Child Development, November/December 2005, Volume 76, Number 6, Pages 1217–1233

Analysis and Interpretation of Twin Studies Including Measures of the Shared Environment

Eric Turkheimer and Brian M. D'Onofrio
University of Virginia

Hermine H. Maes and Lindon J. Eaves
Virginia Commonwealth University

Recent reports using a classical behavior genetic research design in which twin data are combined with a measured characteristic of their shared family environment have made striking claims about estimating environmental influences on behavior with genetic effects controlled. Such claims are overstated for two related reasons. First, when a variable is measured at the family level in a way that makes it necessarily equivalent for twins reared together, it is not possible to partition it into genetic and environmental components. Second, although structural equation modeling and DeFries–Fulker analysis are sound tools for the analysis of many types of twin data, they do not control for genetic or environmental confounds when estimating the effect of measured family-level variables.

Although it has long been promised that behavior genetic studies will help elucidate salient environmental influences on human variation (Heath, Kendler, Eaves, & Markell, 1985; Plomin, 1994; Reiss, Plomin, & Hetherington, 1991), the most frequent conclusion reached about the environment on the basis of genetically informed studies is that environmental factors shared by siblings do not influence children at all (Harris, 1998; Plomin & Daniels, 1987; Rowe, 1994; Turkheimer, in press; see Rutter, 2000, for a balanced review of the controversy).

Previous papers have reviewed the limited success of studies attempting to specify nonshared environmental influences (Turkheimer & Waldron, 2000). This article explores the potential and limitations of twin studies for the exploration of shared environmental factors. By "shared," we simply mean environmental factors that are jointly experienced by siblings raised together, although as we have noted previously (Turkheimer & Waldron, 2000), there are significant definitional problems in understanding shared and nonshared environment. Any purported environmental variable can contain both shared and nonshared variability. Even variables that would appear to apply to entire families, like socioeconomic status (SES) or martial status, may be differentially experienced by children in the same family. In many instances, however, variables are only measured at the level of the family, and therefore can only vary between families, not within them. Our review is motivated by both older and more recent articles that have added such family-level variables to twin studies. Recently, some of these studies have claimed to estimate the environmental influence of a measured family-level variable on an outcome in twin children, controlling for genetic influences on the same outcome.

Whereas an earlier review of the nonshared environment was concerned primarily with the empirical outcome of studies including a measured nonshared environmental variable (Turkheimer & Waldron, 2000), in this paper we demonstrate that twin studies including a family-level variable (which we refer to as the measured C design, C referring to the shared environmental term in traditional twin models) cannot provide estimates of environmental effects unbiased by genetic factors. Based on our review, we specify what twin studies including measures of the shared environment can and cannot do to elucidate environmental processes, and illustrate our conclusions with analyses of twin data from the Virginia 30,000 data set. We conclude with recommendations for future research designs and analytic methods.

Because some of our comments will be quite critical, we should be very explicit about the scope of our concerns. Our comments are limited to the par-

Preparation of this article and the analyses were supported by grants from the National Institute of Mental Health (MH67300) and William T. Grant Foundation. Data collection for the Virgina 30,000 data set was supported by Grants GM-30250, AG-04954, AA-06781, MH-40828, and HL-48148 from the National Institutes of Health and a gift from RJR Nabisco.

Correspondence concerning this article should be addressed to Eric Turkheimer, Department of Psychology, University of Virginia, P.O. Box 400400, Charlottesville, VA 22904-4400. Electronic mail may be sent to ent3c@virginia.edu. A commentary on this paper written by A. Caspi, A. Taylor, and S. Jaffe is available from Dr. Caspi at a.caspi@iop.kcl.ac.uk.

ticular subtype of twin analysis we have already described, in which a classical twin analysis is combined with a measured shared environmental variable. Moreover, we do not intend to suggest that either this research design or the usual statistical methods of analyzing its results are fundamentally flawed. Rather, the difficulties arise in the interpretation and generalization of the results, which in some recent cases have been insufficiently conservative. Another treatment of methodological issues in this design can be found in Purcell and Koenen (in press). More basic reviews of the advantages and disadvantages of genetically informative designs for the study of purported environmental factors can be found elsewhere (D'Onofrio et al., 2003; Eaves, Last, Young, & Martin, 1978; Rutter, Pickles, Murray, & Eaves, 2001).

Standard Twin Design

First, we will very briefly review the standard twin model. Classical twin studies compare the similarity of identical (monozygotic, MZ) twins and fraternal (dizygotic, DZ) twins. MZ twins share all of their genes and, on average, DZ twins share half of their genes. Therefore, to the extent that genetic factors influence a trait, then MZ twins will be more similar than DZ twins. By comparing the covariation among and between MZ and DZ twins, basic twin models partition the variance of a measured trait, called a phenotype, into additive genetic, shared environmental, and nonshared environmental components (Eaves, 1982).

Figure 1 is an example of the structural equation model (SEM) for a basic univariate twin analysis (e.g., Neale & Cardon, 1992). Double-headed arrows represent covariances and single-headed arrows specify regressions of one variable on another. The squares, *T1* and *T2*, represent the phenotypic measures of the twins. The *A* latent variable represents additive genetic effects. Therefore, the parameter connecting the twins' additive genetic variance components is set at 1.0 for the MZ twins and 0.50 for the DZ twins. The parameter *a* represents the influence of genetic component on the phenotype. The *C* latent variable represents the shared environmental component, and its path coefficient *c* influences both twins to the same extent. *E* denotes the nonshared environment. The *e* path estimate represents the influence of environmental variation that is unique to each twin. Twin studies commonly report the proportion of the total phenotypic variance accounted for by each variable. Therefore, the influence of genetic factors (a^2) is referred to as the heritability, the influence of c^2 is the shared environmental influence, and e^2 is the nonshared environmental influence. The path model makes the critical assumption that the effects of genes and the shared environment are independent. If they are correlated ("passive genotype–environment correlation"), estimates of c^2 are biased in studies of twins reared together (Jinks & Fulker, 1970).

Early twin studies were conducted to determine if genetic factors influenced traits or behaviors. Over the last quarter century, researchers have illustrated that genes influence most, if not all, behaviors and traits (Plomin, DeFries, McClearn, & McGuffin, 2000; Turkheimer, 2000) and have demonstrated the ubiquitous importance of the nonshared environmental variance component, in contrast to the relatively

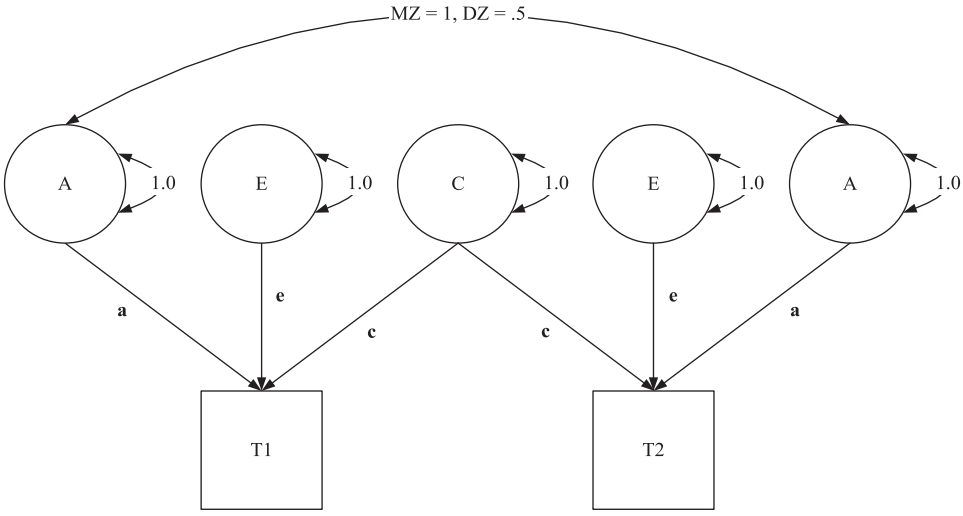

Figure 1. Structural equation model for the standard twin design.

smaller influence of the shared environmental component (Daniels & Plomin, 1985; Dunn & Plomin, 1990; Turkheimer, 2000; Turkheimer & Waldron, 2000). Apparently, most environmental influences cause siblings to be less alike.

Twin Design With Measured Shared Environments

Figure 2 is an example of a model used to study a measured C twin design (Neale & Cardon, 1992). The model is equivalent to the standard twin model described above, except that it includes a specific measure of the family environment (C_f, the f subscript noting that it is a family-level variable, jointly describing both twins). Note that assigning the measured variable the status of ''shared environmental'' is somewhat arbitrary. In designs such as we are considering here, twins are necessarily perfectly concordant for the variable in question, not as a matter of its actual composition but as a simple consequence of the research design. (It is also possible to include covariates that differ within families, such as a measure of parental treatment. These studies present somewhat different issues and will be discussed later.) Parental SES, for example, is necessarily the same for members of a twin pair assessed at the same point in time, so within the limited context of a study of twin children, it is a purely shared environmental measure. One would not want to conclude, however, that SES was in general a completely environmental variable, as it is well known that most purported measures of the environment include genetic variability (Plomin & Bergeman, 1991); twin children are just not genetically informative about their parents' SES. The measured C model can also be parameterized with an arrow between the measured environment and the latent shared environmental factor, but the fit of the two models is exactly the same.

Studies Reporting SEMs of the Measured C Design

Kendler, Neale, Kessler, Heath, and Eaves (1992a) were the first to include a measured family-level variable into a SEM of twins. (A similar model is referred to as the age correction model in Neale & Cardon, 1992.) The analyses included childhood parental loss, either through separation or death, as a ''specified'' shared environment in their univariate twin analyses of various adult psychological disorders, including depression, generalized anxiety disorder, and panic disorder. When basic twin models (Figure 1) were fit without the C_f measure, there was no significant influence of the latent shared environmental variable on the adult disorders (i.e., the confidence intervals around the estimate for the

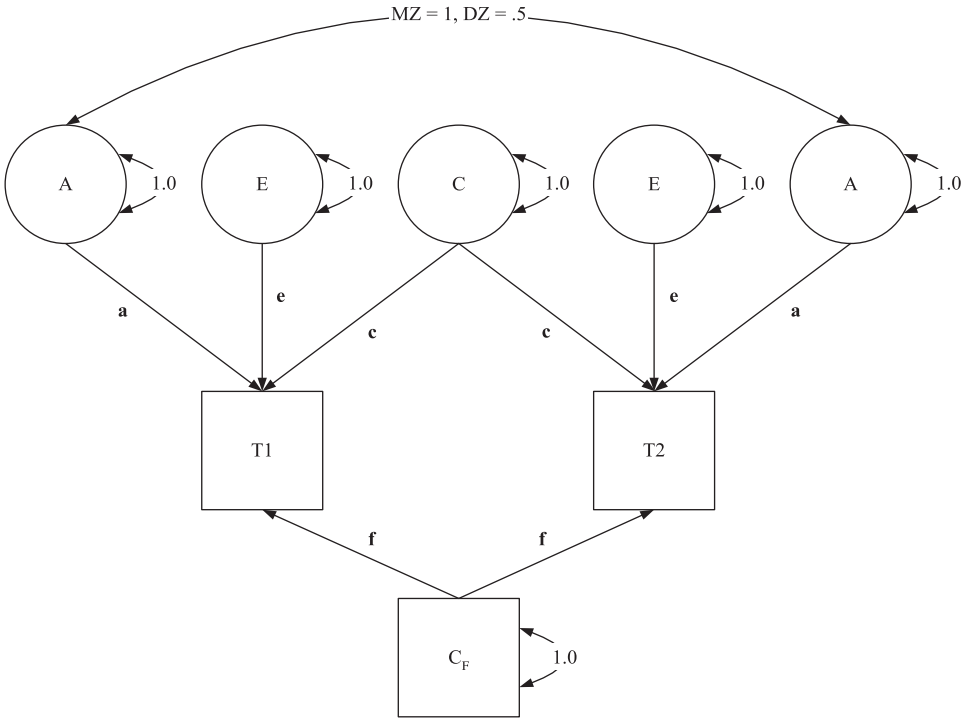

Figure 2. Structural equation model of twin design that includes a measured family-level factor.

shared environment [*c*] included zero). For example, a twin model of major depression that dropped the shared environmental parameter did not result in a significant loss in fit compared with the full twin model (Kendler, Neale, Kessler, Heath, & Eaves, 1992b). However, when the parental loss variables were included as specified shared environmental variables (Figure 2), the parameters associated with the parental loss were statistically significant for major depression, generalized anxiety disorder, and panic disorder. The authors concluded, "A model that includes parental loss as a form of 'specified' family environment shows that, *if it is truly an environmental risk factor for adult psychopathological conditions*, it can account for between 1.5% and 5.1% of the total variance in liability to these disorders. . ." (p. 109, italics added). Note that this conclusion draws attention to the essential assumption that the specified variable is purely environmental. The article went on to discuss how the association between parental loss and the adult conditions could be mediated by other environmental (family dynamics correlated with parental loss) or genetic (predisposition to poor marital processes in the parents and psychiatric problems in the offspring) factors.

As time has gone by, however, researchers have become less assiduous in their attention to the ambiguity of the shared environmental status of measured family-level variables in twin studies. Caspi, Taylor, Moffitt, and Plomin (2000) utilized a similar model to explore the influence of socioeconomic variables on emotional and behavioral problems in 2-year-old twins. In this case, basic twin models illustrated that the shared environmental latent variable accounted for 20% of the variance in behavior problems. The researchers then included a composite index of measured neighborhood factors, which accounted for 5% of the variance in behavioral problems. Although it was noted in the discussion that the approach assumes that the measure of socioeconomic deprivation is a pure environmental variable, the abstract stated, "A nationwide study of 2-year-old twins shows that children in deprived neighborhoods were at increased risk for emotional and behavioral problems *over and above any genetic liability*. The results suggest that the link between poor neighborhoods and children's mental health may be a true environmental effect. . ." (p. 338, italics added).

Jaffee, Moffitt, Caspi, Taylor, and Arsenaeault (2002) used a measured C model to study the relation between domestic violence and internalizing and externalizing problems in twin children. As in the previous study, the authors briefly cited the "possibility that the effect of domestic violence may be partially genetically mediated" (p. 1102) in the discussion section of the manuscript. In contrast, the abstract stated, "a multivariate model showed that adult domestic violence accounted for 2% and 5% of the variation in children's internalizing and externalizing problems, respectively, *independent of genetic effects*" (p. 1095, italics added).

Thapar et al. (2003) used the measured C design to explore the association between maternal smoking during pregnancy and attention deficit hyperactivity disorder (ADHD). The parameter associated with smoking during pregnancy was statistically significant. The discussion includes a careful description of the limitations of the analysis: ". . .even twin or adoption designs cannot be used to test whether maternal smoking during pregnancy has a truly causal relationship with offspring ADHD symptoms, *independent of genetic factors,* not even where maternal ADHD is assessed. . . . Thus, we are careful in stating that we observe an association between maternal smoking during pregnancy and offspring ADHD symptoms and do not conclude that this necessarily implies causality" (p. 1988, italics added). The abstract, however, makes a much stronger claim: "Maternal smoking during pregnancy appears to show an association with offspring ADHD symptoms that is *additional to the effects of genes* and not attributable to [other confounds]" (p. 1985, italics added).

In a study of the relation between domestic violence and IQ in young children, Koenen, Moffitt, Caspi, Taylor, and Purcell (2003) utilized a measured C model to demonstrate a significant relation between domestic violence and children's intelligence. In the abstract, the authors claim, "Structural equation models showed that adult domestic violence accounted for 4% of the variation, on average, in child IQ, *independent of latent genetic influences*" (p. 297, italics added). In this case, no mention was made of the assumption that the measure of adult domestic violence was a purely environmental variable.

Finally, Kim-Cohen, Moffitt, Caspi, and Taylor (2004) investigated the influence of stimulating activities in the home on a measure of cognitive resilience. The phenotype was a measure of cognitive resilience to SES (residuals cognitive ability after SES had been partialed). Although the authors stated, "A caveat is in order. Although we assume that stimulating activities is an environmental variable, it is possible that this variable is also influenced by parental IQ, which is partly heritable" (p. 662), they go on to conclude that their analysis "demonstrate[s] that the environment does play an important role in

children's cognitive resilience to SES adversity *beyond any heritable influences*'' (p. 662, italics added).

Analysis of SEMs of the Measured C Design

Our review of the published articles using the measured C twin SEM reveals that, over time, the assumption that the measured variable is a purely environmental influence on outcome has been somewhat underemphasized. Whereas the earliest articles were very cautious about the interpretation of the results, more recent papers have tended to mention the assumptions only briefly, if at all, and also to include stronger claims about the results' significance. In the following sections, we endeavor to clarify two points about SEMs of the measured C design. First, we specify the quantity estimated by the parameter associated with measured shared environmental variables in SEMs of the measured C design. Second, we show that misspecifying the action of the measured covariate as purely environmental when in fact it is not can have serious consequences for the rest of the analysis, leading to incorrect conclusions about the roles of genes and environment in the genesis of the phenotype being studied.

To answer these questions, it is necessary to derive the estimated values of the path coefficients in Figure 2 from the observed covariances among the twins' phenotypes and the observed covariates in the MZ and DZ twins. Applying standard path diagram tracing rules to Figure 1, we obtain the following system of equations:

$$\begin{aligned} r_{MZ} &= a^2 + c^2 + f^2, \\ r_{DZ} &= \tfrac{1}{2}\, a^2 + c^2 + f^2, \\ r_{COV} &= f, \\ a^2 + c^2 + e^2 + f^2 &= 1. \end{aligned} \tag{1}$$

Solving for the values of the paths, we obtain

$$\begin{aligned} a^2 &= 2(r_{MZ} - r_{DZ}), \\ c^2 &= 2r_{DZ} - r_{MZ} - r_{COV}^2, \\ e^2 &= 1 - r_{MZ}, \\ f &= r_{COV}. \end{aligned} \tag{2}$$

The path for f, the regression of the twins' phenotype on the measured environmental variable, is equal to the observed correlation between the phenotype and the environmental variable, so the genetically informative part of the analysis has no effect on it. In the genetically informative portion, the genetic and nonshared environmental parameters are unchanged from their usual forms. The shared environmental term is discounted by the proportion of the variance accounted for by the covariate. The analysis, therefore, is a way of conducting a twin analysis on a phenotype after removing the portion of the shared variance attributable to a measured variable. Comparing the results of such an analysis to an analogous twin model without the measured variable (Figure 1) tells you how much of the total shared environmental effect on the phenotype is accounted for by the measured covariate, which is the way it was interpreted in the early studies. The regression of the phenotype on the measured shared environmental variable does not control for anything, because it takes exactly the same value whether or not the biometric portion of the model is included.

Conducting this analysis on the actual twin data will help clarify the point. As an example, we used a sample of twins from the VA 30,000 data set (see Truett et al., 1994, for more details of the sample; our concerns here are methodological rather than substantive) The analyses explored years of education completed by twins and a measure of parental separation and divorce (a family-level variable because the twins are necessarily concordant for the measure). The sample included 576 MZ and 742 DZ twin pairs with complete data for both twins and marital status data for their parents. We fit the full model illustrated in Figure 2, as well as two submodels, one with only the biometric decomposition and without divorce (Figure 1), and one with divorce but without the biometric decomposition (figure not shown).

The results are shown in Table 1. The biometric decomposition of educational attainment has no effect at all on the small but significant regression of years of education on parental divorce status. The small percentage of twin covariance for educational status that can be attributed to parental divorce is simply subtracted from the shared environmental component. It bears repeating that even this designation of the divorce effect to the shared environment is arbitrary, a consequence of the fact that twins

Table 1
Parameter Estimates of Twin Models for Years of Education That Include a Measure Parental Divorce

Model	A	C	E	*f*
Full model	.66	.589	.46	– .068
Biometric only	.66	.586	.46	—
Divorce only	—	—	—	– .068

Note. The full model is presented in Figure 2. Figure 1 represents the biometric-only model. The divorce-only model is not shown.

are not genetically informative about parental characteristics that they share. Our conclusion from the analysis is that a small portion of the shared family variance in educational attainment can be explained by the observed regression of educational attainment on parental divorce, if one assumes that the effects of parental divorce are purely environmental. But that is all we can conclude. The observed regression of educational attainment on divorce is just that; the estimate does not control for genetic or environmental effects on either divorce or educational attainment.

Most of the studies we have cited using the measured C design do not report the regression of the phenotype on the measured covariate without the biometric decomposition of the phenotype; therefore, it is not possible to document that the same phenomenon occurs in the published literature. One report that does include such a regression is the paper by Jaffee et al. (2002). These authors described a series of SEM in which they dropped parameters in order to find the best model for the relation between domestic violence and offspring psychopathology. The magnitude of the parameter (f) for the regression of externalizing (.22) and internalizing (.15) on domestic violence did not change as each of the latent biometric variables (A, C, and E) were removed.

We now turn to our second point about structural equation analyses of the measured C design. In the foregoing, we analyzed the effect of simultaneously conducting a biometric decomposition of a phenotype with a regression of the phenotype on a measured family-level covariate, and concluded that the regression on the covariate is exactly the same whether or not the twin information is included. Our next concern is the converse: how does including a family-level covariate affect a biometric analysis of a phenotype? The answer, as we will see, is that rather than having no effect at all, under many circumstances the inclusion of a family-level covariate will lead to incorrect conclusions in the rest of the analysis.

In the analysis above, we showed that the variance explained by the covariate, f^2, is subtracted from the shared environmental component, c^2, when the covariate is included in the model. This attribution of the covariate to the shared environmental term is arbitrary, however, and it is simple to construct circumstances in which it is incorrect. Consider the following circumstances. A twin study of a phenotype shows that $r_{MZ} = .5$ and $r_{DZ} = .25$, resulting in a simple biometric decomposition of $a^2 = .5$, $c^2 = 0$, $e^2 = .5$. A family-level covariate describing both parents (e.g., mid-parent education) is then included in the model, and it correlates with the phenotype at $r = .5$, accounting for .25 of the variability in child phenotype. Because all of the shared variability in the phenotype has been hypothesized to be genetic, the correlation with the parental variable must be mediated genetically. However, an actual researcher would have no way of knowing this, because as we have shown, the measured C design is not informative about the biometric construction of the covariate. So, in practice, the investigator would fit the SEM specified above, and the variance attributable to the covariate would be subtracted from the c term, resulting in $a^2 = .5$, $c^2 = -.25$, $e^2 = .5$, $f^2 = .25$. The negative variance estimate for c^2 would be an obvious violation of the twin model, leading to consideration of additional parameters for genetic dominance or epistasis that would be a spurious result of falsely assuming that the pathway between the covariate and the phenotype was entirely environmental.

The important point, however, is not the particular misspecifications that would result in this or any of the many other examples that might be constructed. Rather, the crucial issue is a conceptual one about statistical modeling: The SEMs used to analyze the measured C design entail an assumption that is, on the one hand, untestable and, on the other hand, is capable of producing misleading results when it is violated. When this is combined with our first point, that the twin analysis has no effect on the estimate of the regression of the phenotype on the covariate, the wisdom of continuing to use these models is called into serious question.

Regression Analyses of Twin Data With Measured Environments

Another analytic technique, called DeFries–Fulker (DF) analysis, has also been used to analyze the measured C design (e.g., Jaffee, Caspi, Moffitt, & Taylor, 2004; Jaffee, Moffitt, Caspi, & Taylor, 2003; McGue & Lykken, 1992; Maughan, Taylor, Caspi, & Moffitt, 2004). DF analyses of twin and family data estimate biometric components without resorting to latent variable structural equation modeling. The approach was originally used to analyze samples of twins in which one twin is selected as a proband (DeFries & Fulker, 1985), and has since been expanded for use in random samples of twins (Cherny, DeFries, & Fulker, 1992; Labuda, DeFries, & Fulker, 1986; Rodgers & McGue, 1994).

Before examining the slightly more complex form of DF analysis that recent studies have utilized for

the measured C design, it will be useful to review the basics of why the method works for the simple case. In DF analysis, one first double enters the twins as two variables in the data set and regresses one on the other. In the dependent variable, which we will refer to as the double-entered vector, the phenotype of an arbitrary member of each pair is included first, followed by the second member of the pair as a separate observation. The independent variable, which we will call the reverse-entered vector, is the same as the first except that the order of the twins is reversed within each pair, so across the two vectors each twin is paired twice with his or her cotwin, once in each order. Then, a regression model is estimated in which the double-entered vector is regressed on the reverse-entered vector, the zygosity of the pair expressed as the degree of genetic relationship (r_g, 1.0 for MZ pairs; .5 for DZ pairs), and the interaction between the two.

$$\hat{y}_{ij} = b_0 + b_1 y_{ji} + b_2 r_g + b_3 y_{ji} r_g. \tag{3}$$

Although recent applications of DF analysis have utilized advances in statistical software to calculate more accurately the standard errors of the estimated parameters in light of the double entering (Kohler & Rodgers, 2001), the most important issues involving the analysis are not estimation and statistical inference, but rather substantive interpretation of the results.

The simplest way to derive the expectations of the DF model is to consider the simple regression of one (double-entered) twin on the other, separately for the MZ and DZ pairs. The covariance between the double- and reverse-entered vectors is simply the covariation between members of twin pairs. In the MZ pairs, the pair covariance is equal to

$$a^2 + c^2 \tag{4}$$

and the regression of one of the double-entered vectors on the reverse-entered vector is equal to this covariance over the phenotypic variance, that is,

$$\frac{a^2 + c^2}{a^2 + c^2 + e^2}. \tag{5}$$

In the DZ twins, the equivalent regression is equal to

$$\frac{\frac{1}{2}a^2 + c^2}{a^2 + c^2 + e^2}. \tag{6}$$

In the full model encompassing both kinds of twins, including the r_g term for zygosity and the interaction between the reverse-entered vector and r_g, the two regressions are expressed as a main effect b_1 common to both zygosities and an in interaction term b_3 that expresses the difference between them. So we have

$$\begin{aligned} MZ &: \frac{a^2 + c^2}{a^2 + c^2 + e^2} = b_1 + 1.0 b_3, \\ DZ &: \frac{\frac{1}{2}a^2 + c^2}{a^2 + c^2 + e^2} = b_1 + .5 b_3, \end{aligned} \tag{7}$$

and therefore

$$\begin{aligned} \frac{c^2}{a^2 + c^2 + e^2} &= b_1, \\ \frac{a^2}{a^2 + c^2 + e^2} &= b_3, \end{aligned} \tag{8}$$

which shows that the b_1 term estimates the standardized shared environmental component, and the interaction term (b_3) estimates the standardized additive genetic component.

Published Regression Analyses of Measured C Design

It has long been known that the DF model can be extended in a variety of ways. Rowe and Waldman (1993) suggested that researchers add a shared environmental measure (F) to the regular DF equation to explore how specific measures of the shared environment could mediate the influence of the latent factor.

$$y_{ij} = b_0 + b_1 y_{ji} + b_2 r_g + b_3 y_{ji} r_g + b_4 F \sum_{i=1}^{n} X_i. \tag{9}$$

Rowe and Waldman (1993) stated correctly that when the measured shared environmental variable is added to the basic DF model, the influence of the shared environmental parameter (b_1) should decrease because the "measured variable 'takes up' some of the explanatory variance associated with abstract shared environmental influence" (p. 366). However, the authors did not analyze a measured C data set in their review chapter.

McGue and Lykken (1992) used a sample of twins to explore genetic and environmental contributions to divorce. The authors explained their use of regression analyses as follows: "In order to consider simultaneously the influence of all family background data, we used a logistic function of birth cohort, zygosity, co-twin's divorce status, parents' divorce status, spouse's parents' divorce status, and all two-way interaction terms to predict divorce risk" (p. 37). Note that parents' divorce status was necessarily shared by both twins, but the twins could differ on the spouse's parents' divorce status. The regression analyses resulted in a significant parameter for the interaction between zygosity and the cotwin's divorce status and the main effects of the other predictors. The authors concluded, "spouse's family background of divorce and respondent's

family background of divorce contributed *independently* to the prediction of marital dissolution'' (p. 370, italics added).

Recent use of the expanded DF model has also focused on family-level factors. Jaffee et al. (2003) explored the impact of father absence and parental antisocial behavior (ASB) on young children's ASB. The regression analyses added measures of father's ASB, mother's ASB, father's presence, and the interaction of father's presence and ASB to the standard DF regression analysis. Each of these familial variables was necessarily shared by the twin children. The full DF model resulted in significant regression coefficients associated with the interaction between genetic relatedness and twin's ASB, the main effects of father's ASB and mother's ASB, and the interaction between father's ASB and father's presence. Based on the results of the regression models, the authors concluded: ''Children experience a double whammy of risk for antisocial behavior. They are at genetic risk because antisocial behavior is highly heritable. In addition the same parents who transmit genes also provide the children's rearing environment. We found that a father's antisocial behavior accounted for his children's behavior problems *independent of any genetic risk* he may have imparted. . .'' (p. 120, italics added).

The relation between parental maltreatment and young children's ASB has also been explored using expanded DF analysis (Jaffee et al., 2004). Although parental maltreatment can differ among twins, these analyses appeared to enter the variable as a single family-level variable. The regression weights associated with physical maltreatment and the interaction of genetic relatedness and the cotwin's ASB were significant. Consequently, the authors concluded, ''the effect of physical maltreatment was significant *after controlling for any genetic transmission* of antisocial behavior. . .'' (p. 50, italics added).

Finally, the relation between smoking during pregnancy and childhood conduct problems was explored using DF analyses (Maughan et al., 2004). Although many studies have documented an association between smoking during pregnancy and offspring difficulties, a number of researchers have questioned whether genetic factors may confound the intergenerational relation (e.g., D'Onofrio et al., 2003; Fergusson, 1999; Sexton, Fox, & Hebel, 1990; Silberg, Parr, Neale, Rutter, Angold, & Eaves, 2003; Thapar et al., 2003). To account for genetic factors and other processes that may mediate the intergenerational association, Maughan et al. (2004) included smoking during pregnancy (as a family-level variable represented by a set of dummy codes to account for different levels of smoking), other potential confounding variables (maternal ASB, paternal ASB, maternal depression, and SES disadvantage), and the basic variables from the DF. Once again, the smoking during pregnancy and the confounding variables were shared influences by definition. The full regression model illustrated that the parameters associated with smoking during pregnancy were reduced substantially when the confounding and DF variables were included but were still sizeable. The paper concludes, ''Around half of the observed association between prenatal smoking and young children's conduct problems was attributable to correlated genetic effects. But the results were also clear in showing that, *even after controlling for genetic influences*, prenatal smoking continued to be significantly linked to children's behavior outcomes'' (p. 841, italics added).

Analytical Review of DF Analysis of Measured C Design

As was the case with SEM analysis of the design, the important question is how to interpret the parameter associated with the measured family-level ''environmental'' variable (b_4). It should be noted from the outset that adding a covariate shared by both twins (by definition) to a model in which the phenotypes of the twins are already regressed on each other seems to be a paradoxical thing to do. Conceptually, the reason double-entered twin vectors are regressed on each other in DF analysis is as a way of estimating the phenotypic variance the twins have in common, with the difference between the magnitude of this shared variance in the MZ and DZ twins estimating the heritability, and the residuals from the regression estimating the variance not shared by members of the twin pair, or e^2. So by adding the shared family-level covariate to this model, one is attempting to predict shared twin variability in a model that already has a term specifically designed to account for all of it. What could be left to estimate? It turns out that the inner workings of DF analysis are somewhat more complex than the results it produces, so to understand exactly what takes place we need to consider the details of the regressions on which DF analysis is based.

In a DF analysis of the measured C design, one regresses the double-entered vector on the reverse-entered vector and on a family-level covariate. It is simplest to begin by considering a sample of MZ twins, and a measured covariate that has been standardized to a mean of zero and a variance of one. Figure 3 is a path diagram of the model. The double-

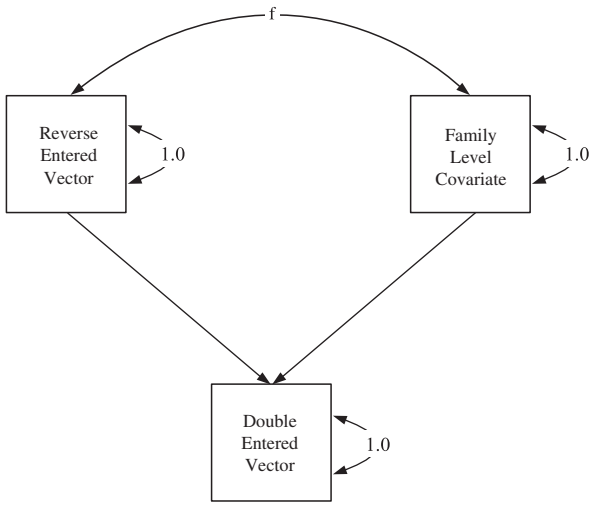

Figure 3. Structural equation model of DF analysis.

entered vector, reverse-entered vector, and measured family covariate have a variance covariance matrix equal to

$$\begin{bmatrix} \sigma_y^2 & & \\ a^2+c^2+f^2 & \sigma_y^2 & \\ f & f & 1.0 \end{bmatrix}. \tag{10}$$

Applying standard formulas for the values of bivariate regression coefficients, we find that the regression of the double-entered vector on the reverse-entered vector is given by

$$\frac{a^2+c^2}{\sigma_y^2-f^2}, \tag{11}$$

which is the proportion of the y variance not explained by the covariate that is shared between the twins. The regression of the double-entered vector on the family-level covariate is given by

$$f\left(\frac{\sigma_y^2}{\sigma_y^2-f^2}\right)(1-r_{\mathrm{I}}), \tag{12}$$

which is the simple regression of the phenotype on the measured family variable, rescaled by the proportion of the phenotypic variance not attributable to the covariate that is not shared by the twins. As was the case with SEM models of the measured C design, the DF estimate of the regression of the phenotype on the covariate does not control for anything, genetic or environmental, although in DF analysis it is linearly rescaled.

To understand why this result occurs, consider the simple regression of a set of double-entered twins on the reverse-entered vector, without the covariate. If the twin scores are expressed as deviations from the population mean, the score of the ith member of the jth pair can be expressed in terms of the pair mean and a deviation from the pair mean:

$$\overline{y_{.j}} + \frac{(y_{1j}-y_{2j})}{2}. \tag{13}$$

This twin is to be predicted from its cotwin, with deviation $(y_{2j}-y_{1j})/2$ from the pair mean. The regression coefficient, as we have seen, is equal to the intraclass correlation r_{I}. The residual from this regression, therefore, is equal to

$$\begin{aligned} &\overline{y_{.j}} + \frac{(y_{1j}-y_{2j})}{2} - r_{\mathrm{I}}\left[\overline{y_{.j}} + \frac{(y_{2j}-y_{ij})}{2}\right] \\ &\quad = (1-r_{\mathrm{I}})\overline{y_{.j}} + \frac{(1+r_{\mathrm{I}})}{2}(y_{1j}-y_{2j}) \end{aligned} \tag{14}$$

and the mean of residuals for twin pair j is simply equal to

$$(1-r_{\mathrm{I}})\overline{y_{.j}}. \tag{15}$$

Although the procedure of double-entry regression succeeds in estimating r_{I}, it does not control for anything: in the residuals from the double-entered regression, the pair means, which are the only portion of the data with which a family-level variable can possibly covary, are correlated 1.0 with the pair means in the original data. So a measured covariate with covariation f with phenotype in the original scores has the same relation with phenotype in the residuals, except that it is rescaled by the proportion of nonshared variance. This coefficient is zero when either f or the nonshared variance is zero.

The uninformativeness of the f parameter is not redeemed by consideration of the MZ and DZ twins. The expectation of $1-r_{\mathrm{I}}$ in Equation (12) in terms of ACE parameters is different in the MZ and DZ twins. In the MZ twins, $1-r_{\mathrm{I}}$ is equal to e^2, whereas in the DZ twins it is equal to $\frac{1}{2}a^2+e^2$. Because the expectations for the main effects of the covariate are different in the MZ and the DZ pairs, one could include both a main effect of the covariate and an interaction with genetic relatedness (r_{g}), although the interaction has not generally been included in actual analyses. If it were however, the main effect of the covariate would have the expectation

$$f\left(\frac{\sigma_y^2}{\sigma_y^2-f^2}\right)(a^2+e^2), \tag{16}$$

and the interaction with zygosity would have the expectation

$$-f\left(\frac{\sigma_y^2}{\sigma_y^2-f^2}\right)(a^2), \tag{17}$$

which still only reexpresses the original regression parameter in terms of a different variance.

This whole point is crucial to our argument, so we will take the time to repeat it at a conceptual level before proceeding to an example from the education and divorce data in the VA 30,000 data set. When twin children vary in some outcome, a sibling model allows us to partition the variability in outcome into a portion varying between families (essentially the variance of the family means) and a portion varying within them (essentially the variance of children around their family mean). Using twin zygosity allows us to further partition the between and within variances into ACE components. The variance of a family-level covariate, however, which is necessarily equal for members of a twin pair, is by definition completely shared in both MZ and DZ twins (there is no variation around the family mean), and therefore cannot be further partitioned into genetic and environmental components. It does not make sense to ask whether the deviations of twins' education scores from the pair means covary with their parents' divorce status, because if one twin has a deviation of $+x$, the other twin has a deviation of $-x$, whereas their parent's marital status is exactly the same. On the bottom line, a variable that has no within-family variation cannot covary with the within-family variation in another variable.

Working the analysis in the divorce-years of education data will help clarify the point. We double entered the twins, and then conducted three analyses: (1) a regression of one double-entered twin vector on the other; (2) a regression of the twin vector on a standardized dummy variable coding for divorce; and (3) a regression in which the twin vector was regressed simultaneously on the reverse-entered vector and the dummy variable for divorce.

Analysis 1 results in a regression equation of

$$Twin_1 = 5.02 + .655\,Twin_2. \tag{18}$$

The regression coefficient of .655 is the intraclass correlation, or the proportion of the variance that is shared by twin pairs. In Analysis 2, the simple regression of standardized divorce status on the double-entered twins results in

$$Educ = 14.7 - .19Div. \tag{19}$$

Therefore, without controlling for anything, each standard deviation of divorce status is associated with a decrease of about a fifth of a year in the children's education. Then, in Analysis 3, the double-entered twin vector is regressed on the reverse-entered vector and divorce status, with the result

$$Educ_1 = 5.11 + .65Educ_2 - .065Div. \tag{20}$$

At first glance, it might seem as though the reduction in the magnitude of the effect of divorce status between Analyses 2 and 3 ($-.19$ to $-.065$) is a result of statistical control of shared genetic and/or environmental effects achieved by including the reverse-ordered twin vector, but it is not. In fact, the coefficient of .065 is simply the product of the original regression coefficient for divorce status in Analysis 2 ($-.189$), and the proportion of phenotypic variance not attributable to the covariate that is not shared between siblings. (The apparent similarity between the coefficient of .655 in Equation (18) and .065 in Equation (20) is coincidental.) The variance of education in the combined sample is 8.00, of which $.189^2 = .036$ is accounted for by divorce status. The intraclass correlation for education equals .655. Therefore, the coefficient for standardized divorce status in the model including the reverse-entered twin vector will be equal to

$$-.189\left(\frac{8.0(1-.655)}{8.0-.0357}\right) = -.065. \tag{21}$$

This is exactly the value estimated in Analysis 3.

Discussion

Our review of the literature on the measured C design and the various statistical and interpretive approaches to it was motivated by a number of factors. First, we agree with researchers in behavior genetics (e.g., Eaves, Last, Martin, & Jinks, 1977; Plomin, DeFries, & Loehlin, 1977; Rutter et al., 1997; Rutter & Silberg, 2002; Scarr & McCartney, 1983) and developmental psychology (e.g., Booth, Carver, & Granger, 2000; Collins, Maccoby, Steinberg, Hetherington, & Bornstein, 2000) that studies need to explore how genetic and environmental factors act and interact. Although the debate over whether genetic factors influence individual differences on behavioral characteristics is over (Turkheimer, 2000), the need to explore the mechanisms between what are considered to be environmental risk factors and individual adjustment is paramount. Our own research program, outlined below, is representative of our desire to explore such processes.

Second, we understand the fundamental motivation behind both the older and more recent reports describing twin studies with measured family-level factors. The theoretical rationale outlined by many of the researchers (clearly expressed in Jaffee et al., 2004) needs to be appreciated by all social scientists, because genetic factors may indeed confound statistical associations between risk factors and child adjustment. Third, we are sympathetic with the desire to respond to researchers who have claimed that shared environmental factors are unimportant

(e.g., Plomin & Daniels, 1987; Harris, 1998; Rowe, 1994), based, in part, on the null findings for shared environmental influences in classical twin designs. Properly qualified and analyzed, standard twin studies in general do have the potential to demonstrate the importance of both genetic and environmental factors and to show how they can confound statistical associations between purported risk factors and child adjustment.

Nevertheless, it must be recognized that twin studies with family-level environmental characteristics are unable to explore genetic mediation of associations between the family-level variables and individual adjustment, and moreover that widely employed methods for analyzing the design can lead to inaccurate results. Many of the conclusions that have been reached about the implications of such studies have been overstated. We are not claiming that all statistical associations that have been explored in the measured C reports are attributable to genetic factors; rather, we conclude that the claims made in the recent literature cannot be supported by the methodology and analyses that were used. In the remainder of this paper, we apply what we have learned about the design, analysis, and interpretation of the measured C design to make some recommendations for future progress in the area.

SEM analyses of family-level variables and outcomes in twins only provide an estimate of the uncontrolled phenotypic relation between them, and cannot determine whether the effect of the measured variable is genetic or environmental. Therefore, researchers should desist from describing the results of such analyses as showing the influence of an environment ''independent of genetic factors'' (e.g., Thapar et al., 2003, p. 1988), which represents an assumption of the model, not a testable hypothesis. In general, we recommend that future analyses refrain from estimating the biometric parameters (a random variance parameter can be estimated to take into account that there are two children per family; see the discussion of hierarchical linear models) when shared environmental influences are included in a twin data set. Doing so only fosters the false impression that something is being controlled in the estimation of the relation between the family-level variable and the twin outcome, and as we have shown inclusion of the covariate can produce seriously misleading results in the biometric analysis.

Expanded DF analyses do not overcome the limitations of the SEM approach. In some ways, DF analysis has proved even more misleading, because the rescaling of the regression parameter associated with the covariate in terms of the nonshared variance when the reverse-entered twins are included in the regression has the effect of reducing its magnitude from the value in the simple regression without the reverse-entered vector. This reduction has convinced researchers that they are controlling for genetic or environmental factors when in fact they are not.

Hierarchical Linear Models

A much greater degree of clarity about the structure of this research design can be achieved by turning to a more modern form of regression than DF analysis: hierarchical linear modeling (HLM). For a review of HLM see Raudenbush and Bryk (2002), and for applications of HLM to family data see van den Oord (2001) and Guo and Wang (2002). There is increasing recognition in the field that HLM is ideally suited to the analysis of family data, in which sibling or twin children are nested within families, and therefore correlated with each other. Therefore, we can say for the ith child in the jth family, the predicted score is given by an intercept for the family b_j and random variation of children in the family around the family intercept:

$$y_{ij} = b_j + \sigma_w^2. \tag{22}$$

The family intercept b_j can then be modeled as a function of covariates like maternal smoking or parental education, plus a second error term describing the variability of the family means around the grand mean:

$$b_j = b_0 + b_C COV + \sigma_B^2. \tag{23}$$

Several conceptual benefits of this approach are immediately apparent. First of all, the approach provides a name for the type of covariate we have been discussing, one that is necessarily shared by both members of a twin pair: it is a family level, or level 2, variable, which describes the data at the higher level of nesting rather than at the level of individual participants. It is exactly like a variable describing the classroom in the paradigmatic HLM model in which individual students are nested within classrooms. Second, in an HLM structure it becomes very clear that a family-level variable can only be used to predict variability in y at the family level, that is, the variability of the family means around the grand mean. Neither genetic nor environmental variability in y can be controlled.

Other Extensions of DF Analysis

Apparently, the recent analyses of the measured C design using DF analysis have confused two things that DF analysis can do with something it cannot.

One common extension of DF analysis, explored most fully by Rodgers and colleagues, can be applied to research designs including a covariate that varies both between and within families (Rodgers, Rowe, & Li, 1994; Rodgers, Rowe, & May, 1994), for example, a measure of maternal interaction that can be measured separately for each of the children within a family (e.g., Caspi et al., 2004), or a measure of a family-level construct like SES or parental marital status that is expressed in terms of the experience of individual siblings within a family. In this situation, the within-family regression of child outcome on the covariate does control for genetic and environmental between-family covariation. This is clearest in the case of identical twins discordant for a predictor within the family: perhaps they were differentially exposed to an environmental toxin. The within-family regression of an outcome on the predictor describes the relation between the two independent of genetic and shared environmental variance, because both of these equal zero within MZ twin pairs. In an HLM model of the design, one would simply include the covariate at both the between- and within-family levels of the model and interpret the resulting parameters separately (Lynch et al., submitted; Mendle, Turkheimer, D'Onofiro, Lynch, & Emery, submitted).

The other applicable extension of DF analysis, less commonly used, is to use it to detect gene by environment interaction (Rodgers, Rowe, & Li, 1994; Rowe, Almeida, & Jacobson, 1999; Rowe, Jacobson, & van den Oord, 1999). In this analysis, one adds to the basic DF model containing a two-way interaction between genetic relatedness and the reverse-entered twin vector a three-way interaction between this term and the measured covariate. The three-way interaction estimates the extent to which the two-way interaction (i.e., the heritability of phenotype) varies as a function of the covariate, which is one way of characterizing gene by environment interaction. In HLM, such models are referred to as heteroscedasticity models, because they involve modeling the magnitude of the between- and within-family error terms as a function of a covariate. See Guo and Wang (2002) for examples of gene by environment interaction analyses via HLM, and Purcell (2002) for related analyses using SEM.

Extended Twin Designs

Thapar et al. (2003) lament the limitations of standard twin models for exploration of intergenerational relations, in particular offspring characteristics related to smoking during pregnancy: ''Strictly speaking, teasing apart genetic and environmentally mediated effects would require examining offspring exposed to maternal smoking in utero where the intrauterine environment was provided by a non-genetically related 'mother' (surrogate; that is a 'before-birth' adoption design). That is clearly not a feasible design in humans'' (p. 1988).

We certainly agree that a maternal smoking study with ''experimental control'' is impossible, but there are designs that can separate genetic and environmental processes in the consequences of maternal behavior. The main limitation of the twin studies reviewed here is that, like smoking in women pregnant with twins, there is by definition no genetic or nonshared environmental variation in the parental measure. One way to overcome this limitation is to include measures of the twins' parents along with family-level environmental factors and twin characteristics (Eaves et al., 1978), a design referred to as the twin-family design. Genetic and shared environmental variation in the family-level variable is based on the biometric parameters of the offspring's phenotype, which are assumed to be the same in the parental generation (e.g., Kendler et al., 1996; Meyer et al., 2000; Taylor, McGue, & Iacono, 2000). The strength of this design lies in its ability to estimate environmental effects while controlling for genetic effects on both the parents and children (Rutter et al., 1997), but it also entails some major assumptions and limitations (D'Onofrio et al., 2003; Rutter et al., 2001). The same genetic and environmental factors are assumed to influence child characteristics and parental behavior (requiring the same phenotype to be measured in both generations), and twin-family models with only one parental characteristic may underestimate the noncausal, intergenerational genetic pathway (Kendler et al, 1996; Meyer et al., 2000).

Instead of including parents of twins as in the twin-family approach, twin studies can also be expanded by adding the children of twins, referred to as the children of twins (CoT) design. The major advantage of the CoT Design is that it provides genetic and environmental variation in family-level risk factors, because unlike parents of twins, twin parents can vary in family-level variables like marital status and smoking during pregnancy. Therefore, the statistical association between a family-level variable and offspring adjustment can be decomposed into environmental processes specifically related to the risk factor (consistent with a causal hypothesis), environmental influences common to both twins and the offspring (consistent with the selection hypothesis), and genetic factors (also

consistent with the selection hypothesis) that are shared by the parents and the offspring (reviews in D'Onofrio et al., 2003, submitted-a; Gottesman & Bertelsen, 1989; Heath et al., 1985; Rutter et al., 2001).

Our laboratory is actively engaged in CoT analyses of genetic and environmental pathways from parental behavior to outcomes in children. One analysis of the association between father absence and age of menarche in the female offspring concluded that the phenotypic relation is not due to environmental processes related to the risk factor; rather familial confounds (either genetic or common environmental) mediate the relation (Mendle et al., submitted). In contrast, analyses using the CoT design suggest that the association between harsh punishment and offspring drug, alcohol, and behavioral problems is due to environmental processes specifically associated with the parenting behavior (Lynch et al., submitted). Analyses of the consequences of parental divorce have also illustrated that the processes underlying intergenerational relations depend on the specific association being explored. Whereas the results of CoT analyses of parental divorce are consistent with a causal relation with lifetime history of psychopathology (D'Onofrio et al., submitted-a), some life course patterns associated with parental marital dissolution, such as cohabitation, appear to be due to genetic and environmental confounds (D'Onofrio et al., submitted-b). The limitations of the CoT design include relatively low statistical power and difficulties in accounting for effects arising in the spouses of twins (reviews in D'Onofrio et al., 2003; Heath et al., 1985; Rutter et al., 2001).

Measured Genotype Designs

Some very interesting recent studies have reported interactions between measured family-level risk factors and measured genotypes (i.e., DNA was collected and analyzed) using singletons (Caspi et al., 2002, 2003) or twins (Foley et al., 2004). These studies investigated the main effect of the family-level measure, the main effect of the measured genotype, and the interaction of the two. These studies are able to test the hypothesis that the expression of the measured genotype depends on the family-level covariate. As the interaction between the measured genotype and the family-level covariate does not depend on the partitioning of genetic and shared environmental variability in the covariate, the criticisms we have outlined in this paper do not apply to this design, although researchers must be cautious about simply assuming that family-level variables are environmental in origin. For an analytical approach that incorporates gene by environment interaction and gene–environment correlation, see Eaves and Erkanli (2003) and Eaves, Silberg, and Erkanli (2003).

Combining Behavior Genetic Designs

Tests of causal effects of parenting behavior on children are greatly strengthened when several family designs are combined, with their varied strengths and weaknesses (Rutter et al., 2001). For example, adoption studies represent a powerful method for exploring these processes (e.g., Plomin, 1995). The adoption design also includes major assumptions (Rutter et al., 2001), but combining the adoption and CoT designs could take advantage of the unique advantages of each, while limiting the assumptions when each is analyzed separately. The CoT design can also be combined with traditional studies of twins as children to explore both passive and active rGE (Neiderhiser et al., 2004).

Studies using many levels of genetic relatedness among family members also have the power to discriminate genetic from environmental transmission. For example, a number of investigators have used the family data available in the National Longitudinal Survey of Youth to study a variety of normative and pathological developmental processes (Rodgers, Rowe, & Buster, 1999; Rodgers, Buster, & Rowe, 2001). Another extensive combination of different behavior genetic designs is the "Stealth Model" developed by Eaves and colleagues to explore the intergenerational transmission of religious practices, personality characteristics, and body mass index (e.g., D'Onofrio et al., 1999; Eaves, Heath, Martin, Maes, et al., 1999; Eaves, Heath, Martin, Neale, et al., 1999; Kirk et al., 1999; Lake, Eaves, Maes, Health, & Martin, 2000; Maes, Neale, & Eaves, 1997; Truett et al., 1994). The design can also be used explore bivariate intergenerational associations (Maes, Neale, Martin, Heath, & Eaves, 1999). The model uses upwards of 30 different family relationships that vary in their degree of environmental and genetic relatedness to test for the mechanisms involved in intergenerational associations while including estimates of assumptions that hinder many behavior genetic designs.

Conclusion

Differentiating actual social causation (i.e., direct environmental influences) from genetic or environmental confounds remains one of the fundamental problems facing the social sciences. Conclusions

about causal mechanisms linking environmental risk factors to childhood adjustment are crucial because they inform both public policy and personal decisions about persistent and vexing choices, like parental divorce (Amato, 2000; D'Onofrio et al., submitted-a), father absence (Jaffee et al., 2003; Mendle et al., submitted), parenting practices (Baumrind, Larzelere, & Cowan, 2002), parental maltreatment (Jaffee et al., 2004; Lynch et al., submitted), parental psychopathology (Gottesman, 1991; Jacobs et al., 2003), prenatal smoking and drug use (Fergusson, 1999), social support (House, Landis, & Umberson, 1988), and poverty (Dohrenwend, 1992).

Evaluating potential causal mechanisms without the use of randomized experiments is extremely difficult (Cook & Shadish, 1994). Merely correlating environmental risk factors with child outcomes using traditional family studies is of limited utility because of ubiquitous environmental and genetic confounding factors. Therefore, multiple research strategies and designs are required, especially genetically informed approaches (Rutter et al., 2001). Although the current review has focused on the limitations of one widely used behavior genetic design, we remain excited by the potential of the various extended family designs to advance the methodology of twin studies and the empirical science of parenting and child development. Nevertheless, vigilance regarding the assumptions and limitations of research designs and statistical analyses will remain the sine qua non of continued empirical progress.

References

Amato, P. R. (2000). The consequences of divorce for adults and children. *Journal of Marriage and the Family, 62*, 1269–1287.

Baumrind, D., Larzelere, R. E., & Cowan, P. A. (2002). Ordinary physical punishment: Is it harmful? *Psychological Bulletin, 128*, 580–589.

Booth, A., Carver, K., & Granger, D. A. (2000). Biosocial perspectives on the family. *Journal of Marriage and the Family, 62*, 1018–1034.

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science, 297*, 851–854.

Caspi, A., Moffitt, T. E., Morgan, J., Rutter, M., Taylor, A., Arseneault, L., et al. (2004). Maternal expressed emotion predicts children's antisocial behavior problems: Using monozygotic-twin differences to identify environmental effects on behavioral development. *Developmental Psychology, 40*, 149–161.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrinton, H., et al. (2003). Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science, 301*, 386–389.

Caspi, A., Taylor, A., Moffitt, T. E., & Plomin, R. (2000). Neighborhood deprivation affects children's mental health: Environmental risks identified in a genetic design. *Psychological Science, 11*, 338–342.

Cherny, S. S., DeFries, J. C., & Fulker, D. W. (1992). Multiple regression analysis of twin data: A model-fitting approach. *Behavior Genetics, 22*, 489–497.

Collins, W. A., Maccoby, E. E., Steinberg, L., Hetherington, E. M., & Bornstein, M. H. (2000). Contemporary research on parenting: The case for Nature and Nurture. *American Psychologist, 55*, 218–232.

Cook, T. D., & Shadish, W. R. (1994). Social experiments: Some developments over the past fifteen years. *Annual Review of Psychology, 45*, 545–580.

Daniels, D., & Plomin, R. (1985). Differential experience of siblings in the same family. *Developmental Psychology, 21*, 747–760.

D'Onofrio, B. M., Eaves, L. J., Murrelle, L., Maes, H. H., & Spilka, B. (1999). Understanding biological and social influences on religious affiliation, attitudes and behaviors: A behavior-genetic perspective. *Journal of Personality, 67*, 953–984.

D'Onofrio, B. M., Turkheimer, E., Eaves, L. J., Corey, L. A., Berg, K., Solaas, M. H., et al. (2003). The role of the Children of Twins design in elucidating causal relations between parent characteristics and child outcomes. *Journal of Child Psychology and Psychiatry, 44*, 1130–1144.

D'Onofrio, B. M., Turkheimer, E. N., Emery, R., Slutske, W., Heath, A., & Martin, N. (in press a). A genetically informed study of marital instability of offspring psychopathology? *Journal of Abnormal Psychology.*

D'Onofrio, B. M., Turkheimer, E. N., Emery, R., Slutske, W., Heath, A., & Martin, N. (in press b). A genetically informed study of the association between parental divorce and offspring life course patterns. *Developmental Psychology.*

DeFries, J. C., & Fulker, D. W. (1985). Multiple regression analysis of twin data. *Behavior Genetics, 15*, 467–473.

Dohrenwend, B. P., Levav, I., Shrout, P. E., Schwartz, S., Navch, G., Link, B. G., et al. (1992). Socioeconomic status and psychiatric disorders: The causation-selection issue. *Science, 255*, 946–952.

Dunn, J., & Plomin, R. (1990). *Separate lives: Why siblings are so different?* New York: Basic Books.

Eaves, L. J. (1982). The utility of twins. In V. E. Anderson, W. A. Hauser, J. K. Penry, & C. G. Sing (Eds.). *Genetic basis of the epilepsies* (pp. 249–276). New York: Raven.

Eaves, L., & Erkanli, A. (2003). Markov Chain Monte Carlo Approaches to analysis of genetic and environmental components of human developmental change and G × E interaction. *Behavior Genetics, 33*, 279–299.

Eaves, L., Heath, A., Martin, N., Maes, H., Neale, M., Kendler, K., et al. (1999). Comparing the biological and cultural inheritance of personality and social attitudes in

the Virginia 30,000 study of twins and their relatives. *Twin Research*, *2*, 62–80.

Eaves, L. J., Heath, A. C., Martin, N. G., Neale, M. C., Meyer, J. M., Silberg, J. L., et al. (1999). Biological and cultural inheritance of stature and attitudes. In C. R. Cloninger (Ed.). *Personality and psychopathology* (pp. 269–308). Washington, DC: American Psychiatric Association.

Eaves, L. J., Last, K., Martin, N. G., & Jinks, J. L. (1977). A progressive approach to non-additivity and genotype-environmental covariance in the analysis of human differences. *British Journal of Mathematical and Statistical Psychology*, *30*, 1–42.

Eaves, L. J., Last, L. A., Young, P. A., & Martin, N. B. (1978). Model-fitting approaches to the analysis of human behavior. *Heredity*, *41*, 249–320.

Eaves, L., Silberg, J., & Erkanli, A. (2003). Resolving multiple epigenetic pathways to adolescent depression. *Journal of Child Psychology and Psychiatry*, *44*, 1006–1014.

Fergusson, D. M. (1999). Prenatal smoking and antisocial behavior. *Archives of General Psychiatry*, *56*, 223–224.

Foley, D. L., Eaves, L. J., Wormley, B., Silberg, J. L., Maes, H. H., Kunh, J., et al. (2004). Childhood adversity, monoamine oxidase A genotype, and risk for conduct disorder. *Archives of General Psychiatry*, *61*, 738–744.

Gottesman, I. I. (1991). *Schizophrenia genesis: The origins of madness*. New York: Freeman & Company.

Gottesman, I. I., & Bertelsen, A. (1989). Confirming unexpressed genotypes for schizophrenia. *Archives of General Psychiatry*, *46*, 867–872.

Guo, G., & Wang, J. (2002). The mixed or multilevel model for behavior genetic analyses. *Behavior Genetics*, *32*, 37–49.

Harris, J. R. (1998). *The nurture assumption: Why children turn out the way they do*. London: Bloonsbury.

Heath, A. C., Kendler, K. S., Eaves, L. J., & Markell, D. (1985). The resolution of cultural and biological inheritance: Informativeness of different relationships. *Behavior Genetics*, *15*, 439–465.

House, J. S., Landis, K. R., & Umberson, D. (1988). Social relationship and health. *Science*, *241*, 540–545.

Jacobs, T., Waterman, B., Heath, A., True, W., Bucholz, K. K., Haber, R., et al. (2003). Genetic and environmental effects on offspring alcoholism: New insights using an offspring-of-twins design. *Archives of General Psychiatry*, *60*, 1265–1272.

Jaffee, S. R., Caspi, A., Moffitt, T. E., & Taylor, A. (2004). Physical maltreatment victim to antisocial child: Evidence of an environmentally mediated process. *Journal of Abnormal Psychology*, *113*, 44–55.

Jaffee, S. R., Moffitt, T. E., Caspi, A., & Taylor, A. (2003). Life with (or without) father: The benefits of living with two biological parents depend on the father's antisocial behavior. *Child Development*, *74*, 109–126.

Jaffee, S. R., Moffitt, T. E., Caspi, A., Taylor, A., & Arseneault, L. (2002). Influence of adult domestic violence on children's internalizing and externalizing problems: An environmentally informative twin study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*, 1095–1103.

Jinks, J. L., & Fulker, D. W. (1970). A comparison of the biometrical genetical MAVA and classical approaches to the analysis of human behavior. *Psychological Bulletin*, *73*, 311–349.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992a). Childhood parental loss and adult psychopathology in women. *Archives of General Psychiatry*, *49*, 109–111.

Kendler, K. S., Neale, M. C., Kessler, R. C., Heath, A. C., & Eaves, L. J. (1992b). A population-based twin study of major depression in women. The impact of varying definitions of illness. *Archives of General Psychiatry*, *49*, 257–266.

Kendler, K. S., Neale, M. C., Prescott, C. A., Kessler, R. C., Heath, A. C., Corey, L. A., et al. (1996). Childhood parental loss and alcoholism in women: A causal analysis using a twin-family design. *Psychological Medicine*, *26*, 79–95.

Kim-Cohen, J., Moffitt, T. E., Caspi, A., & Taylor, A. (2004). Genetic and environmental processes in young children's resilience and vulnerability to socioeconomic deprivation. *Child Development*, *75*, 651–668.

Kirk, K. M., Maes, H. H., Neale, M. C., Heath, A. C., Martin, N. G., & Eaves, L. J. (1999). Frequency of church attendance in Australia and the United States: Models of family resemblance. *Twin Research*, *2*, 99–107.

Koenen, K. C., Moffitt, T. E., Caspi, A., Taylor, A., & Purcell, S. (2003). Domestic violence is associated with environmental suppression of IQ in young children. *Development and Psychopathology*, *15*, 297–311.

Kohler, H. P., & Rodgers, J. L. (2001). DF-analyses of heritability with double-entry twin data: Asymptotic standard errors and efficient estimation. *Behavior Genetics*, *31*, 179–191.

Labuda, M. C., DeFries, J. C., & Fulker, D. W. (1986). Multiple regression analysis of twin data obtained from selected samples. *Genetic Epidemiology*, *3*, 425–433.

Lake, R. E., Eaves, L. J., Maes, H.H, Heath, A. C., & Martin, N. G. (2000). Further evidence against the environmental transmission of individual differences in neuroticism from a collaborative study of 45,850 twins and relatives on two continents. *Behavior Genetics*, *30*, 223–233.

Lynch, S. K., Turkheimer, E., Emery, R. E., D'Onofrio, B. M., Mendle, J., Slutske, W., et al. (2005). A genetically informed study of the association between harsh punishment and offspring behavioral problems. Manuscript submitted for publication.

McGue, M., & Lykken, D. T. (1992). Genetic influence on risk of divorce. *Psychological Science*, *3*, 368–373.

Maes, H. M., Neale, M. C., & Eaves, L. J. (1997). Genetic and environmental factors in relative body weight and human adiposity. *Behavior Genetics*, *27*, 325–351.

Maes, H. H., Neale, M. C., Martin, N. G., Heath, A. C., & Eaves, L. J. (1999). Religious attendance and frequency of alcohol use: Same genes or same environments: A bivariate extended twin kinship model. *Twin Research*, *2*, 169–179.

Maughan, B., Taylor, A., Caspi, A., & Moffitt, T. E. (2004). Prenatal smoking and early childhood conduct problems. *Archives of General Psychiatry, 61*, 836–843.

Mendle, J., Turkheimer, E., D'Onofrio, B. M., Lynch, S. K., & Emery, R. E. (2005). Stepfather presence and age at menarche: A children of twins approach. Manuscript submitted for publication.

Meyer, J. M., Rutter, M., Silberg, J. L., Maes, H. H., Simonoff, E., Shillady, L. L., et al. (2000). Familial aggregation for conduct disorder symptomatology: The role of genes, marital discord, and family adaptability. *Psychological Medicine*, *30*, 759–774.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer.

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

Plomin, R. (1994). Genetic research and identification of environmental influences: Emanuel Miller Memorial Lecture 1993. *Journal of Child Psychology and Psychiatry, 35*, 817–834.

Plomin, R. (1995). Genetics and children's experiences in the family. *Journal of Child Psychology and Psychiatry, 37*, 695–704.

Plomin, R., & Bergeman, C. (1991). The nature of nurture: Genetic influence on "environmental" measures. *Behavioral and Brain Sciences*, *14*, 373–427.

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from each other? *Behavior and Brain Sciences*, *10*, 1–16.

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., & McGuffin, P. (2000). *Behavioral Genetics*. New York: Worth Publishers.

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

Purcell, S., & Koenen, K. (2005). Environmental mediation and the twin design. *Behavior Genetics*, *35*, 491–498.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Reiss, D., Plomin, R., & Hetherington, E. M. (1991). Genetics and psychiatry: An unheralded window on the environment. *American Journal of Psychiatry, 148*, 283–291.

Rodgers, J. L., Buster, M. A., & Rowe, D. C. (2001). Genetic and environmental influences on delinquency: DF analysis of NLSY kinship data. *Journal of Quantitative Criminology, 17*, 145.

Rodgers, J. L., & McGue, M. (1994). A simple algebraic demonstration of the validity of DeFries–Fulker analysis in unselected samples with multiple kinship levels. *Behavior Genetics*, *24*, 259–262.

Rodgers, J. L., Rowe, D. C., & Buster, M. A. (1999). Nature, nurture, and first sexual intercourse in the USA: Fitting behavioral genetic models to NLSY kinship data. *Journal of Biosocial Science*, *31*, 29–41.

Rodgers, J. L., Rowe, D. C., & Li, C (1994). Beyond nature versus nurture: DF analysis of nonshared influences on problem behaviors. *Developmental Psychology, 30*, 374–384.

Rodgers, J. L., Rowe, D. C., & May, K. (1994). DF analysis of NLSY IQ/achievement data: Nonshared environmental influences. *Intelligence*, *19*, 157–177.

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

Rowe, D. C., Almeida, D. M., & Jacobson, K. C. (1999). School context and genetic influences on aggression in adolescence. *Psychological Science*, *10*, 277–280.

Rowe, D. C., Jacobson, K. C., & van den Oord, E. (1999). Genetic and environmental influences on vocabulary IQ: Parental education level as moderator. *Child Development*, *70*, 1151–1162.

Rowe, D. C., & Waldman, I. D. (1993). The question of "How?" reconsidered. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture, & psychology* (pp. 355–374). Washington, DC: American Psychological Association.

Rutter, M. (2000). Psychosocial influences: Critiques, findings, and research needs. *Development and Psychopathology, 12*, 375–405.

Rutter, M., Dunn, J., Plomin, R., Simonoff, E., Pickles, A., Maughan, B., et al. (1997). Integrating nature and nurture: Implications of person–environment correlations and interactions for developmental psychopathology. *Development and Psychopathology, 9*, 335–364.

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

Rutter, M., & Silberg, J. (2002). Gene–environment interplay in relation to emotional and behavioral disturbance. *Annual Review of Psychology, 53*, 463–490.

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

Sexton, M., Fox, N.L, & Hebel, J. R. (1990). Prenatal exposure to tobacco: II Effects on cognitive functioning at age three. *International Journal of Epidemiology, 19*, 72–77.

Silberg, J. L., Parr, T., Neale, M. C., Rutter, M., Anglold, A., & Eaves, L. J. (2003). Maternal smoking during pregnancy and risk to boys' conduct disturbance: An examination of the causal hypothesis. *Biological Psychiatry, 53*, 130–135.

Taylor, J., McGue, M., & Iacono, W. (2000). Sex differences, assortative mating, and cultural transmission effects on adolescent delinquency: A twin family study. *Journal of Child Psychiatry and Psychology, 41*, 433–440.

Thapar, A., Fowler, T., Rice, F., Scourfield, J., van den Bree, M., Thomas, H., et al. (2003). Maternal smoking during pregnancy and attention deficit hyperactivity disorder symptoms in offspring. *American Journal of Psychiatry, 160*, 1985–1989.

Truett, K. R., Eaves, L. J., Walters, E. E., Heath, A. C., Hewitt, J. K., Meyer, J. M., et al. (1994). A model system for analysis of family resemblance in extended kinships of twins. *Behavior Genetics*, *24*, 35–49.

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science*, *9*, 160–164.

Turkheimer, E. (in press). Mobiles: A gloomy view of the future of research into complex traits. In E. Parens, N. Press, & A. Chapman (Eds.) *Wrestling with behavioral genetics: Science, ethics, and the prospect of public conversation*. Baltimore: Johns Hopkins University Press.

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

van den Oord, E. (2001). Estimating effects of latent and measured genotypes in multilevel models. *Statistical Methods in Medical Research*, *10*, 393–407.