Psychological Bulletin
2009, Vol. 135, No. 4, 608–637

0033-2909/09/$12.00 DOI: 10.1037/a0015702

Rethinking Environmental Contributions to Child and Adolescent Psychopathology: A Meta-Analysis of Shared Environmental Influences

S. Alexandra Burt
Michigan State University

Behavioral genetic research has concluded that the more important environmental influences result in differences between siblings (referred to as *nonshared*; e^2), whereas environmental influences that create similarities between siblings (referred to as *shared*; c^2) are indistinguishable from zero. However, there is mounting evidence that during childhood and adolescence, c^2 may make important contributions to most forms of psychopathology. The aim of the meta-analysis was to empirically confirm this hypothesis. The author examined twin and adoption studies ($n = 490$) of internalizing and externalizing psychopathology prior to adulthood. Analyses revealed that c^2 accounted for 10%–19% of the variance within conduct disorder, oppositional defiant disorder, anxiety, depression, and broad internalizing and externalizing disorders, regardless of their operationalization. When age, informant, and sex effects were considered, c^2 generally ranged from 10%–30% of the variance. Importantly, c^2 estimates did not vary across twin and adoption studies, suggesting that these estimates reflect actual environmental influences common to siblings. The only exception was attention-deficit/hyperactivity disorder, which appeared to be largely genetic (and particularly nonadditive genetic) in origin. Conceptual, methodological, and clinical implications of these findings are discussed.

Keywords: shared environment, nonshared environment, internalizing, externalizing

Supplemental materials: http://dx.doi.org/10.1037/a0015702.supp

Just over 20 years ago, Plomin and Daniels (1987) published a paradigm-shifting article in which they persuasively argued that despite clear evidence for prominent genetic influences, the environment was also critically important to psychological and behavioral outcomes. However, this environmental influence did not function as expected according to prominent socialization theories (i.e., resulting in sibling similarity). Instead, it appeared that environmental influences resulted primarily in differences between siblings (i.e., were *nonshared* or child specific; Plomin & Daniels, 1987), whereas genetic influences were almost fully responsible for sibling similarities. Their argument was based primarily on findings from behavioral genetics research, which had converged in suggesting that estimates of the nonshared environment were moderate to large across personality, cognitive abilities, and several forms of psychopathology. By contrast, they noted that environmental influences that create similarities between siblings (i.e., those that were *shared* or family-wide; c^2) had generally been found to be statistically indistinguishable from zero.

The reconceptualization of environmental influences as predominantly nonshared or child specific in origin has since been widely accepted across developmental, personality, and abnormal psychology (indeed, it has been referred to as a law of behavioral genetics; Turkheimer, 2000). Such acceptance across disciplines is particularly noteworthy, given that this notion of the environment challenged long-standing and deeply held assumptions (essentially since Freud) that parents had a family-wide or shared environmental effect on children's development. For example, the absence of shared environmental influences in adulthood led prominent behavioral geneticists to argue that to the extent that parents impact child outcomes, they do so at a child-specific rather than a family-wide level (McGue & Bouchard, 1998; Reiss et al., 1995). Others took the implications of this finding further still, arguing that because parents are the most obvious source of environmentally mediated similarities among siblings and because only nonshared environmental influences are statistically significant in adulthood, therefore parenting has minimal influence on psychological and behavioral outcomes (Harris, 1998, 2000; Rowe, 1994).[1]

Despite its widespread influence and acceptance, there is a recent and growing body of research suggesting that the theory of the environment as responsible only for sibling differences needs revamping, at least with regard to child and adolescent

I would like to thank Kelly Klump, Matt McGue, Robert Plomin, and Eric Turkheimer for their valuable and insightful comments on earlier versions of the article.

Correspondence concerning this article should be addressed to S. Alexandra Burt, Department of Psychology, Michigan State University, 107D Psychology Building, East Lansing, MI 48823. E-mail: burts@msu.edu

[1] In her 1998 book, Harris referred specifically to "personality." She has since clarified this point, noting that her group socialization theory was intended to explain all habitual patterns of behavior (and primarily social behaviors) not readily modifiable in adulthood, rather than just personality per se (Harris, 2000).

psychopathology.[2] First, the original theory was based almost exclusively on studies of personality and cognitive ability (Plomin & Daniels, 1987). The studies of psychopathology they reviewed were limited to a few studies of schizophrenia, manic-depressive psychosis, neuroses, and alcoholism (disorders that rarely manifest before adulthood). In a subsequent review on the same topic, Plomin, Chipuer, and Neiderhiser (1994) noted that "it is more difficult to draw conclusions from behavioral genetic research on psychopathology than it is in the area of personality . . . as samples are generally not large and results are thus less consistent" (p. 12). Nonetheless, after reviewing the evidence in question, they went on to conclude that available data "converge on the conclusion that environmental influence is almost exclusively nonshared for most areas of psychopathology" (Plomin et al., 1994, p. 12).[3] To their credit, the data they reviewed were largely consistent with this interpretation. Critically, however, the use of small sample sizes is a fundamental limitation of twin studies when studying shared environmental influences alongside genetic influences, particularly when the former are more modest in magnitude (Martin, Eaves, Kearsey, & Davies, 1978). For example, Martin et al. (1978) concluded that 725 to 1,233 twin pairs were necessary to reliably detect shared environmental variance contributions of 30% in the presence of nonshared environmental and genetic influences. More modest shared environmental estimates[4] require even larger samples (i.e., at least 7,000 pairs are needed to reliably detect shared environmental influences of 10%).[5] Any conclusions regarding the shared environment should thus be made cautiously (if at all) when examining small samples.

Moreover, very few twin and adoption studies available at the time were focused on psychopathology in children or adolescents in particular (as reviewed by LaBuda, Gottesman, & Pauls, 1993), a critical consideration given that heritability estimates for many phenotypes are thought to change with age (i.e., genetic influences are typically smaller in childhood and increase over time, whereas shared environmental influences are thought to evidence the opposite pattern; Bergen, Gardner, & Kendler, 2007). My own literature search for this meta-analysis yielded only eight twin studies published before 1987 that examined psychopathology (i.e., delinquency, attention problems/hyperactivity, depression, or anxiety) in child and adolescent samples. Of these, the largest sample size was 265 sibling pairs, with an average of 148 sibling pairs per study (range = 38–265). Etiological conclusions regarding child and adolescent psychopathology at the time were thus more speculative than conclusive. Indeed, a quick glance at findings from some of the more recent and very large child and adolescent twin studies (i.e., the Netherlands Twin Registry; $N > 1{,}000$ pairs) suggests that shared environmental influences may make moderate and significant contributions (i.e., 20%–30% of the total phenotypic variance, as reviewed below) to most forms of child and adolescent psychopathology.

Another salient consideration is that efforts to identify these nonshared environmental influences have largely failed, even when very large samples and methodologies specifically designed to identify nonshared environmental sources of variance were used. For example, it has been widely acknowledged that the Nonshared Environment and Adolescent Development study (a project funded with the express goal of identifying these nonshared environmental influences; Reiss, Neiderhiser, Hetherington, & Plomin, 2000) was not able to identify the nonshared environment, despite their best efforts to do so (Neiderhiser, Reiss, & Hetherington, 2007; Plomin, Asbury, & Dunn, 2001). Indeed, specific nonshared environmental factors typically account for no more than 2% of the variance in the outcome (Turkheimer & Waldron, 2000).[6] Recent theorists have thus suggested that, rather than being a function of important and identifiable environmental influences that serve to differentiate siblings, the nonshared environment is largely composed of idiosyncratic and/or transient environmental influences with little to no long-term explanatory power (Rutter, Silberg, O'Connor, & Simonoff, 1999; Turkheimer & Waldron, 2000). Consistent with this more recent conceptualization, longitudinal data suggest that nonshared environmental influences rarely persist over time. Instead, they appear to be largely specific to a given assessment period (Burt, McGue, Iacono, & Krueger, 2006; Rutter et al., 1999; Turkheimer & Waldron, 2000). Given this collective absence of tangible results for the nonshared environment, it may be time to reconsider shared environmental effects.

Indeed, shared environmental influences (though not often examined) appear to be both identifiable and persistent over time, at least prior to adulthood (Rutter et al., 1999). Several independent studies using different samples and methodological design have now suggested that the origin of the association between parental divorce and adolescent behavior problems is largely shared environmental in origin (Burt, Barnes, McGue, & Iacono, 2008; D'Onofrio et al., 2005, 2007; O'Connor, Caspi, DeFries, & Plomin, 2000), as is the association between the parent–child relationship and adolescent externalizing, at least in part (Burt, Krueger, McGue, & Iacono, 2003; Burt, McGue, Krueger, & Iacono, 2007; McGue, Sharma, & Benson, 1996; Pike, McGuire, Hetherington, Reiss, & Plomin, 1996). Moreover, these measured

[2] Note that my argument is exclusive to the emphasis on the nonshared environment as the sole source of meaningful environmental variance. The role of genetic influences, by contrast, is not contested here.

[3] The only clear exception that they noted was for antisocial behavior in high-school-age twins, though they speculated that this could be a function of twins "partnering" to commit delinquent acts.

[4] To my knowledge, there are no published guidelines for semantically describing the proportion of variance explained within the field of behavioral genetics. I thus sought to develop my own for the purpose of this review. I examined descriptors of genetic and environmental influence in roughly 30 recent articles, making an effort to include as many different authors as possible. A (somewhat) consistent pattern emerged, in which effects smaller than 20% (i.e., 5%–19%) were described as "modest" or "small"; effects ranging from 20% to 49% were described as "moderate; and effects of 50% or greater were described as "large." Although not a precise interpretative framework, it does generally conform to the (unwritten) standards of the field, and was thus applied throughout.

[5] Note that these sample size estimates depend on the proportion of monozygotic to dizygotic twin pairs. The estimates reported here assumed that the sample was split evenly across the two twin types.

[6] Recent studies have suggested that the proportion of variance accounted for by specific nonshared environmental factors may be somewhat larger when only extremely discordant siblings are examined (Asbury et al., 2003). I do not contest this conclusion (Burt, McGue, Iacono, & Krueger, 2006). Even so, it remains unclear what factors account for the prominent nonshared environmental variance observed in epidemiological samples.

environmental factors typically account for moderate proportions (e.g., 25%) of the shared environmental variance, even in population-based samples (and could be more pronounced in high-risk samples, as discussed below). Shared environmental influences thus appear to be identifiable sources of environmental variance, particularly relative to nonshared environmental influences.

Shared environmental influences also appear to persist over time (at least across childhood and adolescence), suggesting that these environmental influences are generally systematic in nature prior to adulthood. In one of the largest longitudinal twin studies to date (i.e., data at four waves were available for over 1,000 twin pairs), Bartels, van den Oord, et al. (2004) examined etiological stability and change in internalizing and externalizing spectrum problems across ages 3, 7, 10, and 12. Shared and nonshared environmental proportions of variance were significant at each age and ranged from 15% to 35% (Bartels, van den Oord, et al., 2004). Critically, however, nonshared environmental influences were largely (and, in some cases, exclusively) age specific, accounting for only 6%–10% of the stability in these behaviors. By contrast, a common set of shared environmental influences operated at all ages, accounting for 37%–43% of the stability in internalizing and externalizing problems. In short, though they may dissipate by adulthood, shared environmental influences appear to be persistent sources of individual differences in psychopathology prior to adulthood.

Brief Review of Behavioral Genetic Methodology

I next discuss the possible relevance of the shared environment to child and adolescent psychopathology. However, because this discussion involves some knowledge of the methodology used to calculate genetic and environmental influences, a brief tutorial of this material is in order. Typical behavioral genetic analyses make use of the difference in the proportion of segregating genes shared between reared-together siblings. Monozygotic (MZ) or identical twins result from a single fertilized zygote that splits in two and hence share 100% of their segregating genes. Dizygotic (DZ) or fraternal twins are the result of two independent conceptions and so, like all full siblings, share an average of 50% of their segregating genes. Half-siblings share only one of their two parents, and thus share an average of 25% of their segregating genetic material. Adoptive siblings and step-siblings do not share any segregating genetic material.

Using these differences between siblings, the variance within observed behaviors or characteristics (i.e., phenotypes) is partitioned into three of four components: additive genetic (a^2), dominant genetic (d^2), shared environment (c^2), and nonshared environment plus measurement error (e^2). Note that it is not possible to simultaneously estimate c^2 and d^2 in traditional decompositions of variance between reared-together siblings, because these parameters are estimated using the same information (i.e., the differences in sibling similarity with genetic relatedness).[7] The additive genetic component is the effect of individual genes summed over loci. If acting alone, a^2 would create MZ correlations that are double those of DZ–full-sibling correlations. Indeed, correlations would decrease linearly with decreasing genetic relatedness. Dominant genetic influences index nonadditive or gene-to-gene interactive effects, either at a single genetic locus (referred to as dominance; i.e., the interaction between dominant and recessive genes) or across multiple loci (referred to as epistasis). Because they involve interactions between genes, d^2 would yield MZ correlations that were more than twice as large as those of DZ–full-siblings. The nonshared environment is that part of the environment that differentiates members of a sibling pair, making them less similar. Nonshared environmental influences do not differ by proportion of genes shared, and thus reduce all sibling correlations proportionally to the same degree. Measurement error, which similarly acts to reduce sibling correlations, is also contained within e^2. Finally, the shared environment is that part of the environment that is common to both members of a sibling pair and acts to make siblings within a pair similar to each other. Shared effects do not differ by zygosity or proportion of segregating genes shared, and if acting alone, would make all sibling correlations similar in magnitude. Correlations between genetically unrelated but reared-together siblings (e.g., adoptive and step-siblings) function as "direct" estimates of shared environmental effects.

Gene–environment correlation (rGE; Plomin, DeFries, & Loehlin, 1977; Scarr & McCartney, 1983) is relevant when considering shared environmental influences. rGE is defined as nonrandom, genetically influenced exposure to particular environmental experiences, such that individuals elicit (i.e., evocative rGE) or select (i.e., active rGE) environmental experiences consistent with their genotype that then (presumably) go on to further activate this genotype. The most common example of an active rGE is assortative mating, in which individuals seek out and mate with others similar to themselves (i.e., those with similar leisure interests, intellectual ability, physical attractiveness, etc). To the extent that the traits influencing mate selection are genetically influenced, individuals have thus selected partners whose genotype is similar to their own. Passive rGE, by contrast, reflects the fact that the environment provided to one's biological children reflects the genetically influenced preferences or tendencies of the parent. Because parents share genes with their biological children, the child's genes are correlated with his or her environmental experiences. As an example, if conflictual parent–child relationships are in part a function of the tendency to be antisocial, and if antisocial behavior has a genetic component, then biological parents and children could share both the genes for the antisocial behavior and the corresponding tendency to be conflictual in their relationships. Accordingly, in biological families (such as those used in child-based twin designs), passive rGE can mimic shared environmental influences (Neiderhiser et al., 2004) when the origins are in fact a function of common parent–child genes. Fortunately, the presence of passive rGE confounds in shared environmental effects can be easily evaluated by comparing adoptive and biological siblings (the approach taken in the current study). Because adoptive siblings do not share genes with their adoptive family members, passive genotype–environment correlations are entirely eliminated, providing a direct estimate of shared environmental influ-

[7] The inability to simultaneously calculate c^2 and d^2 is specific to the sibling designs examined herein. Designs that also examine relatives across multiple generations can simultaneously estimate both effects because this design provides multiple sources of information about the shared environment (Rhee & Waldman, 2002).

ences. If c^2 is equivalent across adoptive and biological families, it strongly argues against the presence of passive rGE.

The Scarcity of Shared Environmental Research

As reviewed above, there is growing evidence that, at least with regard to child and adolescent psychopathology, shared environmental influences may be more important than was originally suggested by Plomin and Daniels (1987). Even so, relatively little research to date has seriously examined these shared environmental influences. There are several overlapping reasons for this, but one is simply that there are no review or meta-analytic articles that conclusively document a role for shared environmental influences prior to adulthood. The current meta-analysis aims to rectify this with regard to common psychopathological syndromes in childhood and adolescence.

Second, for most phenotypes, nonshared environmental influences are notably larger in magnitude than are shared environmental influences, even prior to adulthood. As Plomin and Asbury (2005) have stated, "the fact remains that most environmental variance affecting the development of psychological dimensions and psychiatric disorders is not shared by children growing up in the same family" (p. 225). Nonshared environmental influences thus simply appear to be a more important source of environmental variance than do shared environmental influences. Moreover, because the effects are smaller (and because genetic effects are estimated using the same information), twin studies require relatively large samples to detect even moderate shared environmental influences. These power demands may make shared environmental influences less attractive targets for twin researchers (particularly without any review or meta-analytic articles documenting their effect).

Third, because of the genes shared between children and their biological parents in twin samples, similarity between siblings could reflect either assortative mating or passive rGE "in disguise," rather than true shared environmental influences. As noted, assortative mating is thought to reflect an active rGE in which individuals seek out and mate with others phenotypically similar to themselves. To the extent that these phenotypic similarities between spouses reflect genetic similarities, assortative mating has an important implication for the study of shared environmental influences: It would increase the proportion of genes shared by DZ twins (but not MZ twins, who are already genetically identical). By doing so, it would render the heritability estimation procedures reviewed above invalid, and would serve to artifactually inflate shared environmental estimates and suppress genetic estimates. Under this scenario of assortative mating, however, we would expect shared environmental influences to persist across the lifespan (because, in this case, DZ twins share more than 50% of their segregating genes). This developmental pattern does not hold, thereby indirectly arguing against assortative mating as a widespread explanation for observed shared environmental influences.

The theory of passive rGE, by contrast, is consistent with observed changes in genetic and shared environmental influences with age (see Scarr & McCartney, 1983). Passive rGEs are thought to be prominent in childhood and nonexistent by adulthood (the same pattern thought to describe shared environmental influences). Building on this point, the final reason underlying the scarcity of shared environmental research to date may be that shared environmental influences are thought to largely dissipate by adulthood. This sort of developmental pattern could raise doubts for many researchers about the long-term importance of the shared environment. Indeed, the developmental shift from shared environmental influences to genetic influences that is thought to take place from childhood to adulthood has been conceptualized as an indication of decreasing passive rGE, combined with increasing active rGE (including, but not limited to, assortative mating), from childhood to adulthood (as nicely described by Scarr & McCartney, 1983). In short, the finding that shared environmental effects typically dissipate by adulthood, particularly when combined with the theory of passive rGE, raises significant doubts about the shared environment as a consequential source of environmental influence.

Why Is It Important to Study the Shared Environment?

Despite these concerns, I argue that the shared environment is more important than is generally assumed; moreover, it has several advantages for environmental research. First, perhaps the biggest conceptual weakness in inferring the importance of nonshared environmental effects from the magnitude of explained variance is that nonshared environmental components of variance also contain measurement error. It is very difficult to distinguish true nonshared environmental effects from measurement error in these designs. This is a well-known difficulty with studying the nonshared environment, and I will not belabor the point here. However, I note that shared environmental variance components are inherently free of this sort of unsystematic measurement error (although they are not free of systematic measurement error, such as rater biases, as discussed below).

Second, neither gene–environment interactions (G×E) nor active or evocative rGE (on the part of the child) directly contribute to shared environmental sources of variance (though, as noted, they do contain any passive rGE effects that are present).[8] As ably described by Purcell (2002), nonshared environmental components of variance contain G×E in which the child's genes interact with child-specific environmental events (because these child-specific experiences would activate the relevant gene[s] only in one sibling and in this way make even genetically identical siblings different from each other). Genetic components of variance contain G×E in which the child's genes interact with family-wide environmental events (because family-wide events that interact with the child's genes would increase the similarity of genetically identical siblings relative to other sibling types; Purcell, 2002). Similarly, because they ultimately involve genetic activation (which will necessarily be more similar in MZ twins than in DZ twins), active and evocative rGEs (in the child) typically load on genetic components of variance (Purcell, 2002). In stark contrast, provided passive rGE has been ruled out, shared environmental influences are very likely to reflect environmental main effects free from the direct influences of active and evocative rGE and G×E. As main effects are generally more replicable and straightforward than are interactions, I suggest that the shared environment may prove to be more

[8] Note that this reference to active rGE refers specifically to the child's niche-picking, and does not include parental niche-picking. As described above, assortative mating on the part of the parents would have consequences for shared environmental estimates in their children.

clinically useful than the nonshared environment. Studies of the shared environment are thus well positioned to yield meaningful conclusions regarding environmental contributions to etiology and may also be better able to inform prevention efforts.

Third, there has been some recent discussion of indeterminacy within the classical twin sibling design (Coventry & Keller, 2005; Keller & Coventry, 2005). Shared environmental and nonadditive genetic effects cannot be estimated in the same model so as to avoid problems with underidentification; one of the two parameters must always be set to zero. However, as nicely argued by Keller and Medland (2008), the presence of both nonadditive genetic effects and shared environmental effects indirectly serve to inflate additive genetic estimates at the expense of shared environmental influences. They simulated data in which additive genetic, nonadditive genetic, and shared environmental variances were equal to .40, .15, and .15, respectively. Using the classical twin design, the ACE model (and not the ADE model) was chosen as the better fitting model. Additive genetic influences were estimated at .60, whereas c^2 was estimated at .02 (Keller & Medland, 2008). In short, the simultaneous presence of shared environmental and nonadditive genetic influences (a biologically plausible scenario) would lead to overestimates of additive genetic effects and marked underestimates of shared environmental effects within the classical twin design. (Fortunately, however, as with passive rGE, the examination of adoptive siblings circumvents this underestimation of shared environmental effects.)

Next, shared environmental influences appear to be particularly pronounced in high-risk environments, a point first made by Plomin & Daniels (1987) in their original article. More recent studies have supported this claim (Burt, McGue, Demarte, Krueger, & Iacono, 2006; Cleveland, 2003; Turkheimer, Haley, Waldron, D'Onofrio, & Gottesman, 2003; Tuvblad, Grann, & Lichtenstein, 2006). For example, Cleveland (2003) found that adolescent aggression was largely shared environmental in origin within high-risk, disadvantaged neighborhoods, a pattern that did not extend to more advantaged neighborhoods. This finding of very strong shared environmental influences only in disadvantaged neighborhoods was then replicated in Tuvblad et al. (2006). Similarly, Burt, McGue, Demarte, et al. (2006) found that conduct disorder was predominantly shared environmental in origin among girls with an early timing of menarche (i.e., ages 8–11; a well-documented risk factor for antisocial behavior in girls) and largely genetic in origin for those with an average timing of menarche. The above studies collectively suggest that genetic influences may be less influential and shared environmental influences more influential in high-risk or disadvantaged environments. The specific mechanisms at work remain unclear, but one possibility is that some environmental experiences are so risky that they can elicit psychopathological outcomes even in the absence of genetic risk. In any case, such results again highlight the importance of the shared environment to child and adolescent psychopathology.

Finally, and perhaps most important, the fact that shared environmental effects are not evidenced in adulthood does not therefore imply that they are inconsequential. Indeed, this finding says little about the importance of the shared environment prior to adulthood, a particularly salient point given that child and adolescent psychopathology is a significant problem in its own right. More than 20% of children suffer from at least one common mental illness (Shaffer et al., 1996), and consequently may suffer from low educational achievement, peer difficulties, family conflict, incarceration, and a host of other deleterious outcomes (Mash & Wolf, 2005). Understanding family-level environmental forces involved in child and adolescent psychopathology is thus likely to have significant ramifications both on understanding youth outcomes and on the design and implementation of prevention and treatment interventions. This point is further underscored by the (as yet untapped) potential of the shared environment (i.e., if shared environmental influences ultimately prove to be particularly identifiable and persistent forms of environmental influence, as suggested above).

Furthermore, many forms of adult psychopathology (i.e., anxiety, depression, antisocial behavior) begin to manifest, at least in prodromal forms, prior to adulthood. By influencing the onset or early development of psychopathological conditions, the family-level environment may thus have a profound effect, even if its role in the maintenance of these conditions is less certain once the adolescent has passed into adulthood. A nice example of this phenomenon is seen in the study of intelligence. Like psychopathology, intelligence is moderately influenced by the shared environment in childhood, an effect that has fully dissipated by adulthood (McGue, Bouchard, Iacono, & Lykken, 1993). Even so, a recent meta-analysis suggested that the IQs of adoptees are considerably higher than those of their nonadopted biological siblings (i.e., siblings raised by their birth parents; effect size = 1.17; van IJzendoorn & Juffer, 2005). Such results indicate that adoption into high-functioning homes has formative (and probably shared environmentally mediated) consequences on the early development of IQ, despite the fact that shared environmental influences on IQ are no longer relevant after adulthood. Moreover, some of these shared environmental influences on childhood IQ have already been identified (i.e., socioeconomic status and parental warmth; Duyme, Dumaret, & Tomkiewicz, 1999; Petrill & Deater-Deckard, 2004). In short, the absence of shared environmental effects in adulthood need not undermine the contribution of shared environmental effects to the initial onset or childhood development of (what will eventually be) adult outcomes.

The etiological pattern observed for IQ from childhood to adulthood also has another, more subtle, implication. Because shared environmental influences on childhood IQ have been observed in both twin and adoption studies, passive rGE cannot explain the presence of shared environmental influences in childhood (i.e., because adoptive parents do not share genes with their children, passive rGE confounds are eliminated). Given this, the observed reduction in shared environmental influences on IQ from childhood to adulthood is also not likely to be a function of corresponding reductions in passive rGE. As of now, however, no other explanation of this developmental trend has been put forward. Accordingly, a new theory appears to be needed to explain the shift from shared environmental to genetic influences during development.

One possibility, mentioned here only in hypothetical terms (I know of no study explicitly examining this possibility), is that the influence of persistent shared environment factors on psychological and behavioral outcomes may become increasingly mediated by genetic processes over the course of development. In other words, persistent shared environmental influences may activate or deactivate relevant genes, a dynamic process that unfolds slowly over time and gradually results in increased genetic influence and

decreased shared environmental influence (as the latter's effects on the former are expressed). Note that this is not inconsistent with my prior assertion that the shared environment largely represents environmental main effects. Rather, I am suggesting that environmental influences that begin as main effects may ultimately serve to activate or deactivate relevant genes over time. If true, this process would act to steadily increase the correlation between MZ twins more so than between DZ twins (as the former share more genetic material), and in this way, eventually deplete shared environmental components of variance to the benefit of genetic components of variance. Of course, this theory is mere speculation, but it does bolster my point that the lack of significant shared environmental effects in adulthood need not necessarily undermine their ultimate contribution to the disorder.

Current Study

Available evidence thus suggests that shared environmental influences on child and adolescent psychopathology are potentially meaningful targets of future research, particularly as they may yield more etiologically useful information than did studies of nonshared environmental effects. Even so, the presence of shared environmental influences on child and adolescent psychopathology, although promising, is not yet conclusive, as no study to date has attempted to comprehensively and empirically document the presence and magnitude of these effects. The current series of meta-analyses attempted to do just this. I conducted seven separate meta-analyses of the psychiatric syndromes common in childhood and adolescence: (a) conduct problems (CP), (b) oppositional defiant problems (ODP), (c) attention-deficit/hyperactivity problems (ADHP), (d) anxiety (ANX), (e) depression (DEP), (f) broad internalizing difficulties (INT), and (g) broad externalizing difficulties (EXT).

Of note, I did not include substance use disorders, in part because these are seen only in adolescence and thus are not entirely in keeping with the primary goal of this project. The second reason for omitting substance use disorders was more pragmatic: Behavioral genetic studies of substance use and abuse dwarf any of the included literatures (with the possible exception of CP), and thus it was simply not practical to include this body of research in addition to the seven other literatures already under examination. I also omitted eating disorders, as there are very few relevant studies prior to adulthood. Analyses in the current study were thus confined to CP, ODP, ADHP, ANX, DEP, INT, and EXT.

Analyses were first conducted for all data to estimate overall levels of the shared environment for each phenotype. I then compared estimates of the shared environment across twin and adoption studies for each disorder. If shared environmental estimates from twin studies were not higher than those from adoption studies, it would suggest that the estimates of shared environment obtained here are not a function of passive rGE but are instead likely to reflect actual environmental influences common to siblings. I then examined a series of potential moderators of shared environmental effects. Building on prior findings that shared environmental effect estimates vary with age, I sought to compare shared environmental estimates across age. In this way, I was able to determine whether estimates of the shared environment declined with age (albeit using cross-sectional data), as is typically assumed. Next, because adolescent girls typically exhibit more internalizing symptoms whereas boys generally exhibit more externalizing behaviors, I also examined the impact of sex on shared environmental parameter estimates.

I next examined whether estimates of the shared environment varied by informant. This is important for two reasons: First, it is widely acknowledged that genetic and environmental estimates vary markedly by informant (Burt, McGue, Krueger, & Iacono, 2005b; Eaves et al., 1997). Differences in genetic influences across informant have generally been interpreted as consistent with situational specificity (i.e., each informant is exposed to a unique slice of the child's behaviors, and these behaviors may be more or less heritable in different contexts). By contrast, parent-specific shared environmental influences have been thought to represent rater bias—specifically, the inability to discriminate between one's MZ or DZ twins (Bartels et al., 2003; Bartels, Boomsma, et al., 2004).[9] It would thus be quite important to demonstrate that shared environmental effects are not confined to one informant or to parental reports alone.

Second, different informants are subject to different sources of rater bias or measurement error. Children are more likely than adult informants to be affected by increased unreliability and/or idiosyncratic interpretations of the items—a pattern that can serve to reduce twin correlations and thus increase estimates of the nonshared environment. By contrast, parents are subject to possible shared method variance or shared informant effects, as a given parent typically provides reports on both twins. As a result, higher correlations for parent reports may simply reflect the fact that one informant is reporting on two participants. Evidence of shared environmental influences across multiple informants would suggest that the shared environment is largely robust to these various informant effect possibilities.

I also examined the impact of assessment method (i.e., diagnostic interviews vs. questionnaires) on estimates of the shared environment. Diagnostic interviews are typically considered the "gold standard" for assessment of psychopathology. Even so, questionnaires are far more frequently used to assess child and adolescent behavioral and emotional problems. This decision likely stems, at least in part, from the fact that questionnaires are much easier to administer and score. However, questionnaires have several other advantages over diagnostic interviews, particularly for behavioral genetic research. Questionnaires are typically dimensional instruments (i.e., items assess both normal and abnormal permutations of the behavior). This dimensionality results in both increased statistical power to detect effects and a more normally distributed phenotype (the latter is an important statistical assumption in

[9] Other explanations are also possible, however. Rather than being generally unable to discriminate between their twins, it may be that parents have more difficulty distinguishing between their MZ twins than between their DZ twins (given that the latter can look quite different from each other, whereas the former usually do not). If true, this would serve to increase MZ correlations relative to DZ correlations and thus artifactually inflate genetic influences (see Burt et al., 2005b). Furthermore, differences in shared environmental influences across informants need not necessarily reflect rater bias, but could instead reflect situational specificity. For example, children's behavior may be more or less environmentally influenced in various contexts (e.g., peers may be more important influences on the child's behavior at school than at home).

behavioral genetic biometric modeling). Questionnaires also generally include a forced-choice response format that could increase reliability. Moreover, the disadvantage of questionnaire assessments (i.e., reduced clinical applicability) is softened by their frequent use as supplemental assessment tools in child clinical settings, which gives some questionnaires clear clinical relevance. In any case, although they are potentially important, the effect of assessment method on genetic and environmental influences has not yet been systematically evaluated (to my knowledge). The current study will thus be the first to specifically contrast parameter estimates obtained using these two assessment techniques.

In sum, the primary goal of the current series of meta-analyses was to evaluate the role of the shared environment in common disorders of childhood and adolescence: namely, CP, ODP, ADHP, DEP, ANX, and broad EXT and INT phenotypes. Positive evidence of shared environmental influences should serve to both refine psychologists' conceptualization of environmental influences and alter the trajectories of future environmental and intervention research.

Method

Search Strategy

To identify relevant journal articles, published abstracts, and dissertations, I first examined the PsycINFO and Medline databases (in June and July 2007). I combined the following phenotype search terms (*delinquency, delinquent, conduct, antisocial, oppositional, aggression, aggressive, behavior problems, externalizing, depression, depressed, depressive, mood, anxious, anxiety, phobia, internalizing, hyperactive, hyperactivity, inattentive, inattentiveness, attention*) with each of the following genetically informative study terms (*twin, twins, adoptee, adoptees, adoptive, genetic, environment*). Only articles in which abstracts clearly reported the use of a child or adolescent sample or, alternately, did not specify subject age were examined. Those clearly indicating that they used an adult sample were omitted. The reference section of each empirical and review article identified in this way was then closely examined to identify any studies that may have been missed or that were published before these databases were established. Authors of eight manuscripts provided unpublished data.

This strategy yielded a total of 490 studies across all phenotypes, of which 31 were reviews, project overviews, or meta-analyses. Inclusion criteria (i.e., age, construct requirements, inability to calculate effect size, use of other genetically informative designs) are detailed below. Using these criteria, I included 100 studies for CP, 20 for ODP, 36 for EXT, 33 for DEP, 43 for ANX, 29 for INT, and 74 for ADHP (note that some studies provide information for more than one phenotype [e.g., Eaves et al., 1997] and are thus indicated in the count for each; this is appropriate, because phenotypes will be analyzed independently). These twin and adoption studies are listed separately by phenotype, along with effect sizes, construct operationalization, sample age, sample sex, informant, and number of pairs by zygosity or familial relationship, in supplementary online material. The supplementary online material also includes information on study inclusion related to nonindependence (an issue addressed at length below). After accounting for nonindependence, I ultimately included 38 samples for CP, 9 for ODP, 16 for EXT, 17 for DEP, 23 for ANX, 17 for INT, and 26 for ADHP.

Inclusion Criteria

Age. As the current study aimed to clarify the role of the shared environment in child and adolescent psychopathology, participants were required to be younger than age 18. In those samples that included some individuals older than age 17, at least 50% of the sample was required to be younger than 18. This criterion was more typically addressed early in the study selection process via abstract examination (i.e., only those articles clearly using child or adolescent samples or those that did not specify participant age were examined). Even so, 73 of the 490 identified studies included primarily or exclusively adult samples. These were omitted.

Construct validity and related issues. The current study focused on phenotypes related to child and adolescent psychopathology. Included studies met at least one of the following criteria: (a) The study clearly examined child or adolescent psychopathology (i.e., *DSM–III–R* and *DSM–IV* diagnoses, or *DSM–III–R* and *DSM–IV* symptom counts); (b) the measure successfully discriminates clinical and normative samples (i.e., scales on the Achenbach Child Behavior Checklist or Teacher Report Form; see Achenbach & Rescorla, 2001); and/or (c) the measure was associated with a validated measure of the phenotype. The application of these inclusion criteria is presented below, separately by phenotype. Studies that measured obliquely related constructs, such as temperament, personality, self-concept, overall psychopathology, measures peripheral to the phenotype in question (e.g., social aggression as an indicator of CP), specific behavioral acts with a relatively loose association to the phenotype (e.g., number of Bobo doll hits as a measure of aggression or CP), or neuropsychological measures of underlying brain functioning (e.g., assessments of impulsivity as proxies for ADHP) were omitted ($n = 40$ studies).

CP was assessed via the following: *DSM–III–R* and *DSM–IV* diagnoses or symptom counts of conduct disorder and antisocial personality disorder (the latter only in late adolescence), as well as more general measures of antisocial behavior, delinquency, physical aggression, and/or rule-breaking. The general measures were typically questionnaires such as the Aggression and Rule-Breaking scales on the Achenbach family of instruments (i.e., the Child Behavior Checklist, the Teacher Report Form, the Youth Self-Report; Achenbach & Rescorla, 2001), the Delinquent Behavior Inventory (Farrington & West, 1971; Gibson, 1967), or the Conduct Problems scale on the Strengths and Difficulties Questionnaire (Goodman, 1997), among others. Though constructed using items from the Multidimensional Personality Questionnaire (Patrick, Curtin, & Tellegen, 2002), the Impulsive Antisociality scale developed by Blonigen, Hicks, Krueger, Patrick, and Iacono (2005) was included, as items tapped specific antisocial acts and symptoms of conduct disorder and antisocial personality disorder (Blonigen et al., 2005).

ODP was assessed using either *DSM–III–R* and *DSM–IV* diagnoses or symptom counts of oppositional defiant disorder or oppositional and defiant behaviors as measured more generally via the Child Behavior Checklist, the Rutter A or B scales (Rutter, Tizard, & Whitmore, 1970), the Conners Rating Scales (Conners, 2001), and the Olweus questionnaires (Olweus, 1989).

ADHP was assessed using one of the following: *DSM–III–R* diagnoses or symptom counts of attention deficit disorder, *DSM–IV* diagnoses or symptom counts of attention-deficit/ hyperactivity disorder (inattentive subtype, hyperactive subtype, and/or

combined subtype), and hyperactivity, inattention, and/or attention problems (as measured more generally). The more general indices were assessed via questionnaires such as the Child Behavior Checklist, the Rutter A or B scales, the Conners Rating Scales, the DuPaul (DuPaul, 1981), and the Strengths and Difficulties Questionnaire.

EXT was indexed via various combinations of *DSM–III–R* and *DSM–IV* diagnoses or symptoms counts of conduct disorder, oppositional defiant disorder, and attention-deficit disorder or attention-deficit/hyperactivity disorder, as well as more general measures of externalizing behaviors. The general measures were almost exclusively assessed using the Achenbach family of instruments. Notably, however, the Achenbach EXT scale for ages 6 to 18 is essentially a composite of the Achenbach Aggression and Rule-breaking scales, and thus in many ways simply represents an extension of CP. To maintain consistency with the literature, in which the Achenbach EXT scale is often considered separately from other measures of CP, I elected to retain the Achenbach EXT studies within the EXT data group for the primary analyses. However, to evaluate the robustness of the results, these studies were removed from the EXT data group and added to the CP data group for additional EXT and CP analyses.

DEP was indexed using *DSM–III–R* and *DSM–IV* diagnoses or symptom counts of major depressive disorder and depressive symptoms (as measured dimensionally via the Child Depression Inventory, Kovacs, 1992, and the Mood and Feelings Questionnaire, Costello & Angold, 1988, among others).

ANX was assessed via *DSM–III–R* and *DSM–IV* diagnoses or symptom counts of separation anxiety disorder, overanxious disorder, generalized anxiety disorder, social phobia, or panic disorder, as well as more general measures of social anxiety, phobias, and general anxiety. The more general measures were assessed using questionnaires such as the Revised Children's Manifest Anxiety Schedule (Reynolds & Richmond, 1978), the Child Behavior Checklist, and the Strengths and Difficulties Questionnaire. Of note however, the scales on the Child Behavior Checklist and the Strengths and Difficulties Questionnaire (i.e., Anxious Depressed and Emotional Problems, respectively) include items indexing both DEP and ANX. This overlap can be partially understood in the context of symptom overlap across major depressive disorder and generalized anxiety disorder (e.g., difficulty falling or staying asleep is a core diagnostic symptom of both disorders). Moreover, the disorders are highly comorbid, such that individuals affected by one are affected by the other 58%–80% of the time (Judd et al., 1998; Kessler et al., 2005). However, inspection of the item content within these scales revealed that most items (e.g., 10 of the 15 items on the Teacher Report Form and 4 of the 5 items on the Strengths and Difficulties Questionnaire) appeared to be specifically tapping ANX. For instance, items on the Strengths and Difficulties Questionnaire include the following: nervous or clingy in new situations, often has headaches/stomachaches, many fears, many worries, and often unhappy. After some deliberation, I elected to include these studies within ANX. However, to evaluate the robustness of the results, these studies were removed from the ANX data group for a series of additional ANX analyses.

INT was assessed via combinations of diagnoses or symptoms counts of major depressive disorder and anxiety disorders, as well as more general measures of combined mood and anxiety difficulties. INT was indexed almost exclusively via questionnaires (including ratings). Only one study (Blonigen et al., 2005) made use of diagnostic symptoms (combining *DSM–III–R* symptom counts of major depressive disorder, social phobia, and simple phobia). As with EXT, the vast majority of studies employed the Achenbach family of instruments (though unlike with EXT, there is little question regarding the affiliation of these items, as the Achenbach INT scale is a composite of the Anxious Depressed scale, the Withdrawn scale, and the Somatic Complaints scale, and thus clearly belongs within INT).

Use of other genetically informative designs. Because the current meta-analysis aimed to estimate shared environmental influences in particular, analyses were restricted to either (a) twin-sibling studies, or (b) adoptive-sibling studies. These sorts of sibling-based designs, and particularly the adoptive-sibling studies, are ideal for the purposes of the present study. As noted in the introduction, adoptive siblings share no segregating genetic material, and thus any similarity between them must index shared environmental effects. Accordingly, they function as a direct estimate of shared environmental influences (and, perhaps more importantly, allow us evaluate the shared environment unconfounded by passive rGE). Twin designs are also common methodological tools for estimating shared environmental influences as the inclusion of genetically identical twins offers a great deal of statistical power for estimating genetic influences, the other source of biological sibling similarity. Any shared environmental effect present in twin studies should thus be unconfounded by genetic influences. As indicated in the Introduction, however, twin designs do have some limitations with regard to the detection of shared environmental effects. First, passive rGE can masquerade as shared environmental effects in twin designs. Another issue is that because twin siblings are necessarily the same age, they may share environments (e.g., classrooms) to a greater extent than other siblings, thereby casting some doubt on the generalizability of shared environmental estimates obtained via twin studies. The comparison of c^2 estimates across twin and adoptive sibling designs addresses both of these limitations, however (i.e., equivalent estimates would suggest that neither issue is influencing estimates of c^2). Finally, twin designs have limited power to simultaneously identify genetic and shared environment influences, at least in smaller samples. Fortunately, current twin studies are typically quite large (most contain at least 700 twin pairs, and some contain thousands of pairs; e.g., the Netherlands Twin Registry), and thus have more than enough power to detect both genetic and shared environment influences.

Adoption studies that compute associations between adoptive mothers and their adoptive children ($n = 3$) were omitted. Although any similarity between adoptive mothers and their adoptive children must also index the shared environment, the interpretation of this estimate is less clear than that between siblings. Mothers and their children were de facto exposed to different families of origin, different generational or cohort influences (growing up in the 1960s instead of the 1990s), and so forth, making interpretation of these effects unclear. Next, studies examining MZ twin differences ($n = 25$) were omitted here. These designs involve the calculation of difference scores between MZ twins and thus do not allow estimation of shared environmental influences (indeed, they specifically control for genetic and shared environment influences so as to directly examine nonshared environmental effects).

Studies of twins reared apart were also ineligible for inclusion, as they do not independently estimate shared environmental influences. No such studies were selected during the literature review, however, as their samples were composed of adults. Finally, children-of-twins designs are a recent and exciting tool for tracking the influence of specific environmental factors on youth development (e.g., divorce; D'Onofrio et al., 2005 2007). However, because the children of the respective twins rarely grow up in the same household (they are cousins), latent shared environmental influences can be estimated only in the twin parents, rather than in their children. These samples were thus ineligible for inclusion (as with twin-reared-apart studies, they were excluded during the literature review).

Inability to compute study effect sizes. The study effect sizes used in this meta-analysis were intraclass correlations or tetrachoric correlations (though the latter were included only a few times). Heritability estimates (Li, Cheng, Ma, & Swan, 2003) were not used since earlier studies (with smaller samples) typically constrained shared environmental parameters to be zero. Of the 490 studies, 126 were excluded because effects sizes were not reported and I was unable to calculate effect sizes. In most such cases ($n = 89$ studies), the studies in question were focusing on phenotypic level analyses and were not examining heritability coefficients (e.g., Tully, Arseneault, Caspi, Moffitt, & Morgan, 2004, examined whether maternal warmth moderated the effect of birth weight on observed symptoms of ADHD symptoms). Such studies were not relevant to the current meta-analysis. For 36 of the 37 remaining studies, effect sizes were available in other publications. Accordingly, there was only one independent study (Murray & Sines, 1996) for which effect sizes were not reported and could not be calculated. Murray and Sines (1996) evaluated the origins of maternal reports of depression in 364 twin pairs, ages 4–12. Shared environmental influences did not contribute significantly to depression in these data (c^2 was estimated to be 0% in boys and 19% in girls).

Nonindependent samples. The final justification for study exclusion from the meta-analysis was nonindependent sampling. Sample effect sizes were deemed nonindependent for several reasons. Many authors examined more than one dependent measure of the phenotype in the same sample, either within publications (e.g., Eaves et al., 1997) or across multiple publications (e.g., other publications using the Virginia Twin Study of Adolescent Behavioral Development dataset). This could take the form of questionnaire and diagnostic interview data examined separately, data from multiple informants examined separately, and/or data for more than one questionnaire examined separately. Several publications also sampled longitudinal follow-up data on the same set of subjects (e.g., Burt et al., 2005a).

Experts recommend several options for dealing with nonindependent samples. These include averaging effect sizes of the different dependent measures, selecting one measure (presumably the best measure using the largest sample) and omitting the others, or conducting separate meta-analyses (Lipsey & Wilson, 2001). Faced with these same choices in their excellent meta-analysis of the heritability of antisocial behavior, Rhee and Waldman (2002) chose to average the effect sizes when the samples in question were identical in size. If they were not identical, they chose the effect size from the largest sample (Rhee & Waldman, 2002).

Though these choices were appropriate for Rhee and Waldman (2002), given that most samples in their meta-analysis were adults, they are less appropriate here. In particular, longitudinal samples that contain multiple informants and measures are relatively common in these data. Because of attrition, the intake (or youngest) sample is typically the largest. Further, because questionnaires are less labor-intensive to collect than are diagnostic interviews and can be completed at home if necessary, they are typically better represented in the data. Finally, because mothers are more likely than fathers to attend the testing session, and are more reliable than their children as informants, maternal reports are typically available on all or almost all participants, whereas other informant reports are not. Should c^2 estimates vary by informant, age or measure, simply analyzing the largest sample would be quite problematic.

I thus made use of the following strategy: when nonindependent samples varied across age, informant-report, and/or dependent measure, I made use of weighted averages to compute the study effect size (i.e., the sample size is used to weight the contribution of a given effect size to the average effect size). This allowed me to accommodate different sample sizes without biasing results by the consistent selection of maternal reports, young age, and questionnaire. If nonindependent samples contained multiple dependent measures but did not vary by sample size, simple averages were computed. If nonindependent samples did not vary by age, informant report, or dependent measure, however, the largest sample was chosen, albeit with two caveats. Given that sex is a potential moderator, I placed a value on studies where samples were analyzed separately by sex. Thus, if sample sizes were equal, I included the one with more information on sex. I also placed a value on studies that reported results using continuous measures versus dichotomous measures (i.e., diagnostic symptom counts versus diagnoses), as considerable power and information are lost when sub- and suprathreshold variation in diagnostic status is collapsed into a dichotomous diagnostic variable (Krueger & Finger, 2001). Thus, if sample sizes were equal, I included the intraclass correlation effect sizes rather than the tetrachoric correlation effect size. Finally, like Rhee and Waldman (2002), when I could not determine whether samples were independent (i.e., a description of the sample was not reported; Parker, 1989), I assumed independence. The results of this strategy are detailed in the supplementary on-line material.

Analyses: Theoretical and Methodological Overview

The behavioral genetic analyses employed here make use of the difference in the proportion of segregating genes shared between reared-together siblings. Using these differences, the variance within observed behaviors or characteristics (i.e., phenotypes) is partitioned into three of four components: additive genetic (a^2), dominant genetic (d^2), shared environment (c^2), and nonshared environment plus measurement error (e^2). These variance components are defined at some length in the introduction. Crucial to twin methodology is the equal environments assumption, which assumes that MZ twin pairs are no more likely to share the environmental factors that are etiologically relevant to the phenotype under study than are DZ twin pairs. Under this assumption, any differences in MZ and DZ correlations are due to differences in their genetic similarity. The equal environments assumption has

been repeatedly tested and found to be valid for numerous phenotypes, including many mental disorders (as reviewed in Plomin, DeFries, McClearn, & McGruffin, 2008), but it remains an assumption for any particular phenotype until subjected to empirical testing. Adoption studies, by contrast, are susceptible to environmental range restriction, because adoptive parents are typically better educated, more affluent, and perhaps less prone to psychopathology. However, a recent study of the impact of range restriction showed that the range restriction that was present in adoptive families had no effect on adoptive-sibling similarity for several adolescent outcomes, including delinquency, drug use, or IQ (McGue et al., 2007).

One common approach to testing causal influences within the field of behavioral genetics is to fit a series of alternative biometric models, comparing them on their fit to the observed data. In the present meta-analyses, the ACE and ADE models were compared. I did not fit the AE or CE models, as such reduced models serve to artifactually tighten confidence intervals on the remaining parameters (because fewer parameters are being estimated). Furthermore, although nonsignificant parameters are not greater than zero in statistical terms, it is not the case that they are always estimated to be zero, and it is worthwhile to know what these estimates might be, given the goal of the current meta-analyses. Note that, like the two-stage structural equation modeling method (Cheung & Chan, 2005), structural equation modeling is used both to effectively synthesize the observed correlation matrices and fit the proposed models, as outlined in detail by Rhee and Waldman (2002).

Mx, a structural-equation modeling program developed by Neale (1997), was used to perform the model-fitting analyses. Mx uses maximum-likelihood model-fitting techniques to fit models to the observed correlation matrices (as done in Rhee & Waldman, 2002; see Appendix A for an example script). The chi-square test statistic provides a goodness-of-fit index of the model to the observed correlation matrices. These chi-square values are then converted to the Akaike information criterion [AIC; AIC = χ^2 − (2 × *df*); Akaike, 1987]. AIC measures model fit relative to parsimony and is the most commonly employed fit index within the field of behavioral genetics. AIC is used to determine the best-fitting model among a set of fitted models, with the lowest (or most negative) AIC considered the best.

Mx is a particularly useful program for the current meta-analysis. Mx allows for the computation of 95% confidence intervals (CIs) for all proportions of variance estimated in the model. CIs allow researchers to determine whether a specific variance estimate is significantly greater than zero (i.e., if the CI does not overlap with zero, then the variance estimate is statistically significant). I could also constrain various variance parameter estimates to be equal or allow them to vary freely; in this way, I could statistically test whether they were distinguishable (e.g., across different values of the moderator). A significant change in the chi-square goodness-of-fit between the constrained and unconstrained models indicates that the estimates are significantly different from each other.

Order of analyses. I first estimated overall levels of c^2 for each phenotype, comparing the fit of the ACE and ADE models. The better-fitting model, as indicated by a lower AIC, is presented and discussed. Second, I compared overall estimates of c^2 across twin and adoption studies. If c^2 estimates are equivalent in adoptive and twin data, it would suggest that the estimates of c^2 obtained here are not a function of passive rGE but are instead likely to reflect actual environmental influences common to siblings. Finally, I examined whether a series of possible moderators impacted estimates of c^2. As noted, these moderators included age (ages 1–5, ages 6–10, ages 11–18), sex (male–male sibling pairs vs. female–female sibling pairs), assessment method (diagnostic interview vs. questionnaire), and informant (mother, father, teacher, child, observer, and peer reports).

When sex was examined as a potential moderator, analyses were restricted to those studies in which correlations were presented separately by sex. Opposite-sex pairs were omitted for the sex moderation analyses, thereby allowing me to compute and compare estimates of c^2 separately across boys and girls. When examining age as a potential moderator, I omitted those studies that spanned multiple age categories (ages 1–5, 6–10, and 11–18). For example, participants in the Virginia Twin Study of Adolescent Behavioral Development ranged from ages 8 to 16, and thus fell in both the 6- to 10-year-old and 11- to 18-year-old ranges. As their age categorization was unclear, these participants were omitted from the age-moderation analyses. Correlations within a sample that fell cleanly into multiple age categories were included. For instance, publications from the Netherlands Twin Registry often reported correlations at ages 3, 5, 7, 10, and 12. In such cases, I created and analyzed weighted averages for ages 3 and 5, for ages 7 and 10, and for age 12. If samples in the same age category were from multiple waves of data collection (e.g., 10–15 at Time 1 and 12–17 at Time 2), I used the largest sample available. When examining informant effects, I restricted analyses to specific informants (e.g., mother, father, teacher, child, peer, and observer ratings). Maternal reports also included those reports under the more ambiguous term of "parent," as careful reading of methods sections revealed that informants for parent reports were almost always mothers.

Results

Stem and Leaf Plots

Stem and leaf plots are presented in Tables 1 and 2 for externalizing and internalizing disorders, respectively. These plots allow for an informal estimation of possible shared environmental influences. If the median correlation is less than twice as large for MZ twins as for DZ twins, and the median correlation for adoptive siblings is greater than zero, shared environmental influences are likely to contribute to the phenotype. The plots also allow for an informal estimation of dominant genetic influences. If the median correlation is more than twice as large for MZ twins as for DZ twins, d^2 is likely to contribute to the phenotype.

As seen in Table 1, the median MZ correlation for CP was .66, whereas the median DZ correlation was .40, results that suggest the presence of modest shared environmental influences on CP. This interpretation is bolstered by the median correlation for genetically unrelated siblings, which was .14. The median MZ correlation for ODP was .63, whereas that for DZ twins was .33, results that are less consistent with shared environmental influences. EXT seemed likely to be influenced by the shared environment, as the median MZ correlation was .70, whereas the median DZ correlation was .44. Moreover, correlations among genetically unrelated siblings were uniformly greater than zero, with a median

Table 1
Stem and Leaf Plot of Effect Sizes (Correlations) for Externalizing Disorders (Overall Data)

	Leaf		
Stem	MZ twin pairs ($r = a^2 + c^2$)	DZ twin pairs/FS pairs ($r = .5a^2 + c^2$)	Unrelated sibling pairs ($r = c^2$)
Conduct problems			
.9			
.8	0348999		
.7	0111234667899		
.6	1122234556666677	12667	
.5	234778999	00011112266668	
.4	5589	001122223334445556	7
.3		00000113334455689	
.2		0114445567799	6
.1		889	3347
.0			9
−.0		2	
−.1			
Oppositional defiant problems			
.9			
.8	25		
.7	377		
.6	134		
.5	238	02	
.4		025569	
.3	78	1336	
.2		4	
.1		0289	
.0		4	
−.0		3	
−.1			
Attention−deficit/hyperactivity problems			
.9	0		
.8	134578		
.7	1234689		
.6	001123455566789		
.5	15778	37	
.4	9	1455678	
.3		122334578	
.2		00001344566799	
.1		011223333345677889	4
.0		2456999	99
−.0		2	
−.1			3
Externalizing			
.9	4		
.8	35789		
.7	0012589		
.6	11556688	378	
.5	46	23555567	
.4		013334446888	
.3		01279	7
.2		668	0
.1		1	1459
.0			
−.0			
−.1		6	

Note. MZ and DZ indicate monozygotic and dizygotic twins, respectively. FS indicates full siblings. Unrelated sibling pairs include adoptive siblings and step-siblings.

correlation of .17. By contrast, the median MZ correlation for ADHP was .66 (although they went as high as .90), whereas the median DZ correlation was .20, results that clearly argue against shared environmental influences and for dominant genetic influences. However, the median correlation for genetically unrelated siblings (.09) may suggest some shared environmental influence.

Informal stem and leaf plot comparisons for the internalizing disorders were comparable to those for EXT, CP, and ODP. The

Table 2
Stem and Leaf Plot of Effect Sizes (Correlations) for Internalizing Disorders (Overall Data)

	Leaf		
Stem	MZ twin pairs ($r = a^2 + c^2$)	DZ twin pairs/FS pairs ($r = .5a^2 + c^2$)	Unrelated sibling pairs ($r = c^2$)
Anxiety			
.8	14		
.7	3458		
.6	222444669	3448	
.5	067889	166	2
.4	123338	123447	
.3	0157889	01122334455678999	3
.2		001112234799	03
.1		058	4
.0		02348	
−.0		4	
−.1		4	
−.2		6	
Depression			
.8			
.7	024		
.6	11223347		
.5	2457	344	
.4	023	22555889	4
.3	01559	0334559	0
.2	9	01122236666678	1
.1		044456678	66
.0		18	
−.0			
−.1			
Internalizing			
.8			
.7	0222455667788999	23	
.6	14799	011245	
.5	025	0011333667	
.4	89	134558	14
.3		0126	
.2	9	56689	8
.1		244569	9
.0			6
−.0			
−.1			

Note. MZ and DZ indicate monozygotic and dizygotic twins, respectively. FS indicates full siblings. Unrelated sibling pairs include adoptive siblings and step-siblings.

median MZ correlation for ANX was .58, whereas the median DZ correlation was .32 (but as high as .68). Moreover, the median genetically unrelated sibling correlation for ANX was .23. Such results collectively highlight the likelihood of shared environmental contributions to ANX. An examination of depression yielded a similar conclusion. The median MZ correlation was .56, whereas the median DZ correlation was .26 (but as high as .54). However, the median correlation between genetically unrelated siblings was .21. Finally, the median MZ correlation for INT was .72, whereas the median DZ correlation was .45. The median correlation between genetically unrelated siblings was .28. These results imply that shared environmental influences also contribute meaningfully to INT.

Overall Analyses

The current discussion centers almost exclusively on the c^2 estimates. However, as some readers may be interested in other aspects of the results (e.g., genetic influences), full results for all fitted models are presented in Appendix B. For the overall analyses, I compared the fit of the ACE and ADE models for each phenotype. Model fit indices are presented in Table 3. As seen there, the ACE model provided the better fit to the data, as indicated by the smaller AIC value, for all phenotypes save ADHP. The ADE model clearly provided the better fit to the ADHP data.

Parameter estimates for the better fitting models are presented in Table 4. Additive genetic influences were, as expected, moderate to large in magnitude (i.e., 44%–59%) for virtually all childhood disorders, though they accounted for somewhat less variance in ADHP (26%). In the latter case, however, nonadditive genetic influences contributed the single largest proportion of variance, indicating that ADHP is largely genetic in origin (i.e., broad genetic influences were estimated at 70% of the variance). Nonshared environmental influences were also moderate in magnitude, contributing 26% to 42% of the phenotypic variance. Most impor-

Table 3
Fit Indices for Overall Model by Phenotype

Phenotype	χ^2	*df*	AIC
Externalizing			
ACE	**572.36**	**61**	**450.36**
ADE	644.88	61	522.88
Conduct problems			
ACE	**1818.78**	**126**	**1566.78**
ADE	1956.55	126	1704.55
Oppositional defiant problems			
ACE	**929.93**	**30**	**869.93**
ADE	953.59	30	893.59
Attention-deficit/hyperactivity problems			
ACE	1466.12	97	1272.12
ADE	**1246.44**	**97**	**1052.44**
Internalizing			
ACE	**576.47**	**70**	**436.47**
ADE	657.65	70	517.65
Depression			
ACE	**967.76**	**74**	**819.76**
ADE	1051.30	74	903.30
Anxiety			
ACE	**903.02**	**91**	**721.02**
ADE	962.48	91	780.48

Note. The model highlighted in bolded font provided the better fit to the data, as indicated by a lower Akaike information criterion (AIC) value.

tant, however, shared environmental influences were uniformly significant across EXT, CP, ODP, INT, DEP, and ANX, accounting for a minimum of 10% of the total variance in these disorders. Such results are clearly consistent with the primary hypothesis of the current meta-analysis. Given that shared environmental influences were estimated to be zero for ADHP, this phenotype was omitted from all subsequent analyses (although these estimates are presented in Appendix B).

Comparison of c^2 Across Twin and Adoption Studies

Overall estimates of c^2 were next compared across twin and adoption studies, thereby allowing for a direct test of the role of passive rGE. If c^2 estimates are larger in twin studies compared to adoption studies, it can be concluded that the estimates of c^2 are likely to be a function of passive rGE. Equivalent estimates, by contrast, argue against passive rGE and suggest that c^2 reflects actual environmental influences that serve to increase sibling similarity. Results are presented in Figure 1. There were no adoption studies of ODP, and thus ODP was omitted from these analyses. Estimates of c^2 from the adoption studies were significantly greater than zero for all phenotypes except DEP (though only one adoption study was available for DEP). Moreover, estimates of c^2 are equivalent across twin and adoption studies for all but one phenotype ($\Delta\chi^2 < 0.2$ on 1 *df*, *ns*, for EXT, CP, INT, and DEP). The only exception, ANX, yielded c^2 estimates that were larger in adoption studies than in twin studies (though this may be a function of imprecision in the adoption study estimates, given that these analyses were based on only 502 sibling pairs). In any case, estimates of c^2 were never larger in twin studies as compared to adoption studies. Such results strongly argue against passive rGE confounds in these estimates of shared environmental influences.

Impact of Potential Moderators on Estimates of c^2

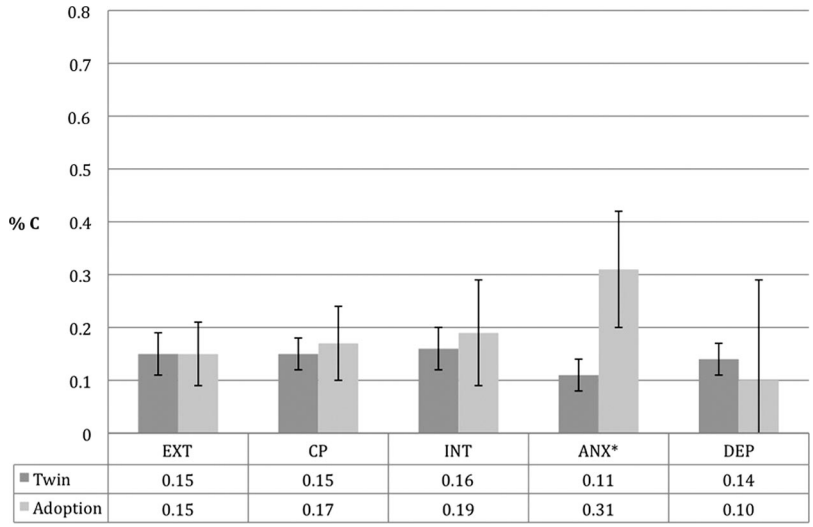
Sex. Estimates of c^2 were next computed separately for male–male and female–female sibling pairs. As seen in Figure 2, estimates of c^2 did not vary significantly across sex for any phenotype save INT (INT: $\Delta\chi^2 = 7.50$ on 1 *df*, $p < .05$). Further, although it did not significantly differ from that of the girls, the c^2 estimate for ANX in boys was not significantly greater than zero. Such results collectively indicate that shared environmental influences generally do not meaningfully vary across sex. Also of note, the proportions of variance in EXT and CP accounted for by c^2 (roughly 20% and 18% on average, respectively) were somewhat larger than those reported in the overall analyses, perhaps suggesting that either the act of collapsing across sex and/or the inclusion of opposite-sex DZ pairs serves to decrease the proportion of shared environmental influence for these phenotypes in particular. This effect was not observed for ODP, INT, ANX, or DEP.

Informant. Estimates of c^2 were also computed separately by informant. Note that one teacher study (Towers et al., 2000) reported very unusual effect sizes not in keeping with those typically observed in sibling studies (e.g., MZ correlations as low as .12, as well as several negative correlations; see supplementary online materials). These outlying data exerted an undue influence on the teacher results (despite the small sample size of 373 pairs) and were thus omitted from the teacher analyses. Results for the four most frequent informants (i.e., mother, father, teacher, and child self-report) are presented in Figure 3. Importantly, significant

Table 4
Parameter Estimates From Better-Fitting Overall Model by Phenotype

Phenotype	a^2	c^2	d^2	e^2
EXT (N = 16 samples with 10,957 sibling pairs)	.590 (.552, .629)	.153 (.118, .187)		.258 (.248, .269)
CP (N = 38 samples with 28,709 sibling pairs)	.576 (.550, .602)	.145 (.121, .169)		.280 (.273, .287)
ODP (N = 9 samples with 12,692 sibling pairs)	.591 (.547, .636)	.101 (.062, .140)		.308 (.297, .319)
ADHP (N = 26 samples with 25,712 sibling pairs)[a]	.259 (.198, .320)		.444 (.383, .505)	.297 (.289, .305)
INT (N = 17 samples with 13,099 sibling pairs)	.507 (.467, .547)	.164 (.129, .198)		.330 (.318, .343)
DEP (N = 17 samples with 21,027 sibling pairs)	.437 (.400, .474)	.139 (.110, .169)		.424 (.411, .438)
ANX (N = 23 samples with 20,786 sibling pairs)	.475 (.438, .512)	.122 (.091, .153)		.404 (.392, .416)

Note. a^2 = additive genetic influences; c^2 = shared environmental influences; d^2 = dominant genetic influences; e^2 = nonshared environmental influences; EXT = externalizing; CP = conduct problems; ODP = oppositional defiant problems; ADHP = attention-deficit/hyperactivity problems; INT = internalizing; DEP = depression; ANX = anxiety.
[a] ADE model is presented.



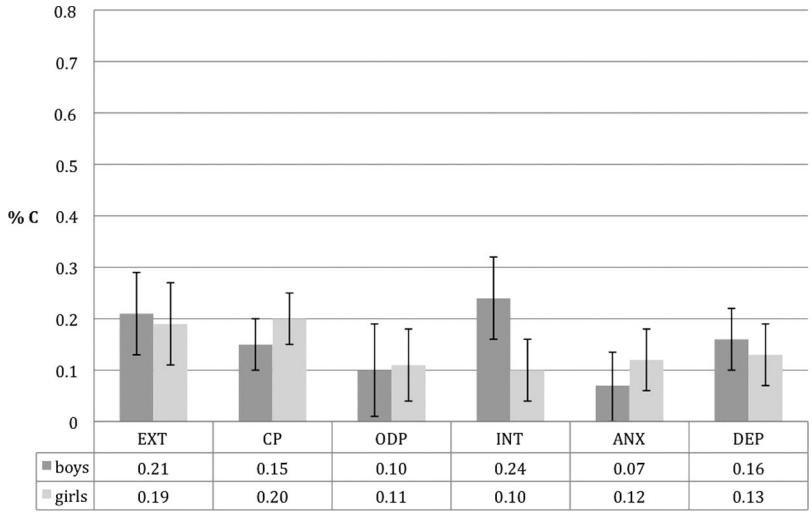
	EXT	CP	INT	ANX*	DEP
Twin	0.15	0.15	0.16	0.11	0.14
Adoption	0.15	0.17	0.19	0.31	0.10

Figure 1. Comparison of shared environmental estimates across twin and adoption studies. EXT = externalizing; twin (N = 13 samples with 9,879 sibling pairs) and adoption (N = 3 samples with 1,078 sibling pairs). CP = conduct problems; twin (N = 34 samples with 27,484 sibling pairs) and adoption (N = 4 samples with 1,225 sibling pairs). INT = internalizing; twin (N = 15 samples with 12,507 sibling pairs) and adoption (N = 2 samples with 592 sibling pairs). DEP = depression; twin (N = 16 samples with 20,899 sibling pairs) and adoption (N = 1 sample with 128 sibling pairs). ANX = anxiety; twin (N = 21 samples with 20,284 sibling pairs) and adoption (N = 2 samples with 502 sibling pairs). ODP (oppositional defiant problems) is not presented as there were no adoption studies available for analysis. The proportion of shared environmental variance (% C) is presented graphically as well as numerically. Error bars indicate the 95% confidence intervals. * Estimates are not equivalent across twin and adoption studies.

estimates of c^2 were obtained for at least three informant reports across all phenotypes. Such results strongly suggest that shared environmental influences on child and adolescent psychopathology are not specific to maternal or parental reports but generalize to some extent across informants. However, estimates of c^2 do appear to vary by informant, even within phenotype. Paternal reports typically yielded higher estimates of c^2 (17%–33%) than did maternal reports (12%–21%), although both were uniformly greater than zero. Indeed, the only exception to significant shared environmental influences on parental reports was found for ADHP, for which c^2 was estimated to be .000 for maternal and paternal informant reports (results not shown). Such findings are quite important in the context of the current study, as they indicate that parent reports are not universally prone to shared environmental mediation. Child self-reports generally yielded relatively small estimates of c^2. These estimates were greater than zero for EXT (8%, but only at $p < .10$), CP (13%), ANX (9%), and DEP (7%), but did not appear to meaningfully contribute to either ODP (0%) or INT (4%). Shared environmental influences on teacher reports were significant only for EXT, INT, ODP, and DEP, but were estimated to be insignificant and near zero for CP and ANX.

When available, I also computed c^2 estimates for observer and peer reports, although these estimates were based on only one or two studies, and thus confidence in the estimates is quite limited. Observer ratings were available only for CP and from only one study (Arsenault et al., 2003). Analyses estimated c^2 at 10%, although this estimate was not statistically significant. Peer reports were available for all phenotypes except ODP, though analyses were typically based on only one sample (the FinnTwin project; Happonen et al., 2002; Pulkkinen, Kaprio, & Rose, 1999). Analyses generally yielded small but nonsignificant estimates of c^2 for peer reports (8%, 11%, 13%, 8%, and 0% for EXT, CP, INT, ANX, and DEP, respectively).

In an effort to better understand differences in shared environmental influences across informants, I examined informant estimates in two additional ways. I first statistically compared child self-report estimates with mother and teacher estimates. Although maternal estimates generally appeared larger than child estimates, these differences were only statistically significant for DEP and ODP (both ps $< .001$). Differences for ANX and EXT were not significant. Estimates for CP and INT were also not significant (although a trend was evident; both ps = .07). The importance of these marginal differences for CP and INT is undermined by the very large number of sibling pairs for CP (29,205 pairs and 9,505 pairs for maternal and child reports, respectively) and, to a lesser extent, INT (11,727 pairs and 886 pairs for child reports, respectively). In short, although maternal estimates of c^2 may appear somewhat larger, there are few to no meaningful differences in shared environmental effect estimates across maternal and child reports for four of the six disorders examined here (CP, INT, EXT, or ANX). Comparisons of child and teacher reports yielded more significant differences, but in inconsistent directions. Teacher reports yielded higher estimates of c^2 than did child reports for two disorders (i.e., DEP and ODP), whereas child reports yielded higher estimates of c^2 than did teacher reports for two other disorders, CP ($p < .001$) and ANX ($p = .048$). There were no significant differences between teacher and child estimates for



	EXT	CP	ODP	INT	ANX	DEP
boys	0.21	0.15	0.10	0.24	0.07	0.16
girls	0.19	0.20	0.11	0.10	0.12	0.13

Figure 2. Comparison of shared environmental estimates across sex. EXT = externalizing (N = 9 samples with 2,875 male–male sibling pairs and 3,128 female–female sibling pairs); CP = conduct problems (N = 21 samples with 7,048 male–male sibling pairs and 9,310 female–female sibling pairs); ODP = oppositional defiant problems (N = 6 samples with 2,686 male–male sibling pairs and 4,690 female–female sibling pairs); INT = internalizing (N = 10 samples with 3,090 male–male sibling pairs and 5,308 female–female sibling pairs); DEP = depression (N = 10 samples with 5,157 male–male sibling pairs and 6,724 female–female sibling pairs); ANX = anxiety (N = 13 samples with 5,209 male–male sibling pairs and 7,292 female–female sibling pairs). The proportion of shared environmental variance (% C) is presented graphically as well as numerically. Error bars indicate the 95% confidence intervals. Estimates are uniformly equivalent across sex.

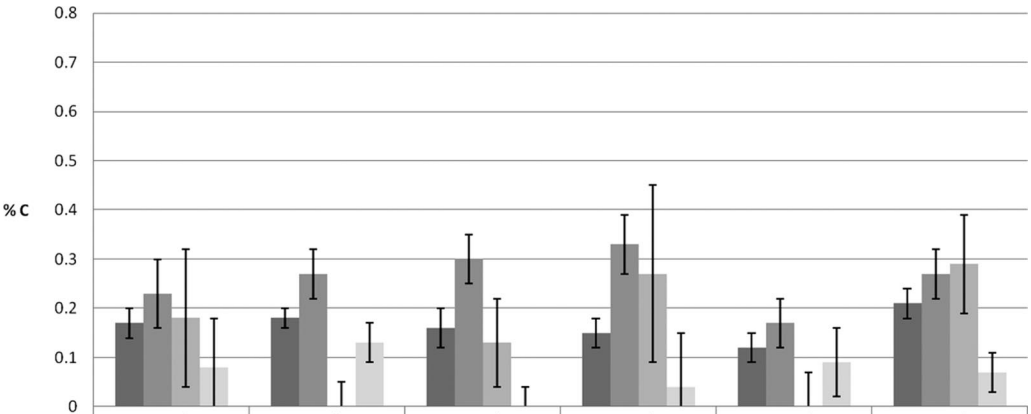
EXT and INT. In sum, there is no consistent pattern of differences in shared environmental effects across teacher and child reports.

I next examined genetic and nonshared environmental estimates by informant. As seen in Figure 4, estimates of e^2 were notably larger for child self-reports (ranging from 42%–64%, with an average of 57%) than were e^2 estimates for other informants (which ranged from 19%–45%, with an average of 30%). Moreover, the estimates of genetic influence on child self-reports were generally smaller than those of the other informants. Across all six disorders, average genetic influences on maternal, paternal, and teacher reports were estimated to be 57%, 47%, and 51%, respectively (precise estimates of a^2 are presented for each disorder in Appendix B). By contrast, the average proportion of variance in child self-reports accounted for by genetic influences was estimated at only 37% (range = 30%–50%). Thus, both genetic and shared environmental influences appear to be somewhat reduced, whereas nonshared environmental influences are increased, for child self-reports compared to other informant reports.

Age. Estimates of c^2 were computed separately for sibling pairs ages 1–5, 6–10, and 11–18 years (see Figure 5). Consistent with expectations, estimates of c^2 do appear to vary with age for all six phenotypes. Shared environmental influences on CP were largest in early childhood (23%), but decreased to 15%–16% in middle childhood and remained there through adolescence. ODP and DEP also evidenced the largest shared environmental influences in early to mid-childhood (25% and 19%, respectively), but estimates decreased dramatically by adolescence (0% and 8%, respectively). Alternately, shared environmental influences on INT were weakest in early childhood (8%) and had increased substantially by middle childhood (32%), although they weakened again in adolescence (23%). The proportion of variance in EXT accounted for by c^2 remained modest to moderate in magnitude across all age ranges (14%–23%), although it was significantly smaller in early childhood than in middle childhood. Finally, ANX estimates were modest across all ages (9%–15%).

Assessment method. Estimates of c^2 were computed separately for questionnaires and diagnostic interviews (see Figure 6). Shared environmental influences were uniformly significant when disorders were assessed using questionnaires. However, when assessed via diagnostic interviews, shared environmental influences were greater than zero only for EXT and CP (and the former was nonsignificant at p = .086), and were estimated to be zero or near-zero for all remaining phenotypes. Even so, the c^2 estimates for the diagnostic interviews did not differ significantly from the questionnaire estimates for any disorder except ODP and ANX ($\Delta\chi^2 < 2.8$ on 1 df, ns, for EXT, CP, INT, and DEP).

Given the relative consistency of shared environmental influences across other moderators, these results are somewhat surprising. As with the informant reports, it may be that part of the answer can be obtained by examining the genetic and nonshared environmental estimates. As seen in Figure 7, estimates of e^2 were uniformly larger when the disorder was assessed via diagnostic interview as opposed to questionnaire. Nonshared environmental influences were particularly large for diagnostic interviews of the internalizing disorders (accounting for as much as 48%–68% of the variance). By contrast, genetic influences on DEP and INT (though not ANX; see Appendix B) were significantly larger when assessed via questionnaire (52%–54% of the variance; average = 53%) as opposed to diagnostic interview (27%–37% of the variance; average =



	EXT*	CP*	ODP*	INT*	ANX*	DEP*
mothers	0.17	0.18	0.16	0.15	0.12	0.21
fathers	0.23	0.27	0.30	0.33	0.17	0.27
teacher	0.18	0.00	0.13	0.27	0.00	0.29
child	0.08	0.13	0.00	0.04	0.09	0.07

Figure 3. Comparison of shared environmental estimates across informant. EXT = externalizing (N = 16 samples with maternal reports on 9,380 sibling pairs, paternal reports on 2,630 sibling pairs, teacher reports on 1,548 sibling pairs, and child self-reports on 1,317 sibling pairs); CP = conduct problems (N = 37 samples with maternal reports on 29,205 sibling pairs, paternal reports on 5,958 sibling pairs, teacher reports on 9,699 sibling pairs, and child self-reports on 9,505 sibling pairs); ODP = oppositional defiant problems (N = 9 samples with maternal reports on 11,871 sibling pairs, paternal reports on 4,893 sibling pairs, teacher reports on 2,309 sibling pairs, and child self-reports on 4,386 sibling pairs); INT = internalizing (N = 16 samples with maternal reports on 11,727 sibling pairs, paternal reports on 3,331 sibling pairs, teacher reports on 634 sibling pairs, and child self-reports on 886 sibling pairs); DEP = depression (N = 17 samples with maternal reports on 15,267 sibling pairs, paternal reports on 7,204 sibling pairs, teacher reports on 1,392 sibling pairs, and child self-reports on 12,439 sibling pairs); ANX = anxiety (N = 23 samples with maternal reports on 19,395 sibling pairs, paternal reports on 5,982 sibling pairs, teacher reports on 3,696 sibling pairs, and child self-reports on 4,961 sibling pairs). The proportion of shared environmental variance (% C) is presented graphically, as well as numerically. Error bars indicate the 95% confidence intervals. * Estimates vary across informant.

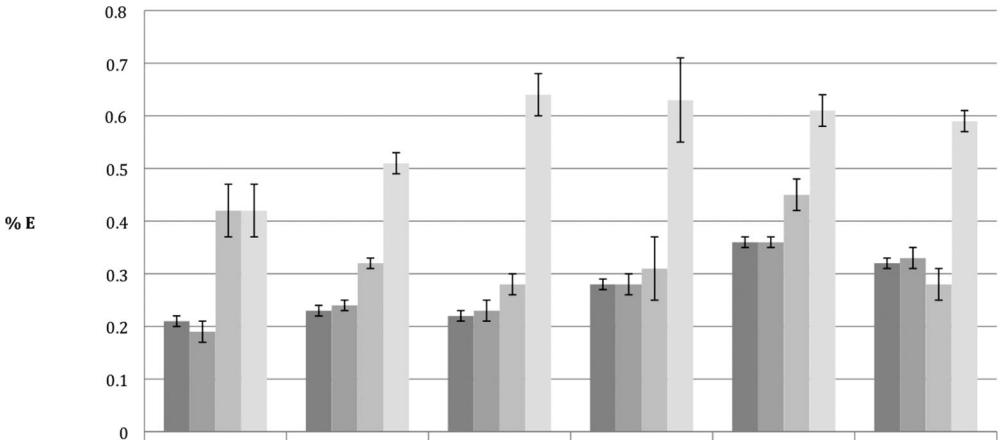
32%). In short, both genetic and shared environmental influences appear to be somewhat reduced, whereas nonshared environmental influences are increased, when using diagnostic interviews in place of informant questionnaires.

Supplemental Analyses

I also sought to address the aforementioned issues of operationalization for CP, EXT, and ANX. The analyses of EXT were largely based on questionnaire measures, the vast majority of which were in the Achenbach family of instruments. However, the Achenbach EXT scale for ages 6 to 18 is a composite of the Aggression and Rule-Breaking scales, and thus is essentially an extension of CP. To maintain consistency with extant literature, in which EXT is often considered separately from CP, I elected to retain the Achenbach studies within the EXT data group for the primary analyses. To evaluate the robustness of the results, however, I removed all EXT studies employing the Achenbach 6–18 scale (14 samples) from the EXT data group and reanalyzed EXT data. Those samples omitted from the EXT analyses that were not already represented in the CP analyses (i.e., 5 samples) were added to the CP data group, and CP data were then reanalyzed.

Results were almost identical to those reported previously. All estimates of c^2 for CP were within 0.9% of those reported above. The only exception to this was that c^2 was estimated to be 16.0% in adoption studies (rather than 17.3%), though the estimate for twin studies did not change (14.4% vs. 14.6%). The conclusions for CP are thus unchanged. Similarly, shared environmental influences on EXT overall were estimated at 19.0% of the variance (rather than 15.3%). Shared environmental estimates also did not substantively vary across operationalization of EXT for sex, informant, assessment method, or age (though there was a stronger indication of decreasing shared environmental influences with age, as c^2 was estimated to be 10% during adolescence, rather than 16%). That the results remained consistent for both EXT and CP suggests that they are generally robust to these operationalization issues.

Operationalization was also a concern for some measures of ANX. As evidenced by their respective titles (i.e., Anxious Depressed and Emotional Problems), the ANX scales on the Child Behavior Checklist and the Strengths and Difficulties Questionnaire include items indexing both DEP and ANX. However, inspection of their item content revealed that most items appeared to be specifically tapping ANX. After some deliberation, I elected to include these studies within ANX. To evaluate the robustness of the results, however, all ANX studies making use of the Child Behavior Checklist or the Strengths and Difficulties Questionnaire (nine samples) were removed for additional ANX analyses.



	EXT*	CDP*	ODP*	INT*	ANX*	DEP*
mothers	0.21	0.23	0.22	0.28	0.36	0.32
fathers	0.19	0.24	0.23	0.28	0.36	0.33
teacher (same)	0.42	0.32	0.28	0.31	0.45	0.28
child	0.42	0.51	0.64	0.63	0.61	0.59

Figure 4. Comparison of nonshared environmental estimates across informant. EXT = externalizing; CP = conduct problems; ODP = oppositional defiant problems; INT = internalizing; DEP = depression; ANX = anxiety. The proportion of nonshared environmental variance (% E) is presented graphically as well as numerically. Error bars indicate the 95% confidence intervals. Sample sizes are the same as those reported in Figure 3. * Estimates vary across informant.

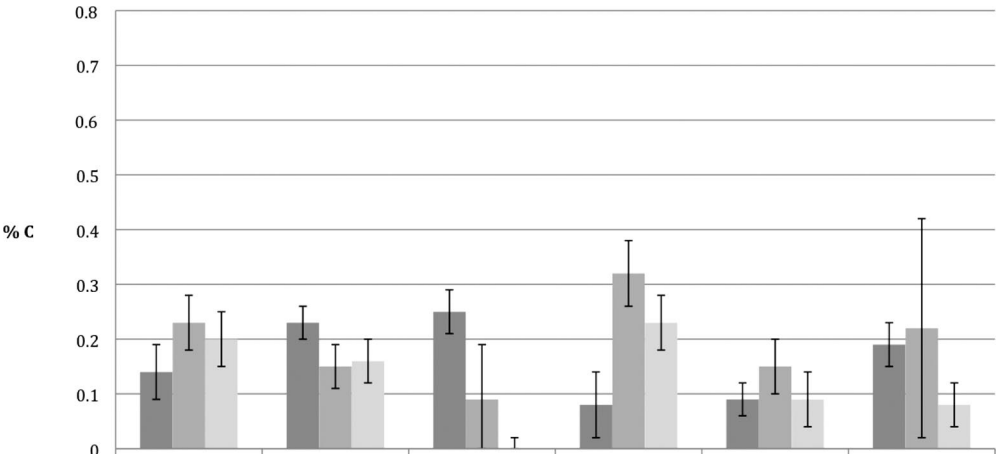
As with CP and EXT, there was little evidence of meaningful differences in estimates of c^2. Overall shared environmental influences were estimated to be 14.3% (as opposed to 12.2%). Results regarding the impact of informant, assessment method, and sex were similarly unchanged. The pattern observed for c^2 with age was generally comparable during early- and middle-childhood (11% and 20%, respectively, as opposed to 9% and 15%), but was no longer significant during adolescence (1% vs. 9%). Such findings again suggest that the results reported here are generally robust to issues of operationalization.

Discussion

The primary goal of the current series of meta-analyses was to evaluate the role of the shared environment in common disorders of childhood and adolescence, namely conduct problems, oppositional defiant problems, attention-deficit/hyperactivity problems, depression, anxiety, and broad externalizing and internalizing phenotypes. I collected 490 twin and adoption studies across all phenotypes, of which I ultimately included 38 samples for CP, 9 for ODP, 16 for EXT, 17 for DEP, 23 for ANX, 17 for INT, and 26 for ADHP. The full ACE model provided the best fit to the data for all disorders except ADHP, for which the ADE model fit best (this disorder was thus omitted from the remaining analyses and discussion except where specifically indicated). Shared environmental influences were uniformly significant across EXT, CP, ODP, INT, DEP, and ANX, and accounted for 10%–15% of the variance in the externalizing disorders and 12%–16% of the variance in the internalizing disorders. In no case was the c^2 estimate larger in twin studies than in adoption studies, results that strongly argue against passive rGE confounds in these estimates of the shared environment. Importantly, however, the proportion of phenotypic variance accounted for by the shared environment varied across informant, age, and/or assessment method for all disorders, results which will be discussed in depth below. In sum, these results are consistent with the core hypothesis of the present study, as they suggest that shared environmental influences play a clear role in the presence of child and adolescent psychopathology, with the sole exception of ADHP.

Limitations

Use of maximum likelihood estimation. These model-fitting analyses assume that the variables under study are normally distributed, an assumption that does not generally hold in the current study. Data for child and adolescent psychological and behavioral problems are usually positively skewed, such that there are relatively few individuals with the disorder in question. When the normal distribution is violated is this way, weighted least squares (WLS) estimation is preferable to maximum likelihood (ML) estimation for obtaining asymptotically correct standard errors and chi-square fit statistics (as described nicely by Rhee & Waldman, 2002). Unfortunately, WLS estimation requires weight matrices (i.e., variance–covariance matrices), to which I rarely had access. Instead, I was limited to examining published data only (i.e., intraclass correlations) for the vast majority of studies. Although the use of correlations (and ML estimation) was thus unavoidable, it does have one important analytical consequence. When Rhee and Waldman (2002) compared parameter estimates obtained on correlations and those obtained on variance-covariance matrices, they found that the exclusion of the weight matrices systematically overestimated genetic influences and systematically underestimated the magnitude of shared environmental influences (i.e., the c^2 estimate for antisocial behavior was reduced from 28% to 17%)



	EXT*	CP*	ODP*	INT*	ANX*	DEP*
ages 0-5	0.14	0.23	0.25	0.08	0.09	0.19
ages 6-10	0.23	0.15	0.09	0.32	0.15	0.22
ages 11-18	0.20	0.16	0.00	0.23	0.09	0.08

Figure 5. Comparison of shared environmental estimates across age. EXT = externalizing (N = 15 samples with 5,661 sibling pairs age 1–5 years, 4,848 sibling pairs age 6–10 years, and 6,023 sibling pairs age 11–18 years); CP = conduct problems (N = 30 samples with 16,954 sibling pairs age 1–5 years, 8,960 sibling pairs age 6–10 years, and 12,091 sibling pairs age 11–18 years); ODP = oppositional defiant problems (N = 6 samples with 7,848 sibling pairs age 1–5 years, 1,490 sibling pairs age 6–10 years, and 5,028 sibling pairs age 11–18 years); INT = internalizing (N = 14 samples with 5,663 sibling pairs age 1–5 years, 4,173 sibling pairs age 6–10 years, and 4,993 sibling pairs age 11–18 years); DEP = depression (N = 13 samples with 8,630 sibling pairs age 1–5 years, 540 sibling pairs age 6–10 years, and 9,519 sibling pairs age 11–18 years); ANX = anxiety (N = 17 samples with 16,145 sibling pairs age 1–5 years, 9,585 sibling pairs age 6–10 years, and 8,378 sibling pairs age 11–18 years). The proportion of shared environmental variance (% C) is presented graphically as well as numerically. Error bars indicate 95% confidence intervals. * Estimates vary across age.

Estimates of the shared environment obtained here should thus be viewed as *conservative*, a particularly striking point in light of the small-to-moderate estimates already observed herein.

Shared environment in adulthood. Although the current study is useful for examining shared environmental influences on behavioral and emotional disturbances in childhood and adolescence, these results have little to say about shared environmental influences in adulthood. Adult samples were intentionally excluded from these analyses, both because prior work has generally concluded that shared environmental influences are minimal in adulthood but also to keep the meta-analyses tractable. It thus remains unclear how these results may generalize to adulthood. Of note, however, shared environmental influences were still modest to moderate in magnitude (i.e., at least 15% of the total variance) during adolescence for EXT, INT, and CP, and were smaller but still significant for ANX and DEP. Given these findings, it would be beneficial to know the point in the lifespan at which shared environmental influences lose their salience. One likely time would be the transition to adult independence that typically occurs just after high school (i.e., moving away from the family home). Future research should specifically evaluate this hypothesis within a longitudinal design.

Implications and Future Directions

ADHP appears to be etiologically unique among psychological disorders of childhood and adolescence. Unlike the other disorders evaluated here, ADHP did not appear to be influenced by the shared environment (i.e., c^2 was estimated to be .000), but was significantly influenced by nonadditive (i.e., d^2 accounted for 44% of the variance, and the ADE model provided an improved fit to the data). Such results suggest that ADHP is etiologically unique among the more common forms of child and adolescent psychopathology. Namely, ADHP appears to be influenced by interactions among genes (and prominently so, particularly early in life), whereas genetic influences on the other disorders appear to be exclusively additive (i.e., a function of multiple genes summed over loci). Further, unlike the other disorders, similarity between siblings appears to be solely a function of common genetic influences. Common environmental experiences, by contrast, appear to make no etiological contribution to ADHP. Despite the strength of these findings, I know of no convincing explanation for the relatively unique etiology of AHDP, especially given its high levels of comorbidity with other externalizing disorders. Rater-contrast effects (i.e., when parents overemphasize the hyperactivity differences in their DZ twins) have been raised as one explanation for the presence of nonadditivity for ADHP (Eaves et al., 2000; Simonoff et al., 1998), but it remains unclear why this phenomenon would apply only to inattention and hyperactivity and not to other child behavioral problems (Kuntsi, Gayan, & Stevenson, 2000). Future research should seek to better understand the seemingly unique etiology of ADHP compared to other forms of child and adolescent psychopathology.

Understanding different environmentality estimates across assessment method. Although shared environmental influences were uniformly significant when disorders were assessed using

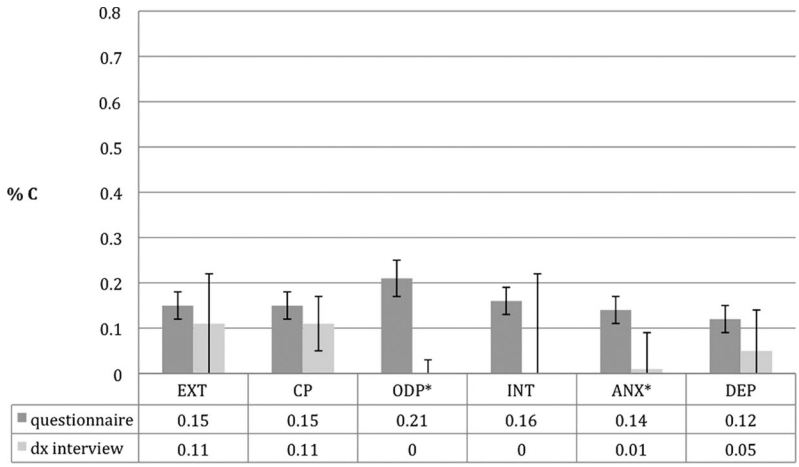


Figure 6. Comparison of shared environmental estimates across assessment method. EXT = externalizing (N = 16 samples with 9,834 sibling pairs assessed via questionnaire, and with 2,218 sibling pairs assessed via diagnostic interview); CP = conduct problems (N = 38 samples with 23,908 sibling pairs assessed via questionnaire, and 7,333 sibling pairs assessed via diagnostic interview); ODP = oppositional defiant problems (N = 9 samples with 8,459 sibling pairs assessed via questionnaire, and 6,235 sibling pairs assessed via diagnostic interview); INT = internalizing (N = 17 samples with 12,473 sibling pairs assessed via questionnaire, and 626 sibling pairs assessed via diagnostic interview); DEP = depression (N = 17 samples with 17,120 sibling pairs assessed via questionnaire, and 3,910 sibling pairs assessed via diagnostic interview); ANX = anxiety (N = 23 samples with 17,239 sibling pairs assessed via questionnaire, and 4,621 sibling pairs assessed via diagnostic interview); dx interview = diagnostic interview. The proportion of shared environmental variance (% C) is presented graphically as well as numerically. Error bars indicate the 95% confidence intervals. * Estimates vary significantly across assessment method.

questionnaires, they were only occasionally significant when assessed via diagnostic interviews. Even so, c^2 estimates for the diagnostic interviews did not differ significantly from those obtained using questionnaires for any disorder except ODP and ANX. This overlap may partially reflect the relatively limited number of participants assessed via formal diagnostic interviews (which naturally results in less precise estimates). However, the consistent pattern of results across disorders suggests that there may be meaningful differences in heritability estimates by assessment method. Estimates of e^2 were uniformly larger when the disorder was assessed via diagnostic interview as opposed to questionnaire, whereas genetic influences were often larger when

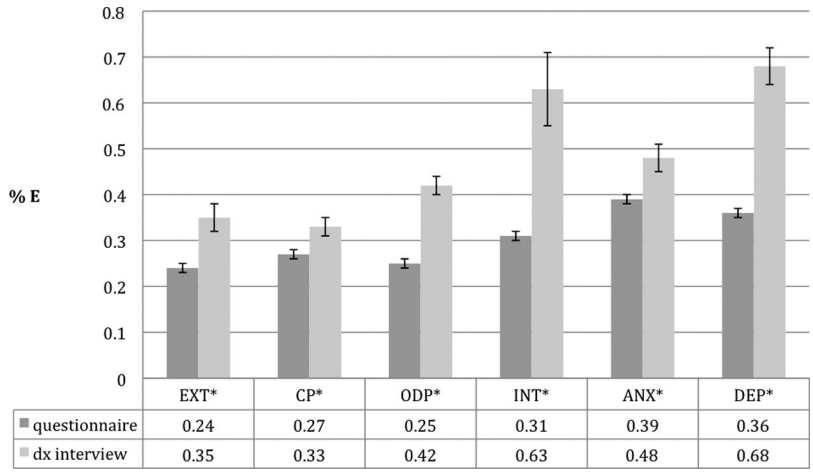


Figure 7. Comparison of nonshared environmental estimates across assessment method. EXT = externalizing; CP = conduct problems; ODP = oppositional defiant problems; INT = internalizing; DEP = depression; ANX = anxiety; dx interview = diagnostic interview. The proportion of nonshared environmental variance (% E) is presented graphically as well as numerically. Error bars indicate the 95% confidence intervals. Sample sizes are the same as those reported in Figure 6. * Estimates vary significantly across assessment method.

assessed via questionnaire (average = 53% for INT and DEP) compared to diagnostic interview (average = 32%). Thus, both genetic and shared environmental influences appear to be somewhat reduced, whereas nonshared environmental influences are increased, when using diagnostic interviews in place of informant questionnaires.

Though the origin of these differences remains unclear, one possibility is that such results reflect an increase in measurement error for diagnostic interviews (because measurement error is contained within e^2), as questionnaires have forced-choice responses that are likely to improve reliability. Alternately, these differences could reflect differences in skew across the two assessment methods. Questionnaires often assess both normative and deviant ranges of the trait in question, whereas diagnostic interviews assess only clinically significant symptoms, resulting in a "floor effect" and prominent skew (i.e., most participants have symptom counts close to the low end of the scale). For example, as reported in Legrand, McGue, & Iacono (1999), although adolescent self-report questionnaires of anxiety are skewed (i.e., −0.10 to 1.0), *DSM–III–R* symptom counts of separation anxiety disorder and overanxious disorder in those same adolescents are far more skewed (i.e., 4.05 to 4.11; Legrand et al., 1999). As noted previously, nonnormality can distort ACE estimates obtained via ML estimation. Diagnostic interviews may therefore simply be less useful for establishing heritability and environmentality.

There is another possible interpretation, however. Rather than reflecting measurement error of some kind, these differences could reflect true differences in genetic and environmental influences across severity, such that nonshared environmental influences are more etiologically salient for those with more disturbed functioning. There are a few recent studies attempting to identify nonshared environmental effects contributing to *extreme* sibling differences using an MZ differences design. Asbury, Dunn, Pike, and Plomin (2003) examined whether sibling differences in parental treatment were linked to child-specific behavior problems (as assessed via the brief Strength and Difficulties questionnaire), and whether extreme sibling differences potentiated this relationship, in a large cross-sectional sample of 4 year-old MZ twin pairs (n = 2,353). They found that differential parental treatment accounted for an average of 3% of the variance in differential sibling outcomes across the full sample, but 11% of the variance in highly discordant pairs.

My colleagues and I then extended these results using a longitudinal and cross-lagged MZ differences design (Burt, McGue, Iacono, et al., 2006). We found that differential parent–child conflict at age 11 uniquely contributed to differential sibling externalizing symptom counts at age 14, but only in the most discordant twin pairs. In the full unselected sample, this relationship was very small and not significant. Such findings are consistent with the notion that nonshared environmental influences may be more important for more severe (i.e., clinically significant) manifestations of child and adolescent psychopathology. However, we examined sibling differences using a clinical measure, namely a combined symptom count of conduct disorder and oppositional defiant disorder as assessed via the Diagnostic Interview for Children and Adolescents–Revised (DICA-R; Reich & Welner, 1988). If the above severity hypothesis were true, we would have expected some association even for smaller symptom differences between siblings (as each symptom was required to be clinically significant in both severity and frequency). Moreover, as noted in the introduction, available data suggest that shared environmental influences are often most prominent in high risk or disadvantaged environments—findings that are not in keeping with this interpretation. Future research should seek to determine whether differences in heritability and environmentality across assessment method stem from measurement error or skew or meaningful differences in these estimates with disorder severity.

Understanding different environmentality estimates across informants. Significant estimates of c^2 were obtained for at least three informant reports across all disorders, results that strongly suggest that shared environmental influences on child and adolescent psychopathology are not unique to specific informant reports but generalize (at least to some extent) across informants. Estimates of shared environmental influences do appear to vary by informant, however, as do estimates of genetic and nonshared environmental influences. These differences may reflect situational specificity or differences in perspective and level or type of exposure to the children's behavior (Achenbach, McConaughy, & Howell, 1987). In other words, it may be that problematic behaviors and emotions are expressed differently and for different reasons in various contexts and accordingly appear more or less heritable or environmental within those contexts (Burt et al., 2005b). This interpretation of situational specificity by informant is consistent with reports of enhanced disorder validity when using multiple informants (Bird, Gould, & Staghezza, 1992; Hart, Lahey, Loeber, & Hanson, 1994; Jensen et al., 1999), as well as reports of higher genetic influences when using multiple informants (Arsenault et al., 2003). However, the particular pattern of genetic and environmental influences observed here suggests that although situational specificity may be one contributing factor, other forces may also be at play. These possibilities are discussed below.

First, paternal reports typically yielded somewhat higher estimates of the shared environment than did maternal reports (average estimates were 26% and 17%, respectively). That estimates vary within the parental unit is somewhat surprising, as parents are likely to be similarly exposed to their children's behaviors, at least in intact families. Differential exposure to children's behaviors in divorced families may partially explain these results, as might the inclusion of families where one parent stays at home with the children while the other works; however, situational specificity has limited explanatory power within the parental unit in general. Another possible explanation is that fathers are somehow less able or willing to distinguish between their children than are mothers—a complication that, if present for both MZ and DZ pairs, would act to inflate estimates of the shared environment. However, this interpretation is undercut by the absence of shared environmental influences identified for ADHP, which suggests that fathers are able to distinguish between their DZ twins, at least for some behaviors. It thus remains unclear what may account for this finding. Future research should seek to more fully understand the origin of differences in maternal and paternal perceptions of their child.

Second, genetic and environmental parameter estimates varied somewhat across child and maternal informant reports (although these differences were typically not quite statistically significant). One possible explanation for these findings is that of shared method variance or shared informant effects. Child reports involve correlating the reports of two separate informants (i.e., the chil-

dren's self-reports are correlated), whereas parental reports typically involve correlating reports by the same informant (i.e., a given parent reports on both children, and these reports are then correlated across the sample). Given this, the higher correlations observed for parental reports could be a function of shared informant effects. If true, this would suggest that although the shared environment still generally contributes to child and adolescent psychopathology, the contribution is somewhat smaller than that reported herein and applies only to CP, ANX, DEP, and perhaps EXT, which yielded significant estimates of c^2 for child self-reports (7%–13% of the variance).

Alternately, these results could stem from the increased unreliability and measurement error common in child self-reports (which would, by definition, decrease sibling similarity and thus increase the nonshared environmental proportion of variance to the exclusion of genetic and shared environmental proportions of variance). This latter interpretation is bolstered in part by the often smaller estimates of genetic influence found for child self-reports (37% on average) compared to those for parental informants (47%–57% on average) and those for teacher informants (51% on average). The estimates of genetic influence for child self-reports also appear to be somewhat smaller than those from adult twin self-reports (with the exception of DEP). For example, heritability estimates for self-reports of adult antisocial behavior are estimated to be approximately 50% (although they range as high as 70%; Burt, Carter, McGue, & Iacono, 2007; Miles & Carey, 1997; Slutske et al., 2001; Vernon, McCarthy, Johnson, Jang, & Harris, 1999), compared to 36.5% in the current meta-analysis. That estimates of genetic influence are reduced only for self-reports in childhood, but not for self-reports in adulthood or when using other informant reports during childhood, implies that child self-reports may indeed be plagued by increased measurement error. Future research should further explore informant effects on genetic and environmental contributions to child and adolescent psychopathology.

The impact of age on shared environmental influences. Although shared environmental estimates appeared to vary by age, the pattern of differences did not always conform to expectations (i.e., decreasing c^2 with age). Estimates of c^2 appeared weaker in adolescence than in childhood for some disorders (i.e., ODP, INT, and DEP), although this pattern did not hold for CP, EXT, or ANX, for which the estimates remained small to moderate in magnitude during adolescence. Such results could suggest that the commonly held belief that shared environmental influences dissipate by adulthood may not be universally true. Importantly, however, the current findings are cross-sectional and thus are potentially confounded by cohort effects. Longitudinal studies are needed to more convincingly support (or reject) hypotheses about etiological change with age.

Fortunately, a few of the studies examined in the current meta-analysis were longitudinal, allowing me to informally examine actual change with age. For EXT, available studies suggested that there was little to no change in the magnitude of either genetic or shared environmental influences with age (Bartels, van den Oord, et al., 2004; Burt et al., 2005a; Haberstick, Schmitz, Young, & Hewitt, 2005; van der Valk, van den Oord, Verhulst, & Boomsma, 2003; van der Valk, Verhulst, Neale, & Boomsma, 1998). For INT, two studies reported evidence of decreasing genetic influences and increasing shared environmental influences with age (Bartels, van den Oord, et al., 2004; van der Valk et al., 2003), whereas others reported no changes with age (Haberstick et al., 2005; van der Valk et al., 1998). Studies were similarly inconsistent for CP. Three reported no evidence of changes with age (Haberstick, Schmitz, Young, & Hewitt, 2006; van Beijsterveldt, Bartels, Hudziak, & Boomsma, 2003; van der Valk et al., 1998), whereas others suggested that either genetic influences decrease with age (Vierikko, Pulkkinen, Kaprio, & Rose, 2006) or that shared environmental influences increase with age (Eley, Lichtenstein, & Moffitt, 2003). The only available study for DEP indicates that genetic influences increase while shared environmental influences decrease with age (Scourfield et al., 2003). Finally, although one study reported a general pattern of decreasing a^2 and increasing c^2 for ANX (in girls only; Topolski et al., 1999), there was little evidence of significant etiological change with age in another study of ANX (van der Valk et al., 1998).

When the above longitudinal findings are examined alongside those of the present study, there appears to be mounting evidence against the expected pattern of dramatically decreasing shared environmental influences (and increasing genetic influences) from early childhood through adolescence, with the clear exception of ODP (and maybe DEP). Given this, I argue that there is a need to reconsider existing theories of the etiological development of the more common forms of child and adolescent psychopathology. Future research should seek to develop and test new theories on this fundamentally important developmental topic.

Identification of the shared environment. Taken together, the results of the current meta-analysis strongly suggest that the shared environment makes consequential contributions to common forms of child and adolescent psychopathology (with the sole exception of attention-deficit/hyperactivity disorder). These shared environmental influences do not appear to represent passive rGE in disguise, as shared environmental estimates were equivalent for nonadoptive versus adoptive youth. Instead, they likely represent persistent and identifiable main effects of the environment on the development of psychopathology. Such results have important implications for those seeking to understand environmental influences on child and adolescent psychopathology, as well as those looking to intervene in the onset, development, and/or maintenance of these conditions. Indeed, variables accounting for 10%–30% of the total phenotypic variance in a given disorder constitute fundamental indices of risk. For example, although the association between parental divorce and adolescent conduct problems is now well accepted (Amato, 2001), closer inspection reveals that divorce actually accounts for just over 1% of the variance in adolescent CP (Burt et al., 2008). Identification of 10%–30% of the phenotypic variance thus seems likely to represent a fundamental gain in our clinical understanding and treatment of child and adolescent psychopathology.

Efforts to identify the risk and protective environmental factors that constitute these shared environmental influences should be given high priority in future research (Rutter, Pickles, Murray, & Eaves, 2001). Existing research on this topic has been quite promising. Several independent samples have now found convincing evidence that the origin of the association between parental divorce and adolescent behavior problems is shared environmental in origin (Burt et al., 2008; D'Onofrio et al., 2005, 2007; O'Connor et al., 2000). Recent evidence also suggests that parenting and the parent–child relationship are associated with adolescent externalizing via shared environmental mechanisms, at least in part (Burt

et al., 2003; Burt, McGue, et al., 2007, McGue et al., 1996; Pike et al., 1996). For example, in an effort to identify shared environmental influences on externalizing disorders, Burt et al. (2003) incorporated a measured psychosocial variable (i.e., mother and adolescent reports of parent–child conflict) into the design. Results indicated that the shared environmental influences contributing to conflict accounted for roughly 12% of the *total* variance (or 23% of the shared environmental variance) in EXT.

The parent–child relationship and parental divorce thus appear to meaningfully account for shared environmental variance adolescent externalizing. Future research should explore other factors (e.g., prenatal influences and toxin exposure, neighborhood effects, school effects, and sibling influences) that may also contribute to child and adolescent psychopathology via shared environmental mechanisms. Future research should also examine the specificity of these effects across the various forms of psychopathology to determine whether shared environment influences differentiate or unite particular disorders. For instance, poverty may act as a shared environmental influence common to all forms of childhood psychopathology, whereas other sorts of shared environmental influences (e.g., conflictive parenting) may be specific to certain disorders or clusters of disorders.

To do this sort of research effectively, researchers should keep the following in mind: To be sure that the contribution is specifically shared environmental in nature, it is necessary to incorporate or control for genetic similarity between siblings. This can be done using a variety of methodological designs (e.g., twin studies, adoption studies) and statistical techniques (e.g., multilevel modeling, structural equation modeling, comparisons of adoptive and nonadoptive youth). Adoptive sibling studies are certainly the most powerful and straightforward design, as any association between nonbiologically related relatives must be shared environmental in nature. Twin studies are also useful, although they require large samples (at least 700 pairs or more) to reliably detect moderate shared environmental influences. Interested readers are referred to Plomin et al. (2008) for more detail.

An important distinction must also be drawn between objective and effective environments (Goldsmith, 1993; Turkheimer & Waldron, 2000). The objective environment refers to the observable environment and does not consider how it may affect the different members of the family. Thus, whether an objective environmental event is shared or nonshared rests only on whether the event is common to the siblings, regardless of whether it acts to increase their similarity. The effective environment, by contrast, refers solely to the outcome of the event (Turkheimer & Waldron, 2000). It is this definition of the environment that was used here (as described in the introduction). Shared environmental factors can thus be either individual-level (e.g., family relationships) or family-level (e.g., divorce or neighborhood effects) forces that effectively act to make family members similar to each other. In short, essentially any environmental experience could conceivably operate at the shared environmental level. However, only individual-level variables can be analyzed using traditional behavioral genetic decompositions of variance. Family-level variables can be examined with multilevel modeling (Burt, McGue, et al., 2007; Guo & Wang, 2002).

Reconciling the theory of the environment as responsible primarily for sibling differences with the results of the current study. Behavioral genetic research has historically concluded that the more important environmental influences on psychological and behavioral outcomes result in differences between siblings (Plomin & Daniels, 1987), a conclusion that continues to influence theory and interpretation of environmental influences up to the present day. For psychopathology, however, this conclusion was based largely on studies using smaller adult samples and thus has little relevance to child and adolescent psychopathology. More recent research has suggested modest to moderate influences of the shared environment on child and adolescent psychopathology. Moreover, shared environmental influences appear to be identifiable and to play a significant role in continuity of symptoms over time (prior to adulthood), findings that stand in stark contrast to those for nonshared environmental effects (Rutter et al., 1999; Turkheimer & Waldron, 2000). One interpretation of these findings is that although shared environmental effects are often a function of relatively persistent, systematic influences, nonshared influences are largely idiosyncratic and unsystematic in nature (Burt, McGue, Iacono, & Kreuger, 2006; Rutter et al., 1999; Turkheimer & Waldron, 2000). The results of the current meta-analyses dovetail nicely with this more recent conceptualization of the environment, strongly supporting the role of shared or family-level environmental factors in the development of most emotional and behavioral problems during childhood and adolescence.

Even so, we should not abandon the nonshared environment as meaningless to child and adolescent psychopathology. Although shared environmental influences account for a significant proportion of the total phenotypic variance, nonshared environmental influences remain particularly potent sources of environmental variance. More recent research has suggested that nonshared environmental influences may be more salient to developmental outcomes in high-risk (as opposed to population-based) samples (Asbury et al., 2003; Burt, McGue, Iacono, & Krueger, 2006). Moreover, environmental influences need not be solely shared or nonshared in origin; rather, a given environmental factor could contribute to child behavioral problems at both the family-wide and child-specific levels, serving both to color the global home environment that surrounds the children and to influence each child individually (see Burt, McGue, et al., 2007, for an example of simultaneous shared and nonshared environmental mediation).

In conclusion, these results suggest that the shared or family-level environment contributes to most forms of child and adolescent psychopathology. Empirical examinations and theoretical incorporation of the shared environment into current theories of the origins of these disorders are consequently of real and critical value for understanding the development and persistence of common mental health issues. Future research should seek to more fully understand shared environmental influences prior to adulthood.

References

Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101,* 213–232.

Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms and profiles.* Burlington: University of Vermont, Research Center for Children, Youth, & Families.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52,* 317–332.

Amato, P. R. (2001). Children of divorce in the 1990s: An update of the Amato and Keith (1991) meta-analysis. *Journal of Family Psychology, 15,* 355–370.

Arsenault, L., Moffitt, T. E., Caspi, A., Taylor, A., Rijsdijk, F. V., Jaffee, S. R., et al. (2003). Strong genetic effects on cross-situational antisocial behavior among 5-year-old children according to mothers, teachers, examiner–observers, and twins' self-reports. *Journal of Child Psychology and Psychiatry, 44,* 832–848.

Asbury, K., Dunn, J. F., Pike, A., & Plomin, R. (2003). Nonshared environmental influences on individual differences in early behavioral development: A monozygotic twin differences study. *Child Development, 74,* 933–943.

Bartels, M., Boomsma, D. I., Hudziak, J. J., Rietveld, M. J. H., van Beijsterveldt, T., & van den Oord, E. (2004). Disentangling genetic, environmental, and rater effects on internalizing and externalizing problem behavior in 10-year-old twins. *Twin Research and Human Genetics, 7,* 162–175.

Bartels, M., Hudziak, J. J., Boomsma, D. I., Rietveld, M. J. H., van Beijsterveldt, C. E. M., & van den Oord, E. J. C. G. (2003). A study of parent ratings of internalizing and externalizing problem behavior in 12-year-old twins. *Journal of the American Academy of Child & Adolescent Psychiatry, 42,* 1351–1359.

Bartels, M., van den Oord, E. J. C. G., Hudziak, J. J., Rietveld, M. J. H., van Beijsterveldt, C. E. M., & Boomsma, D. I. (2004). Genetic and environmental mechanisms underlying stability and change in problem behaviors at ages 3, 7, 10, and 12. *Developmental Psychology, 40,* 852–867.

Bergen, S. E., Gardner, C. O., & Kendler, K. S. (2007). Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: A meta-analysis. *Twin Research and Human Genetics, 10,* 423–433.

Bird, H. R., Gould, M., & Staghezza, B. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *Journal of the American Academy of Child & Adolescent Psychiatry, 31,* 75–85.

Blonigen, D. M., Hicks, B. M., Krueger, R. F., Patrick, C. J., & Iacono, W. G. (2005). Psychopathic personality traits: Heritability and genetic overlap with internalizing and externalizing psychopathology. *Psychological Medicine, 35,* 637–648.

Burt, S. A., Barnes, A. R., McGue, M., & Iacono, W. G. (2008). Parental divorce and adolescent delinquency: Ruling out the impact of common genes. *Developmental Psychology, 44,* 1668–1677.

Burt, S. A., Carter, L. A., McGue, M., & Iacono, W. G. (2007). The different origins of stability and change in antisocial personality disorder symptoms. *Psychological Medicine, 37,* 27–38.

Burt, S. A., Krueger, R. F., McGue, M., & Iacono, W. G. (2003). Parent–child conflict and the comorbidity among childhood externalizing disorders. *Archives of General Psychiatry, 60,* 505–513.

Burt, S. A., McGue, M., Demarte, J. A., Krueger, R. F., & Iacono, W. G. (2006). Timing of menarche and the origins of conduct disorder. *Archives of General Psychiatry, 63,* 890–896.

Burt, S. A., McGue, M., Iacono, W. G., & Krueger, R. F. (2006). Differential parent–child relationships and adolescent externalizing symptoms: Cross-lagged analyses within a twin differences design. *Developmental Psychology, 42,* 1289–1298.

Burt, S. A., McGue, M., Krueger, R. F., & Iacono, W. G. (2005a). How are parent–child conflict and child externalizing behaviors related over time? Results from a genetically-informative cross-lagged study. *Development and Psychopathology, 17,* 1–21.

Burt, S. A., McGue, M., Krueger, R. F., & Iacono, W. G. (2005b). Sources of covariation among child externalizing disorders: Informant effects and the shared environment. *Psychological Medicine, 35,* 1133–1144.

Burt, S. A., McGue, M., Krueger, R. F., & Iacono, W. G. (2007). Environmental contributions to adolescent delinquency: A fresh look at the shared environment. *Journal of Abnormal Child Psychology, 35,* 787–800.

Cheung, M. W., & Chan, W. (2005). Classifying correlation matrices into relatively homogeneous subgroups: A cluster analytic approach. *Educational and Psychological Measurement, 65,* 954–979.

Cleveland, H. H. (2003). Disadvantaged neighborhoods and adolescent aggression: Behavioral genetic evidence of contextual effects. *Journal of Research on Adolescence, 13,* 211–238.

Conners, C. K. (2001). *Conners Rating Scales–Revised.* New York: Multi-Health Systems.

Costello, E., & Angold, A. (1988). Scales to assess child and adolescent depression: Checklists, screens, and nets. *Journal of the American Academy of Child and Adolescent Psychiatry, 27,* 726–737.

Coventry, W. L., & Keller, M. C. (2005). Estimating the extent of parameter bias in the classical twin design: A comparison of parameter estimates from extended twin–family and classical twin designs. *Twin Research and Human Genetics, 8,* 214–223.

D'Onofrio, B. M., Turkheimer, E., Emery, R. E., Slutske, W. S., Heath, A. C., Madden, P. A. F., et al. (2005). A genetically informed study of marital instability and its association with offspring psychopathology. *Journal of Abnormal Psychology, 114,* 570–586.

D'Onofrio, B. M., Turkheimer, E., Emery, R. E., Maes, H. H., Silberg, J. L., & Eaves, L. J. (2007). A children of twins study of parental divorce and offspring psychopathology. *Journal of Child Psychology and Psychiatry, 48,* 667–675.

DuPaul, G. J. (1981). Parent and teacher ratings of ADHD symptoms: Psychometric properties in a community-based sample. *Journal of Clinical Child Psychology, 20,* 245–253.

Duyme, M., Dumaret, A.-C., & Tomkiewicz, S. (1999). How can we boost IQs of "dull children"? A late adoption study. *Proceedings of the National Academy of Sciences, 96,* 8790–8794.

Eaves, L. J., Rutter, M., Silberg, J. L., Shillady, L., Maes, H., & Pickles, A. (2000). Genetic and environmental causes of covariation in interview assessments of disruptive behavior in child and adolescent twins. *Behavior Genetics, 30,* 321–334.

Eaves, L. J., Silberg, J. L., Meyer, J. M., Maes, H. H., Simonoff, E., Pickles, A., et al. (1997). Genetics and developmental psychopathology: 2. The main effects of genes and environment on behavioral problems in the Virginia twin study of adolescent development. *Journal of Child Psychology and Psychiatry, 38,* 965–980.

Eley, T. C., Lichtenstein, P., & Moffitt, T. E. (2003). A longitudinal behavioral genetic analysis of the etiology of aggressive and nonaggressive antisocial behavior. *Development and Psychopathology, 15,* 383–402.

Farrington, D. P., & West, D. J. (1971). A comparison between early delinquents and young aggressives. *British Journal of Criminology, 11,* 341–358.

Gibson, H. B. (1967). Self-report delinquency among school boys and their attitudes towards police. *British Journal of Social and Clinical Psychology, 20,* 303–315.

Goldsmith, H. H. (1993). Nature–nurture issues in the behavioral genetics context: Overcoming barriers to communication. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture, and psychology* (pp. 325–339). Washington, DC: American Psychological Association.

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38,* 581–586.

Guo, G., & Wang, J. (2002). The mixed or multilevel model for behavioral genetic analysis. *Behavior Genetics, 32,* 37–50.

Haberstick, B. C., Schmitz, S., Young, S. E., & Hewitt, J. K. (2005). Contributions of genes and environments to stability and change in externalizing and internalizing problems during elementary and middle school. *Behavior Genetics, 35,* 381–396.

Haberstick, B. C., Schmitz, S., Young, S. E., & Hewitt, J. K. (2006). Genes

and developmental stability of aggressive behavior problems at home and school in a community sample of twins aged 7–12. *Behavior Genetics, 36,* 809–819.

Happonen, M., Pulkkinen, L., Kaprio, J., Van der Meere, J., Viken, R. J., & Rose, R. J. (2002). The heritability of depressive symptoms: Multiple informants and multiple measures. *Journal of Child Psychology and Psychiatry, 43,* 471–479.

Harris, J. R. (1998). *The nurture assumption: Why children turn out the way they do.* New York: Free Press.

Harris, J. R. (2000). Socialization, personality development, and the child's environments: Comment on Vandell (2000). *Developmental Psychology, 36,* 711–723.

Hart, E. L., Lahey, B. B., Loeber, R., & Hanson, K. S. (1994). Criterion validity of informants in the diagnosis of disruptive behavior disorders in children: A preliminary study. *Journal of Consulting and Clinical Psychology, 62,* 410–414.

Jensen, P. S., Rubio-Stipec, M., Canino, G., Bird, H. R., Duncan, M. K., Schwab-Stone, M. E., et al. (1999). Parent and child contributions to diagnosis of mental disorder: Are both informants always necessary? *Journal of the American Academy of Child & Adolescent Psychiatry, 38,* 1569–1579.

Judd, L. L., Kessler, R. C., Paulus, M. P., Zeller, P. V., Wittchen, H. U., & Kunovac, J. L. (1998). Comorbidity as a fundamental feature of generalized anxiety disorders: Results from the National Comorbidity Study (NCS). *Acta Psychiatrica Scandinavica, 393*(Suppl.), 6–11.

Keller, M. C., & Coventry, W. L. (2005). Quantifying and addressing parameter indeterminacy in the classical twin design. *Twin Research and Human Genetics, 8,* 201–213.

Keller, M. C., & Medland, S. E. (2008, June). *Evaluation of the cascade model: A new extended twin family model.* Paper presented at the Behavioral Genetics Association, Louisville, KY.

Kessler, R. C., Nelson, C. B., McGonagle, K. A., Liu, J., Swartz, M., & Blazer, D. G. (2005). Comorbidity of *DSM–III–R* major depressive disorder in the general population: Results from the U.S. National Comorbidity Survey. *British Journal of Psychiatry, 168*(Suppl. 30), 17–30.

Kovacs, M. (1992). *Children's Depression Inventory manual.* North Tonawanda, NY: Multi-Health Systems.

Krueger, R. F., & Finger, M. S. (2001). Using item response theory to understand comorbidity among anxiety and unipolar mood disorders. *Psychological Assessment, 13,* 140–151.

Kuntsi, J., Gayan, J., & Stevenson, J. (2000). Parents' and teachers' ratings of problem behaviours in children: Genetic and contrast effects. *Twin Research and Human Genetics, 3,* 251–258.

LaBuda, M. C., Gottesman, I. I., & Pauls, D. L. (1993). Usefulness of twin studies for exploring the etiology of childhood and adolescent psychiatric disorders. *American Journal of Medical Genetics, 48,* 47–59.

Legrand, L. N., McGue, M., & Iacono, W. G. (1999). A twin study of state and trait anxiety in childhood and adolescence. *Journal of Child Psychology and Psychiatry, 40,* 953–958.

Li, M. D., Cheng, R., Ma, J. Z., & Swan, G. E. (2003). A meta-analysis of estimated genetic and environmental effects on smoking behavior in male and female adult twins. *Addiction, 98,* 23–31.

Lipsey, M. W., & Wilson, D. B. (Eds.). (2001). *Applied Social Research Methods Series: Vol. 49. Practical meta-analysis*. Thousand Oaks, CA: Sage.

Martin, N. G., Eaves, L. J., Kearsey, M. J., & Davies, P. (1978). The power of the classical twin study. *Heredity, 1978,* 97–116.

Mash, E. J., & Wolf, D. A. (2005). *Abnormal child psychology* (3rd ed.). Belmont, CA: Wadsworth.

McGue, M., & Bouchard, T. J. J. (1998). Genetic and environmental influences on human differences. *Annual Review of Neuroscience, 21,* 1–24.

McGue, M., Bouchard, T. J. J., Iacono, W. G., & Lykken, D. T. (1993). Behavioral genetics of cognitive ability: A life-span perspective. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture, and psychology* (pp. 59–76). Washington, DC: American Psychological Association.

McGue, M., Keyes, M., Sharma, A., Elkins, I., Legrand, L., Johnson, W., et al. (2007). The environment of adopted and non-adopted youth: Evidence of range restriction from the Sibling Interaction and Behavior Study (SIBS). *Behavioral Genetics, 37,* 449–462.

McGue, M., Sharma, A., & Benson, P. (1996). The effect of common rearing on adolescent adjustment: Evidence from a U.S. adoption cohort. *Developmental Psychology, 32,* 604–613.

Miles, D. R., & Carey, G. (1997). Genetic and environmental architecture of human aggression. *Journal of Personality and Social Psychology, 72,* 207–217.

Murray, K. T., & Sines, J. O. (1996). Parsing the genetic and nongenetic variance in children's depressive behavior. *Journal of Affective Disorders, 38,* 23–34.

Neale, M. C. (1997). *Mx: Statistical modeling.* (4th ed.). (Available from Virginia Commonwealth University, Box 900126, Richmond, VA 23298)

Neiderhiser, J. M., Reiss, D., & Hetherington, E. M. (2007). The Nonshared Environment in Adolescent Development (NEAD) Project: A longitudinal family study of twins and siblings from adolescence to young adulthood. *Twin Research and Human Genetics, 10,* 74–83.

Neiderhiser, J. M., Reiss, D., Pedersen, N., Lichtenstein, P., Spotts, E. L., Hansson, K., et al. (2004). Genetic and environmental influences on mothering of adolescents: A comparison of two samples. *Developmental Psychology, 40,* 335–351.

O'Connor, T. G., Caspi, A., DeFries, J. C., & Plomin, R. (2000). Are associations between parental divorce and children's adjustment genetically mediated? An adoption study. *Developmental Psychology, 36,* 429–437.

Olweus, D. (1989). Prevalence and incidence in the study of anti-social behavior: Definitions and measurements. In M. W. Klein (Ed.), *Cross-national research in self-reported crime and delinquency.* Dordrecht, The Netherlands: Kluwer.

Parker, T. (1989, August). *Television viewing and aggression in four and seven year old children.* Paper presented at the Summer Multicultural Access to Research Training conference, Boulder, CO.

Patrick, C. J., Curtin, J. J., & Tellegen, A. (2002). Development and validation of a brief form of the Multidimensional Personality Questionnaire. *Psychological Assessment, 14,* 150–163.

Petrill, S. A., & Deater-Deckard, K. (2004). Task orientation, parental warmth and SES account for a significant proportion of the shared environmental variance in general cognitive ability in early childhood: Evidence from a twin study. *Developmental Science, 7,* 25–32.

Pike, A., McGuire, S., Hetherington, E. M., Reiss, D., & Plomin, R. (1996). Family environment and adolescent depressive symptoms and antisocial behavior: A multivariate genetic analysis. *Developmental Psychology, 32,* 590–603.

Plomin, R., & Asbury, K. (2005). Nature and nurture: Genetic and environmental influences on behavior. *Annals of the American Academy of Political and Social Science, 600,* 86–98.

Plomin, R., Asbury, K., & Dunn, J. (2001). Why are children from the same family so different? Nonshared environment a decade later. *Canadian Journal of Psychiatry, 46,* 225–233.

Plomin, R., Chipuer, H. M., & Neiderhiser, J. M. (1994). Behavioral genetic evidence for the importance of nonshared environment. In E. M. Hetherington, D. Reiss, & R. Plomin (Eds.), *Separate social worlds of siblings: The impact of nonshared environment on development* (pp. 1–31). Hillsdale, NJ: Erlbaum.

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

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype–environment

interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84,* 309–322.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGruffin, P. (2008). *Behavioral genetics* (5th ed.). New York: Worth.

Pulkkinen, L., Kaprio, J., & Rose, R. J. (1999). Peers, teachers and parents as assessors of the behavioural and emotional problems of twins and their adjustment: The multidimensional peer nomination inventory. *Twin Research and Human Genetics, 2,* 274–285.

Purcell, S. (2002). Variance components model for gene–environment interaction in twin analysis. *Twin Research, 5,* 554–571.

Reich, W., & Welner, Z. (1988). *Diagnostic Interview for Children and Adolescents–Revised: DSM–III–R version (DICA-R*). St. Louis, MO: Washington University.

Reiss, D., Hetherington, M., Plomin, R., Howe, G. W., Simmens, S. J., Henderson, S. H., et al. (1995). Genetic questions for environmental studies: Differential parenting and psychopathology in adolescence. *Archives of General Psychiatry, 52,* 925–936.

Reiss, D., Neiderhiser, J. M., Hetherington, E. M., & Plomin, R. (2000). *The relationship code: Deciphering genetic and social influences on adolescent development.* Cambridge, MA: Harvard University Press.

Reynolds, C. R., & Richmond, B. O. (1978). What I think and feel: A revised measure of children's manifest anxiety. *Journal of Abnormal Child Psychology, 6,* 271–280.

Rhee, S., & Waldman, I. D. (2002). Genetic and environmental influences on antisocial behavior: A meta-analysis of twin and adoption studies. *Psychological Bulletin, 128,* 490–529.

Rowe, D. C. (1994). *The limits of family influence: Genes, experience, and behavior.* New York: Guilford Press.

Rutter, M., Pickles, A., Murray, R., & Eaves, L. (2001). Testing hypotheses on specific environmental causal effects on behavior. *Psychological Bulletin, 127,* 291–324.

Rutter, M., Silberg, J., O'Connor, T. J., & Simonoff, E. (1999). Genetics and child psychiatry: I. Advances in quantitative and molecular genetics. *Journal of Child Psychology and Psychiatry, 40,* 3–18.

Rutter, M., Tizard, J., & Whitmore, K. (1970). *Education, health and behaviour.* London: Longman.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype–environment effects. *Child Development, 54,* 424–435.

Scourfield, J., Rice, F., Thapar, A., Harold, G. T., Martin, N., & McGuffin, P. (2003). Depressive symptoms in children and adolescents: Changing aetiological influences with development. *Journal of Child Psychology and Psychiatry, 44,* 968–976.

Shaffer, D., Fisher, P., Dulcan, M. K., Davies, M., Piacentini, J., Schwab-Stone, M. E., et al. (1996). The NIMH Diagnostic Interview Schedule for Children Version 2.3 (DISC-2.3): Description, acceptability, prevalence rates, and performance in the MECA study. *Journal of the American Academy of Child and Adolescent Psychiatry, 35,* 865–877.

Simonoff, E., Pickles, A., Hervas, A., Silberg, J. L., Rutter, M., & Eaves, L. (1998). Genetic influences on childhood hyperactivity: Contrast effects imply parental rating bias, not sibling interaction. *Psychological Medicine, 28,* 825–837.

Slutske, W. S., Eisen, S., Xian, H., True, W. R., Lyons, M. J., Goldberg, J., et al. (2001). A twin study of the association between pathological gambling and antisocial personality disorder. *Journal of Abnormal Psychology, 110,* 297–308.

Topolski, T. D., Hewitt, J. K., Eaves, L., Meyer, J. M., Silberg, J. L., Simonoff, E., et al. (1999). Genetic and environmental influences on ratings of manifest anxiety by parents and children. *Journal of Anxiety Disorders, 13,* 371–397.

Towers, H., Spotts, E., Neiderhiser, J. M., Plomin, R., Hetherington, E. M., & Reiss, D. (2000). Genetic and environmental influences on teacher ratings of the Child Behavior Checklist. *International Journal of Behavioral Development, 24,* 373–381.

Tully, L. A., Arseneault, L., Caspi, A., Moffitt, T. E., & Morgan, J. (2004). Does maternal warmth moderate the effects of birth weight on twins' attention-deficit/hyperactivity disorder (ADHD) symptoms and low IQ? *Journal of Consulting and Clinical Psychology, 72,* 218–226.

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science, 13,* 160–164.

Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B., & Gottesman, I. I. (2003). Socioeconomic status modifies heritability of IQ in young children. *Psychological Science, 14,* 623–628.

Turkheimer, E., & Waldron, M. (2000). Nonshared environment: A theoretical, methodological, and quantitative review. *Psychological Bulletin, 126,* 78–108.

Tuvblad, C., Grann, M., & Lichtenstein, P. (2006). Heritability for adolescent antisocial behavior differs with socioeconomic status: Gene–environment interaction. *Journal of Child Psychology and Psychiatry, 47,* 734–743.

van Beijsterveldt, C. E. M., Bartels, M., Hudziak, J. J., & Boomsma, D. I. (2003). Causes of stability of aggression from early childhood to adolescence: A longitudinal genetic analysis in Dutch twins. *Behavior Genetics, 33,* 591–605.

van der Valk, J. C., van den Oord, E. J. C. G., Verhulst, F. C., & Boomsma, D. I. (2003). Genetic and environmental contributions to stability and change in children's internalizing and externalizing problems. *Journal of the American Academy of Child & Adolescent Psychiatry, 42,* 1212–1220.

van der Valk, J. C., Verhulst, F. C., Neale, M. C., & Boomsma, D. I. (1998). Longitudinal genetic analysis of problem behaviors in biologically related and unrelated adoptees. *Behavior Genetics, 28*(5), 365–380.

van IJzendoorn, M. H., & Juffer, F. (2005). Adoption is a successful natural intervention enhancing adopted children's IQ and school performance. *Current Directions in Psychological Science, 14,* 326–330.

Vernon, P. A., McCarthy, J. M., Johnson, A. M., Jang, K. L., & Harris, J. A. (1999). Individual differences in multiple dimensions of aggression: A univariate and multivariate genetic analysis. *Twin Research, 2,* 16–21.

Vierikko, E., Pulkkinen, L., Kaprio, J., & Rose, R. J. (2006). Genetic and environmental sources of continuity and change in teacher-rated aggression during early adolescence. *Aggressive Behavior, 32,* 308–320.

Appendix A

Table A1
Example of Mx Script

```
G1: Model parameters, Conduct Problems
Calculation NG = 7
Matrics
X Lower 1 1 Free                    ! a: additive genetic parameter
Y Lower 1 1 Free                    ! c: shared environmental parameter
Z Lower 1 1 Free                    ! e: unique environmental parameter
W Lower 1 1 Free                    ! d: non-additive genetic influence parameter
I Iden 2 2
H Full 1 1 ! scalar, .5
Q Full 1 1 ! scalar, .25
End Matrics;
fix w 1 1                           !d is fixed to zero because model can include either c or d
                                      but not both

Matrix H .5
Matrix Q .25
Begin Algebra;
A = X * X';                         !a^2: additive genetic variance
C = Y * Y';                         !c^2: shared environmental variance
E = Z * Z';                         !e^2: unique environmental variance
D = W * W';                         !d^2: non-additive genetic variance
V = A + C + E + D;                  ! total variance
p = A|C|E|D;                        ! put parameter estimates in one matrix
S = P@V~;                           ! standardized parameter estimates
End Algebra;
End

Title G2: FS, age 10–18, CHILD ques, Burt et al., 2007
Data NI = 2 NO = 204
KMatrix Symm
1
.33 1
Matrices = Group 1
Covariances A + C + E + D | H@A + C + Q@D _
H@A + C + Q@D | A + C + E + D /
Option Rsiduals
End

Title G3: URT, age 10–18, CHILD ques, Burt et al., 2007
Data NI = 2 NO = 406
KMatrix Symm
1
.17 1
Matrices = Group 1
Covariances A + C + E + D | C _
C | A + C + E + D /
Option Rsiduals
End

Title G4: MZ males, age 11, AVERAGED, Burt et al., 2005b & Johnson et al., 2005
Data NI = 2 NO = 272
KMatrix Symm
1
.63 1
Matrices = Group 1
Covariances A + C + E + D | A + C + D _
A + C + D | A + C + E + D /
Option Rsiduals
End
```

(*Appendixes continue*)

Table A1 (*continued*)

```
Title G5: MZ females, age 11, AVERAGED, Burt et al., 2005b & Johnson et al., 2005
Data NI = 2 NO = 283
KMatrix Symm
1
.59 1
Matrices = Group 1
Covariances A + C + E + D | A + C + D _
A + C + D | A + C + E + D /
Option Rsiduals
End

Title G6: DZ males, age 11, AVERAGED, Burt et al., 2005b & Johnson et al., 2005
Data NI = 2 NO = 130
KMatrix Symm
1
.45 1
Matrices = Group 1
Covariances A + C + E + D | H@A + C + Q@D _
H@A + C + Q@D | A + C + E + D /
Option Rsiduals
End

Title G7: DZ females, age 11, AVERAGED, Burt et al., 2005b & Johnson et al., 2005
Data NI = 2 NO = 176
KMatrix Symm
1
.20 1
Matrices = Group 1
Covariances A + C + E + D | H@A + C + Q@D _
H@A + C + Q@D | A + C + E + D /
Option Rsiduals
Option NDecimals = 4
Option DF = −11
Option Interations = 100
Option Check
End
```

Note. NI = number of variables in correlation matrix (i.e., 2; disorder for Siblings 1 and 2); NO = number of observations (i.e., twin/sibling pairs); MZ = monozygotic; DZ = dizygotic; FS = full sibling; URT = adopted or unrelated reared together.

Appendix B

Table B1
Parameter Estimates for All Primary Fitted Models

Model	a^2	c^2	d^2	e^2
EXT				
Full				
ACE	.590 (.552, .629)*	.153 (.118, .187)*		.258 (.248, .269)*
ADE	.739 (.717, .761)*		.000 (.000, .013)	.250 (.240, .259)*
Study type (ACE)				
Twin	.592 (.548, .637)*	.151 (.109, .192)*		.258 (.248, .269)*
Adoption	.642 (.316, .906)*	.155 (.092, .220)*		.203 (.000, .512)
Sex (ACE)				
Boys	.542 (.461, .628)*	.211 (.129, .291)*		.247 (.231, .265)*
Girls	.533 (.450, .621)*	.194 (.111, .274)*		.274 (.257, .293)*
Informant (ACE)				
Maternal	.617 (.579, .656)*	.171 (.136, .207)*		.212 (.202, .222)*
Paternal	.578 (.511, .649)*	.230 (.164, .296)*		.193 (.176, .212)*
Teacher	.406 (.266, .555)*	.177 (.045, .303)*		.417 (.380, .458)*
Child	.497 (.379, .622)*	.084 (.000, .183)		.422 (.380, .469)*
Peer	.643 (.415, .835)*	.077 (.000, .279)		.280 (.226, .352)*
Observer	—	—		—
Age (ACE)				
Ages 1–5	.627 (.572, .684)*	.142 (.089, .193)*		.231 (.218, .245)*
Ages 6–10	.575 (.527, .625)*	.231 (.183, .279)*		.191 (.180, .204)*
Ages 11–18	.545 (.496, .596)*	.195 (.148, .241)*		.259 (.246, .273)*

(table continues)

Table B1 (*continued*)

Model	a^2	c^2	d^2	e^2
Assessment method (ACE)				
Quest	.617 (.578, .656)*	.147 (.111, .182)*		.237 (.227, .248)*
Dx inter	.541 (.419, .672)*	.106 (.000, .221)~		.353 (.329, .380)*
CP				
Full				
ACE	.576 (.550, .602)*	.145 (.121, .169)*		.280 (.273, .287)*
ADE	.719 (.706, .733)*		.000 (.000, .005)	.270 (.263, .276)*
Study Type (ACE)				
Twin	.575 (.547, .603)*	.146 (.121, .171)*		.279 (.272, .286)*
Adoption	.367 (.160, .574)*	.173 (.102, .245)*		.460 (.303, .632)*
Sex (ACE)				
Boys	.558 (.501, .617)*	.150 (.096, .204)*		.292 (.279, .305)*
Girls	.526 (.478, .576)*	.204 (.157, .251)*		.270 (.260, .280)*
Informant (ACE)				
Maternal	.594 (.570, .619)*	.178 (.155, .201)*		.228 (.222, .234)*
Paternal	.491 (.440, .543)*	.272 (.223, .321)*		.237 (.223, .251)*
Teacher	.680 (.634, .705)*	.000 (.000, .039)		.321 (.308, .334)*
Child	.365 (.308, .423)*	.126 (.079, .173)*		.509 (.489, .530)*
Peer	.572 (.372, .748)*	.086 (.000, .257)		.342 (.289, .409)*
Observer	.506 (.340, .668)*	.102 (.000, .252)		.391 (.351, .438)*
Age (ACE)				
Ages 1–5	.512 (.480, .545)*	.226 (.196, .255)*		.262 (.253, .272)*
Ages 6–10	.629 (.585, .674)*	.148 (.106, .189)*		.223 (.213, .234)*
Ages 11–18	.531 (.491, .571)*	.160 (.124, .196)*		.309 (.298, .321)*
Assessment Method (ACE)				
Quest	.572 (.544, .600)*	.155 (.130, .180)*		.274 (.266, .281)*
Dx inter	.553 (.493, .614)*	.113 (.058, .167)*		.334 (.319, .350)*
ODP				
Full				
ACE	.591 (.547, .636)*	.101 (.062, .140)*		.308 (.297, .319)*
ADE	.693 (.670, .714)*		.000 (.000, .016)	.300 (.290, .311)*
Study Type (ACE)				
Twin	.591 (.547, .636)*	.101 (.062, .140)*		.308 (.297, .319)*
Adoption	—	—		—
Sex (ACE)				
Boys	.589 (.493, .691)*	.101 (.010, .189)*		.310 (.288, .335)*
Girls	.609 (.536, .686)*	.107 (.036, .176)*		.284 (.269, .300)*
Informant (ACE)				
Maternal	.619 (.581, .659)*	.159 (.122, .196)*		.221 (.213, .230)*
Paternal	.472 (.417, .527)*	.296 (.242, .348)*		.233 (.219, .248)*
Teacher	.585 (.489, .687)*	.132 (.040, .221)*		.283 (.260, .308)*
Child	.359 (.301, .397)*	.000 (.000, .042)		.642 (.609, .677)*
Peer	—	—		—
Observer	—	—		—
Age (ACE)				
Ages 1–5	.545 (.504, .588)*	.250 (.209, .291)*		.205 (.194, .216)*
Ages 6–10	.580 (.449, .711)*	.086 (.000, .198)		.334 (.298, .376)*
Ages 11–18	.639 (.596, .674)*	.000 (.000, .032)		.364 (.345, .384)*
Assessment Method (ACE)				
Quest	.538 (.493, .585)*	.211 (.167, .254)*		.251 (.240, .263)*
Dx inter	.587 (.549, .618)*	.000 (.000, .027)		.416 (.397, .435)*
ADHP				
Full				
ACE	.699 (.683, .715)*	.000 (.000, .002)		.317 (.308, .325)*
ADE	.259 (.198, .320)*		.444 (.383, .505)*	.297 (.289, .305)*
Study Type (did not fit since c^2 was estimated to be zero)				
Twin	—		—	—
Adoption	—		—	—

(*Appendixes continue*)

Table B1 (*continued*)

Model	a^2	c^2	d^2	e^2
Sex (ADE)				
Boys	.066 (.000, .197)		.622 (.490, .713)*	.312 (.298, .328)*
Girls	.183 (.065, .300)*		.519 (.403, .639)*	.298 (.286, .310)*
Informant (ADE)				
Maternal	.188 (.129, .246)*		.539 (.481, .599)*	.273 (.266, .280)*
Paternal	.000 (.000, .072)		.581 (.503, .610)*	.420 (.400, .442)*
Teacher	.706 (.609, .745)*		.013 (.000, .108)	.281 (.269, .294)*
Child	.359 (.182, .405)*		.000 (.000, .187)	.641 (.600, .684)*
Peer	.710 (.269, .837)*		.000 (.000, .441)	.290 (.234, .359)*
Observer	—		—	—
Age (ADE)				
Ages 1–5	.000 (.000, .022)		.631 (.602, .653)*	.372 (.358, .387)*
Ages 6–10	.216 (.099, .332)*		.489 (.372, .606)*	.296 (.282, .312)*
Ages 11–18	.415 (.317, .513)*		.277 (.181, .376)*	.308 (.296, .320)*
Assessment method (ADE)				
Quest	.283 (.218, .347)*		.426 (.362, .491)*	.291 (.283, .300)*
Dx inter	.104 (.000, .245)		.514 (.371, .640)*	.382 (.364, .401)*
INT				
Full				
ACE	.507 (.467, .547)*	.164 (.129, .198)*		.330 (.318, .343)*
ADE	.675 (.654, .695)*		.000 (.000, .010)	.316 (.305, .327)*
Study Type (ACE)				
Twin	.506 (.463, .549)*	.165 (.127, .203)*		.330 (.318, .342)*
Adoption	.064 (.000, .437)	.188 (.099, .278)*		.748 (.416, .905)*
Sex (ACE)				
Boys	.445 (.363, .532)*	.241 (.162, .318)*		.315 (.294, .337)*
Girls	.582 (.514, .653)*	.100 (.037, .162)*		.319 (.302, .336)*
Informant (ACE)				
Maternal	.575 (.535, .615)*	.150 (.114, .185)*		.277 (.266, .288)*
Paternal	.388 (.322, .455)*	.331 (.271, .392)*		.281 (.259, .305)*
Teacher	.418 (.198, .648)*	.272 (.071, .467)*		.310 (.250, .390)*
Child	.328 (.183, .455)*	.041 (.000, .151)		.631 (.557, .717)*
Peer	.662 (.467, .871)*	.128 (.000, .317)		.210 (.169, .265)*
Observer	—	—		—
Age (ACE)				
Ages 1–5	.573 (.504, .643)*	.076 (.017, .135)*		.351 (.331, .372)*
Ages 6–10	.371 (.305, .439)*	.320 (.258, .381)*		.309 (.290, .331)*
Ages 11–18	.414 (.356, .473)*	.228 (.178, .278)*		.358 (.337, .380)*
Assessment method (ACE)				
Quest	.538 (.499, .578)*	.156 (.121, .191)*		.306 (.294, .318)*
Dx inter	.373 (.126, .470)*	.000 (.000, .215)		.628 (.553, .715)*
DEP				
Full				
ACE	.437 (.400, .474)*	.139 (.110, .169)*		.424 (.411, .438)*
ADE	.592 (.575, .610)*		.000 (.000, .007)	.402 (.391, .413)*
Study type (ACE)				
Twin	.435 (.398, .472)*	.142 (.112, .171)*		.424 (.411, .437)*
Adoption	.000 (.000, .412)	.097 (.000, .286)*		.903 (.590, 1.00)*
Sex (ACE)				
Boys	.452 (.380, .525)*	.157 (.095, .218)*		.391 (.370, .415)*
Girls	.457 (.391, .524)*	.132 (.073, .189)*		.412 (.393, .432)*
Informant (ACE)				
Maternal	.466 (.429, .504)*	.214 (.182, .247)*		.319 (.308, .331)*
Paternal	.398 (.343, .452)*	.272 (.225, .320)*		.330 (.313, .348)*
Teacher	.431 (.318, .545)*	.289 (.187, .391)*		.281 (.248, .319)*
Child	.342 (.286, .397)*	.069 (.028, .111)*		.589 (.567, .613)*
Peer	.721 (.527, .836)*	.000 (.000, .169)		.281 (.234, .342)*
Observer	—	—		—
Age (ACE)				
Ages 1–5	.515 (.467, .565)*	.190 (.147, .233)*		.294 (.280, .310)*
Ages 6–10	.193 (.000, .475)	.222 (.003, .425)		.585 (.487, .706)*
Ages 11–18	.389 (.327, .450)*	.076 (.031, .121)*		.535 (.511, .562)*

(table continues)

Table B1 (*continued*)

Model	a^2	c^2	d^2	e^2
Assessment method (ACE)				
Quest	.516 (.479, .553)*	.124 (.094, .154)*		.360 (.348, .374)*
Dx inter	.273 (.163, .360)*	.045 (.000, .133)		.682 (.643, .724)*
ANX				
Full				
ACE	.475 (.438, .512)*	.122 (.091, .153)*		.404 (.392, .416)*
ADE	.607 (.590, .624)*		.000 (.000, .010)	.388 (.377, .399)*
Study type (ACE)				
Twin	.492 (.454, .531)*	.107 (.075, .139)*		.401 (.390, .414)*
Adoption	.076 (.000, .378)	.315 (.205, .425)*		.609 (.374, .758)*
Sex (ACE)				
Boys	.538 (.461, .617)*	.066 (.000, .132)		.397 (.376, .420)*
Girls	.532 (.470, .596)*	.116 (.058, .172)*		.353 (.338, .368)*
Informant (ACE)				
Maternal	.515 (.478, .552)*	.123 (.092, .154)*		.363 (.352, .375)*
Paternal	.467 (.401, .533)*	.171 (.115, .226)*		.363 (.343, .384)*
Teacher	.556 (.479, .599)*	.000 (.000, .059)		.445 (.417, .476)*
Child	.296 (.199, .393)*	.092 (.015, .168)*		.613 (.579, .648)*
Peer	.692 (.488, .883)*	.088 (.000, .281)		.220 (.177, .277)*
Observer	—	—		—
Age (ACE)				
Ages 1–5	.577 (.537, .617)*	.087 (.054, .121)*		.336 (.324, .349)*
Ages 6–10	.461 (.407, .516)*	.148 (.102, .193)*		.391 (.374, .409)*
Ages 11–18	.476 (.417, .536)*	.088 (.039, .136)*		.437 (.418, .457)*
Assessment method (ACE)				
Quest	.475 (.435, .515)*	.137 (.104, .170)*		.389 (.376, .402)*
Dx inter	.509 (.418, .558)*	.011 (.000, .087)		.480 (.454, .508)*

Note. a^2 = additive genetic influences; c^2 = shared environmental influences; d^2 = dominant genetic influences; e^2 = nonshared environmental influences; EXT = externalizing; CP = conduct problems; ODP = oppositional defiant problems; ADHP = attention-deficit/hyperactivity problems; INT = internalizing; DEP = depression; ANX = anxiety; dx inter = diagnostic interview. All study type and moderator analyses were conducted using the better-fitting overall model (i.e., ACE for EXT, CP, ODP, INT, DEP, and ANX; and ADE for ADHP). Dashes indicate that data were not available.
* $p < .05$.

Received October 16, 2008
Revision received January 12, 2009
Accepted January 13, 2009 ■