Johnson, 2011).

To be clear, we and others (e.g., Neiderhiser, 2011; Vrieze, Iacono, & McGue, 2012) believe that there is much to be learned about gene–environment interplay in the associations between family processes and child and adolescent outcomes. The use of genetically informed samples allows family scientists to fully analyze specific shared environmental variants, and therefore to better understand them. Behavioral genetic analyses to understand gene–environment correlations and interactions are useful statistical tools available to test theories that account for genetic and environmental effects. To perform a behavioral genetics analysis, information about how children are genetically related (e.g., full biological offspring of parents) or unrelated (e.g., adopted, step, blended) to parents and siblings is needed to determine groups. If large, genetically informed samples are not available to family researchers, there are public genetically informed data sets available (e.g., Add Health; Beaver, DeLisi, Wright, & Vaughn, 2009; Beaver, Gibson et al., 2009). Regardless, it is important to at least consider how study results concerning associations between family processes and child outcomes may be mediated by genetic versus environmental mechanisms, as well as to note how study findings fit with previously published behavioral genetics research.

Switching Gears: A Brief Look at New Methods in Molecular Genetics

Although the focus of this article is on how family researchers can extend quantitative behavioral genetics research by identifying specific shared environment variants and utilizing methods of gene–environment interplay, we believe that it is crucial to briefly review and comment on molecular genetics research because of the increasing number of molecular genetics studies on psychopathology. In fact, it has been argued that the future of understanding genetic and environmental influences depends considerably on the ability to detect specific genetic variants using molecular genetic approaches (McGue, 2008).

Rather than looking at ballpark estimates of genetic effects, new methods are trying to pinpoint specific genetic influences. There are several methods for identifying genetic influences that are beyond the scope of our review. We provide some examples of these methods to begin to familiarize readers with the names and basic ideas involved. One of the earliest examples is linkage analysis for single-gene disorders (Plomin et al., 2008).
Single genes have been found to be associated with very rare disorders, such as Huntington’s disease. However, linkage analysis has been less successful in identifying genes for complex disorders (when many genes are involved), which is thought to include most psychological traits (Plomin et al., 2008). Allelic association and candidate gene studies are also commonly published, which have generally been difficult to replicate (Chabris et al., 2012; Tabor, Risch, & Myers, 2002).

Here, we focus on genome-wide association studies (GWAS; e.g., International Schizophrenia Consortium, 2009; Plomin et al., 2008), which are one of the latest developments among rapidly developing human genomic methodologies. For example, the total number of publications of GWAS findings has increased from less than 100 in 2007 to nearly 1,000 in 2011 (Hindorff et al., 2012). Our goal for this section is to provide a brief introduction to GWAS and landmark molecular G × E studies for the unfamiliar reader in family science. For a comprehensive review of genomic methods and issues related to GWAS, we encourage readers to review Vrieze et al. (2012).

**WHAT IS GWAS?**

Genome-wide association studies examine the specific genetic influence or effect of what varies within our genes: an allele. Readers may be familiar with the word *SNP* (single-nucleotide polymorphism) from brief reports on major GWAS findings. Alleles are measured by examining those SNPs. GWAS is a method for examining associations between all or most SNP markers and a phenotype of interest, sometimes in a case-control format (Corvin, Craddock, & Sullivan, 2010). Usually, more than 300,000 SNP markers are examined in association with a phenotype of interest (Hardy & Singleton, 2009).

The unit of analysis, the SNP, is overwhelmingly binary (e.g., paired adenine and guanine, cytosine, and thymine) and is what varies in the human genome. SNPs are collected through DNA assays, which denote a comprehensive but not all-inclusive set of genetic markers across the human genome. Hundreds of thousands of SNPs are compared in association with a phenotype of interest, such as alcoholism. If, after controlling for multiple testing and replication across multiple data sources, cases (e.g., alcoholics) are more likely to have a certain number of alleles on or across a certain number of genes compared to controls (e.g., nonalcoholics), we could conclude that those genetic variants are central in genetically influencing a given trait (e.g., alcoholism). Major benefits of GWAS is that it does not require an a priori hypothesis, and many genetic variants can be analyzed at once (Hardy & Singleton, 2009).

**GWAS Assumptions**

Our purpose here is not to overwhelm readers with detailed information but to briefly introduce the big picture behind GWAS research (for a detailed review, see Corvin et al., 2010).

First, it is necessary to discuss the common disease and common variant hypothesis (CDCV; Hardy & Singleton, 2009). This hypothesis predicts that common disease-causing genetic variants are found in all human populations who carry that disease. The opposing model is the common disease and rare variant hypothesis (CDRV). This hypothesis proposes that disease-related genetic variants are found in different human populations who carry the disease. If CDRV proves the case for most diseases, then GWAS will be much less valuable, because there will be different genetic causes for the same disease. Therefore, GWAS operates with the assumption of CDCV because it seeks to identify genetic variants individuals have in common in relation to outcomes.

Related to the difference in these hypotheses is the need to control for genetic ancestry (McCarthy et al., 2008). This helps reduce potential confounds of causal genetic variants across different human populations. Genetic ancestry refers to a person’s shared ancestry (not race or ethnicity) and is determined by geographic location of origin. Racial categories are typically socially constructed and have limited biological utility. Differences exist between genetic ancestries mostly because of chance or evolutionary selection. For example, East Asians are at a higher risk for an ALDH2 (enzyme) deficiency because of an inability to break down acetaldehyde into acetate when drinking alcohol. Therefore, some experience a flushing sensation and an unpleasant response to alcohol. Consequently, people with this deficiency have a reduced likelihood of alcoholism (Eng, Luczack, & Wall, 2007). An examination of these genetic variants in a study (not controlling for genetic ancestry) would reveal an effect for this genetic variant on
alcoholism but would fail to recognize that it is significant for only a subpopulation (see Ntzani, Liberopoulos, Manolio, & Ioannidis, 2012).

Finally, and as previously mentioned, GWAS requires large samples and replication as a result of restrictions on statistical power (McCarthy et al., 2008). This is because of the simultaneously testing of thousands of regressions (alleles predicting phenotypes) and the need to correct for multiple testing (to control for a Type I error, or false positive). Because of the need for large sample sizes to detect an effect, many studies are involved in consortiums to pool together data for sufficient power.

**Major GWAS Findings**

Rather than broadly review several findings, we focus on one major research finding that is generally consistent with most GWAS findings on many psychosocial traits to date. As discussed earlier, quantitative behavioral genetics research has shown large additive genetic effects for schizophrenia and bipolar disorder (approximately 60%; Lichtenstein et al., 2009). The International Schizophrenia Consortium (2009), led by S. M. Purcell, conducted a large-scale GWAS analysis (3,000 cases and matching controls without schizophrenia) and found some evidence of 74,000 SNPs in the Chromosome 6 region. Those same SNPs that predicted schizophrenia also predicted bipolar disorder, thus indicating a potential shared etiology of those diseases. In this study, however, only 3% of the variance in schizophrenia could be explained by these genetic variants alone. However, by using simulation techniques (and making certain assumptions), findings showed the variance explained may be nearly 30%. Still, these estimates are remarkably different from the heritability predicted by twin studies. This example reflects what is reported across various GWAS analyses: very small genetic effects that were previously thought to be quite large, based on quantitative behavioral genetics methods. This has been referred to as the case of the missing heritability (Maher, 2008), and newer methods continue to be developed and refined to help deal with some of the assumptions of GWAS.

Recently, there has been attention to methods such as versatile gene-based tests (Liu et al., 2010), genome-wide complex trait analysis (Yang, Lee, Goddard, & Visscher, 2011), pathway analysis (Hong, Pawitan, Magnusson, & Prince, 2009), and methods that use genetic scoring (GIANT Consortium, 2010; Vrieze et al., 2011), which seems especially promising. Genetic scoring methods take the top SNPs from a GWAS (or even better, a meta-analysis of GWAS results), weight them according to their strength of effects, and use them in subsequent analyses. The design and implementation of new methods will clearly continue in the goal of discovering specific genetic influences of various complex traits.

**G × E Using Molecular Genetic Methods**

There is also a large body of research that utilizes molecular genetics techniques in G × E (for a review, see Burt, 2011; Dick, 2011; Uher, 2011). In these types of analyses, a genetic variant is targeted on the basis of a hypothesis (e.g., the connection between serotonin or dopamine transporters) and is modeled to interact with some environmental factor to predict an outcome. Perhaps the best known of these studies is the research by Caspi and colleagues (Caspi et al., 2002; Caspi et al., 2003). In the first molecular genetic G × E report, Caspi et al. (2002) used a candidate gene approach that focused on the MAOA gene, as previous animal and human research has found that the absence of MAOA is associated with aggression in men (Brunner, Nelen, Breakefield, Ropers, & van Oost, 1993; Uher, 2011). Caspi et al. found that while the MAOA gene was not associated with antisocial behavior, those with a particular allele frequency on the MAOA gene and who had also experienced child abuse had an increased risk of adult antisocial behavior. In 2003, Caspi et al. found a similar pattern of results utilizing the SERT gene (involved in serotonin regulation), stressful life events, and depression.

These initial findings led to much excitement across a variety of fields. However, replication of Caspi’s work has been somewhat controversial. In their meta-analysis of eight studies, Taylor and Kim-Cohen (2007) found evidence for the MAOA and antisocial behavior interaction, but the stated evidence was preliminary, and more replication was needed. However, Risch et al. (2009) acquired data from 14 independent studies (N = 14,000+) and found no conclusive evidence of the interaction of SERT, child abuse, and depression. In response, Caspi, Hariri, Holmes, Huher, and Moffitt (2010)
reviewed several human observational studies, neuroscience studies, and animal research studies to provide more conclusive evidence on $G \times E$ with SERT. Uher (2011) also reviewed potential reasons for the inconclusive findings, including methodological (e.g., use of structured interviews versus self-reports of child abuse) and statistical issues that arise when analyzing categorical outcomes in $G \times E$ interactions (for more detail, see Eaves, 2006).

This illustrates the mixed and inconclusive findings in the field and underscores the need for continued replication. Particularly when it comes to $G \times E$ using molecular methods (and allelic association or candidate gene studies that test main genetic variant effects, for that matter), it is important to find out whether the results replicate before generating too much excitement, too quickly. In fact, the journal *Behavioral Genetics* has required direct replication of any candidate gene association study to be considered for publication (Hewitt, 2012). Journals with reviewers who are not as familiar with this field may be seeing an increasing amount of candidate gene studies submitted for publication (that may fail to replicate; see Duncan & Keller, 2011).

**Conclusions on New Methods in Molecular Genetics**

Overall, GWAS findings have found rather small effects on traits that we think of as largely heritable. Several factors may contribute to these findings. First, the effects themselves may actually be quite small (CDCV hypothesis), or the effects may be rare and differ across populations (CDRV hypothesis). Second, correlation is not causation; just because the SNP and phenotype are highly correlated does not mean that the SNP is causing variation in the phenotype. The identified allele may not be the causal mechanism, but rather may be just correlated with the causal genetic variant (called linkage disequilibrium). Third, the small effects may be due to an interaction between alleles and the environment (which, again, are often hard to replicate). Fourth, genetic effects may be confounded if results are not controlled for genetic ancestry. Finally, alleles may not be the appropriate level of analysis for discovering genetic variant effects—hence the need for additional methods such as versatile gene-based tests (Liu et al., 2010), genome-wide complex trait analysis (Yang et al., 2011), path analysis (Hong et al., 2009), and genetic scoring (GIANT Consortium, 2010; Vrieze et al., 2011). Moreover, research that utilizes $G \times E$ via molecular genetics methods such as candidate gene studies will continue to grow, but caution is warranted in interpreting nonreplicated findings.

Molecular genetics methods and findings on complex traits such as psychopathology are only beginning to emerge. Even though only very small effects have been detected using GWAS, this research could still potentially have a large payoff. For example, findings may be helpful in understanding the etiology of mental health disorders that profoundly affect family systems (e.g. schizophrenia and bipolar disorder; International Schizophrenia Consortium, 2009). Family practitioners may someday utilize GWAS or other molecular genetics findings to inform clinical interventions with families that promote healthy family functioning. Such interventions may be achieved through psychoeducation and discussion of research findings with families. At this point, however, we are not there yet, and we have a long way to go.

**Conclusions on Behavioral and Molecular Genetics for Family Researchers**

In summary, we believe that understanding behavioral and molecular genetics research can substantially expand our understanding of families. Moreover, through collaboration across family and behavior genetics disciplines, we can help inform a broader, more complete understanding of the complex etiology of family health and well-being. Members of our field are particularly well placed to contribute to the understanding of shared environmental effects and the impact of gene–environment interplay. The main point is this: If genes and the environment are both having effects on human development (and most of our theories tell us so), then our odds of understanding the environmental influences are better if we can understand and systematically rule out genetic influences.

**Note**

Diana R. Samek was supported by Grant No. MH017069 from the National Institute of Mental Health postdoctoral training grant.
REFERENCES


Haworth, C. M. A., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J. C., van Beijsterveldt, C. E. M., . . . Plomin, R. (2009). The heritability of general cognitive ability increases linearly from...


